Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization

Marco Blanchette and Benoit Chabot¹

Département de Microbiologie et d'Infectiologie, Faculté de Médecine, Université de Sherbrooke, Sherbrooke, Quebec, Canada J1H 5N4

¹Corresponding author
e-mail: b.chabot@courrier.usherb.ca

The RNA-binding protein hnRNP A1 is a splicing regulator produced by exclusion of alternative exon 7B from the A1 pre-mRNA. Each intron flanking exon 7B contains a high-affinity A1-binding site. The A1-binding elements promote exon skipping *in vivo*, activate distal 5' splice site selection *in vitro* and improve the responsiveness of pre-mRNAs to increases in the concentration of A1. Whereas the glycine-rich C-terminal domain of A1 is not required for binding, it is essential to activate the distal 5' splice site. Because A1 complexes can interact simultaneously with two A1-binding sites, we propose that an interaction between bound A1 proteins facilitates the pairing of distant splice sites. Based on the distribution of putative A1-binding sites in various pre-mRNAs, an A1-mediated change in pre-mRNA conformation may help define the borders of mammalian introns. We also identify an intron element which represses the 3' splice site of exon 7B. The activity of this element is mediated by a factor distinct from A1. Our results suggest that exon 7B skipping results from the concerted action of several intron elements that modulate splice site recognition and pairing.

Keywords: alternative splicing/RNA-binding proteins/splice site selection/UPI

Introduction

Most mammalian pre-mRNAs contain several introns that are removed precisely and efficiently to produce mature mRNAs. Despite considerable progress in understanding the mechanisms of splice site recognition in simple splicing units, little is known about how splice sites are recognized and paired in pre-mRNAs containing multiple introns and alternative splicing signals (for reviews, see Black, 1995; Chabot, 1996). Because sequences at splice junctions are not highly conserved, recent efforts have focused on the identification of elements that help select true splice sites amid an abundance of similar sequences. Elements that stimulate splicing to an adjacent splice site are known collectively as splicing enhancers, a subset of which are bound by members of the SR family of splicing factors (Lavigne *et al.*, 1993; Sun *et al.*, 1993; Staknis and Reed, 1994; Ramchatesingh *et al.*, 1995; Achsel and Shimura, 1996; Gallego *et al.*, 1997). The binding of SR proteins to an exon enhancer can stimulate U2AF⁶⁵ binding to a weak upstream 3' splice site (Wang *et al.*, 1995).

Enhancer elements that affect 5' splice site selection may also improve U1 snRNP binding to nearby 5' splice sites (Eperon *et al.*, 1993; Kohtz *et al.*, 1994; Humphrey *et al.*, 1995). In some cases, the assembly of a stable enhancer complex requires additional cell-specific factors (Tian and Maniatis, 1992; Heinrichs and Baker, 1995; Lynch and Maniatis, 1995) or the participation of non-SR proteins (Min *et al.*, 1997). Elements that repress the use of splice sites have also been identified (Caputi *et al.*, 1994; Amendt *et al.*, 1995; Siebel *et al.*, 1995; Staffa and Cochrane, 1995; Del Gatto *et al.*, 1996; Kanopka *et al.*, 1996; Blanchette and Chabot, 1997; Chan and Black, 1997a; Si *et al.*, 1997). In *Drosophila*, the binding of Sxl to 3' splice site sequences interferes with U2AF binding (Valcarcel *et al.*, 1993). In mammals, the binding of polypyrimidine tract-binding protein (PTB) to 3' splice site sequences has also been implicated in splicing repression by competing for U2AF-binding sites (Lin and Patton, 1995; Singh *et al.*, 1995; Ashiya and Grabowski, 1997; Chan and Black, 1997b). Although a variety of different types of exon and intron elements can modulate the use of nearby splice sites, it remains unclear whether the majority of these elements affect splice site recognition, commitment between a pair of splice sites, or some other events of spliceosome assembly.

Following the recognition of individual splice sites, appropriate partners must find each other. This commitment step occurs through a network of interactions across the intron in which SR proteins and/or SF1 mediate communication between factors bound at splicing signals (Fu, 1993; Wu and Maniatis, 1993; Zuo and Maniatis, 1996; Abovich and Rosbash, 1997). The frequency with which splice site partners find each other is dependent on the distance separating a pair of splice sites. Possibly to increase the rate of commitment, the spatial distance separating splice sites in several yeast introns is reduced through base pairing interactions between sequences downstream from the 5' splice site and sequences upstream from the branch site (Parker and Patterson, 1987; Libri *et al.*, 1995; Charpentier and Rosbash, 1996; Howe and Ares, 1997). How efficient splice site pairing is achieved in some of the very large mammalian introns is unknown.

Heterogeneous nuclear ribonucleoprotein (hnRNP) A/B proteins counteract the effect of SR proteins by promoting distal 5' splice site utilization *in vitro* and *in vivo* (Mayeda and Krainer, 1992; Cáceres *et al.*, 1994; Mayeda *et al.*, 1994; Yang *et al.*, 1994). We have shown previously that hnRNP A1 modulates 5' splice site selection on an A1 model pre-mRNA by binding to a conserved intron sequence (CE1a) located in between competing 5' splice sites (Chabot *et al.*, 1997). In contrast to SR proteins, which increase U1 small nuclear (sn)RNP binding to both 5' splice sites, the A1–CE1a interaction does not affect U1 snRNP binding to competing 5' splice sites, suggesting

Fig. 1. Schematic representation of the hnRNP A1 alternative splicing unit. The size and position of the CE1 and CE4 elements are shown (top). An alignment of the portion of the mouse and human sequences from exon 7B to the *StuI* site in the mouse intron is shown below. The sequence of CE4 is underlined. The region between the flag and the *StuI* site has been deleted in the pmA1ΔCE4 construct (see Figure 2).

that A1 and SR proteins target different events in splice site selection. Here we document the presence of an additional A1-binding site located in the CE4 element downstream of alternative exon 7B. CE4, like CE1a, promotes exon 7B skipping *in vivo* and distal 5' splice site selection *in vitro*. The ability of A1 complexes to interact simultaneously with two high-affinity binding sites supports the notion that an interaction between bound A1 molecules brings in close proximity the 5' splice site of exon 7 and the 3' splice site of exon 8 to favor exclusion of exon 7B. Another sequence element within CE4 can promote distal 3' splice site selection by repressing the 3' splice site of exon 7B in a process that does not require A1. Our studies suggest that efficient exclusion of exon 7B requires the participation of distinct intron elements which modulate pre-mRNA conformation and splice site utilization.

Results

CE4 modulates exon 7B skipping *in vivo*

Previously, we reported the identification of a 150 nucleotide element (CE1) which promotes distal 5' splice site utilization *in vitro* and exon 7B skipping *in vivo* (Chabot *et al.*, 1997). CE1 is located in the intron upstream of alternative exon 7B (Figure 1), and a minimal portion of CE1 (CE1a; 17 nucleotides) can activate distal 5' splice site selection *in vitro* (Chabot *et al.*, 1997). The activity of CE1a is likely to be mediated through an interaction with the hnRNP A1 protein since a mutation that compromises A1 binding to CE1a prevents activation of the distal 5' splice site *in vitro* (Chabot *et al.*, 1997). Visual inspection of the mouse sequence downstream of exon 7B revealed a putative A1-binding site identical to the one found in CE1a (UAGAGU). This sequence is part of a conserved 24 nucleotide region that we named CE4 (Figure 1). HnRNP A1 binds to CE4 *in vitro* (S13 RNA in Chabot *et al.*, 1997). To verify whether CE4 also modulates splicing of the A1 pre-mRNA *in vivo*, we transiently expressed in HeLa cells a genomic portion of the mouse hnRNP A1 alternative splicing unit under the control of

the cytomegalovirus (CMV) promoter (Figure 2A). RT-PCR analysis of total RNA indicated that expression of the wild-type A1 minigene preferentially yielded transcripts lacking exon 7B (Figure 2B, lanes 4 and 12). This splicing profile reproduces the levels of endogenous A1 and A1^B mRNAs (data not shown). As shown previously (Chabot *et al.*, 1997), deletion of CE1 improved the relative frequency of exon 7B inclusion (Figure 2B, lane 5). Reinsertion of the minimal 17 nucleotide CE1a element nearly restored the wild-type profile of exon skipping (lane 8). Likewise, the deletion of a 113 nucleotide region containing CE4 promoted exon 7B inclusion (A1ΔCE4; Figure 2B, lanes 6 and 13). While reinsertion of an unrelated sequence had no effect (data not shown), insertion of CE4 (24 nucleotides) restored efficient exon 7B skipping (A1RCE4; lane 14). Notably, the deletion of CE4 was as efficient as the deletion of both CE1 and CE4 (Figure 2B, A1ΔΔ; lane 7, see Figure 2C for quantitation). These results indicate that CE4 contains sequences that promote exon 7B skipping.

CE4 promotes distal 5' splice site selection *in vitro*

The activity of CE4 on 5' splice site selection *in vitro* was investigated using model transcripts that contain the 5' splice sites of exon 7 and 7B in competition for the 3' splice site of the adenovirus L2 exon (Figure 3A). The basic model pre-mRNA lacks CE1a and CE4 and is spliced almost exclusively to the proximal 5' splice site (C5' -/-; Figure 3B, lane 1). Insertion of CE1a and CE4 at the upstream and downstream position, respectively, stimulated splicing to the distal 5' splice site (C5' 1a/4; lane 6). Transcripts containing CE4 or CE1a at both positions were spliced to the distal 5' splice site with an efficiency similar to C5' 1a/4 (C5' 1a/1a and C5' 4/4; lanes 7 and 8, respectively). Substrates carrying CE1a or CE4 at the upstream or downstream position only were spliced to the distal 5' splice site at an intermediate but comparable level (Figure 3B, C5' 1a/-, C5' 4/-, C5' -/4 and C5' -/1a; lanes 2-5; see Figure 3A for quantitation). Thus, CE1a and CE4 are interchangeable and display a similar ability to promote distal 5' splice site utilization

Fig. 2. Modulation of alternative splicing by CE1a and CE4 *in vivo*. **(A)** Schematic representation of the A1 minigene and derivatives. The positions of oligonucleotides E1a-569 and A1E9 used for RT-PCR assays in **(B)** are shown. The position of CE1 and CE4 is indicated. **(B)** RT-PCR assays on total RNA were performed to measure the efficiency of exon 7B inclusion following transfection into HeLa cells. The A1c and A1^Bc lanes correspond to PCR assays performed directly on minigenes carrying the cDNA of A1 and A1^B, respectively. **(C)** Autoradiograms were scanned and the intensities of the bands were measured by densitometry in at least five independent transfection assays. Changes in the A1^B/A1 ratio of amplified products are indicated by histograms with the standard deviation. Values are expressed relative to the A1^B/A1 ratio obtained with the wild-type A1 minigene which has been given the arbitrary value of 1.

in vitro. These results suggest that CE1a and CE4 modulate 5' splice site choice through a common mechanism.

The effect of CE4 on 5' splice site selection is mediated by hnRNP A1

We have shown that a mutation in the UAGAGU sequence of CE1a reduces A1 binding and prevents activation of the distal 5' splice site (Chabot *et al.*, 1997). If the binding of A1 is also responsible for the effect of CE4 on 5' splice site selection, pre-mRNA molecules carrying CE4 should be more responsive to increases in the concentration of A1 than a pre-mRNA lacking high-affinity A1-binding sites. When increasing amounts of recombinant GST-A1 protein (rA1) were added to a splicing mixture containing C5' 4/4 RNA, selection of the proximal 5' splice site gradually decreased until the distal 5' splice site became the exclusive choice (Figure 4A, lanes 5–8). In contrast, the largest amount of rA1 only promoted a modest reduction in proximal 5' splice site utilization on a pre-mRNA lacking strong A1-binding sites (lanes 1–4). An intermediate effect was seen with pre-mRNAs containing only one copy of CE1a or CE4 (data not shown). Thus, pre-mRNAs carrying high-affinity A1-binding sites are

more sensitive to an increase in the concentration of A1 *in vitro*.

To test the effect of a reduction in the concentration of A1 on 5' splice site selection *in vitro*, we added increasing amounts of an oligonucleotide carrying the DNA version of three contiguous A1-binding sites. A1 binds to this oligonucleotide directly and with specificity (F.Dallaire and B.Chabot, unpublished results). Pre-incubating splicing mixtures with the oligonucleotide shifted selection toward the proximal 5' splice site of C5' 4/4 RNA (Figure 4B, lanes 8–11). Supplementing the mixture containing the highest concentration of oligonucleotide with rA1 restored preferential distal 5' splice site utilization (lanes 12–14). As expected for a pre-mRNA lacking high-affinity A1-binding sites (C5' –/– RNA), the already efficient splicing to the proximal 5' splice site was not improved further by the addition of increasing amounts of oligonucleotide, and the addition of rA1 did not stimulate distal 5' splice site selection (lanes 1–7). The oligonucleotide also increased proximal 5' splice site selection in mixtures containing C5' 4/– RNA (data not shown). However, lower concentrations of oligonucleotide were required to obtain a shift similar to the one obtained with

Fig. 3. CE1a and CE4 are interchangeable and display a similar activity in 5' splice site selection. (A) Schematic representation of the basic model pre-mRNA containing competing 5' splice sites. The identity of the elements inserted downstream of each 5' splice site is indicated below. The percentage of distal products relative to the sum of proximal and distal products was obtained by densitometric scanning of the autoradiograms shown in (B). (B) Labeled pre-mRNAs were incubated in HeLa nuclear extracts and splicing products were run on a 10.25% acrylamide-8 M urea gel. The position of distal and proximal lariat molecules is indicated. Note that the proximal lariat introns derived from C5' -/-, C5' 1a/- and C5' 4/- migrate below the pre-mRNA. For other substrates, proximal lariat molecules migrate above the pre-mRNAs. The origin of the gel is indicated (ori).

C5' 4/4 RNA. Thus, a pre-mRNA carrying two high-affinity A1-binding sites is more resistant to a reduction in the effective concentration of A1 than a pre-mRNA containing only one optimal A1-binding site. Overall, these results demonstrate that the effect of CE4 on 5' splice site selection is mediated through an interaction with the hnRNP A1 protein.

Importance of the glycine-rich domain

The N-terminal portion of A1 (called UP1) contains two RNA-binding domains but lacks the C-terminal glycine-rich domain (Figure 5). UP1 binds inefficiently to a human β -globin pre-mRNA and, in contrast to A1, does not affect 5' splice site selection when added to an S100 extract supplemented with ASF/SF2 (Mayeda *et al.*, 1994). To determine whether UP1 displayed a similar behavior on transcripts carrying high-affinity A1-binding sites, we first monitored the binding of recombinant A1 and UP1 proteins to CE4. rUP1 bound to naked CE4 RNA at least as efficiently as rA1, and both proteins bound inefficiently to the complementary sequence of CE4 (Figure 5A). Thus, the absence of the glycine-rich domain of A1 does not abrogate its capacity to interact with CE4. When increasing amounts of rUP1 were added in nuclear extracts incubated with the C5' 4/4 pre-mRNA, 5' splice site selection shifted from predominantly distal to primarily proximal (Figure 5B). The presence of high-affinity A1-binding sites and endogenous A1 proteins therefore allowed UP1 to display

a dominant-negative effect on 5' splice site selection. Thus, although not essential for binding to CE4, the glycine-rich domain plays a crucial role in promoting distal 5' splice site selection.

Bridging activity of A1 proteins

We have shown previously that the effect of CE1 on 5' splice site selection is not associated with differences in U1 snRNP binding to competing 5' splice sites (Chabot *et al.*, 1997). Identical results were obtained with CE4 (data not shown). Moreover, the important shifts in 5' splice site selection obtained by the addition of rA1 or by pre-incubating splicing mixtures with the DNA oligonucleotide carrying A1-binding sites did not affect U1 snRNP binding to competing 5' splice sites (data not shown). To explain the effect of CE1a and CE4 on 5' splice site selection, we propose that the binding of A1 to CE1a and CE4 is followed by protein contacts between A1 molecules bound to these elements (Figure 6A). This interaction would bring the more distant pair of splice sites into close proximity, thereby improving their rate of commitment. This model implies that A1 affects 5' splice site selection not by altering 5' splice site recognition but rather by changing the conformation of the pre-mRNA.

To address whether A1 molecules can interact simultaneously with two RNA-binding elements, we first incubated rA1 or rUP1 with 32 P-labeled CE1a RNA and a control K+ RNA. The mixture was then incubated with

Fig. 4. A pre-mRNA carrying CE4 elements is more sensitive to an increase in the concentration of A1. **(A)** C5' $-/-$ and C5' 4/4 RNAs were incubated in HeLa extracts containing purified recombinant GST-A1 (rA1). Lanes 1–4 and 5–8 contain 0, 0.25, 0.5 and 1.0 μ g of rA1, respectively. **(B)** Distal 5' splice site selection in the presence of increasing amounts of an oligonucleotide bound by A1. Increasing amounts of (TAGGGT)₃ oligonucleotide were pre-incubated for 10 min at 30°C in splicing mixtures. Lanes 1–5, 6–10, 11–15 and 16–20 contain 0, 1, 2, 4 and 8 pmol of oligonucleotide. Labeled C5' $-/-$ and C5' 4/4 RNAs were added and incubation was continued for 2 h. Splicing products in (A) and (B) were loaded onto an 11% acrylamide–8 M urea gel.

cold CE1a RNA covalently coupled to an adipic acid hydrazide column. Each column was washed at increasing salt concentration to elute bound complexes. As shown in Figure 6B, the rA1–[³²P]CE1a complex was specifically retained by the column, labeled CE1a RNA co-eluting with rA1 (Figure 6B, middle panels). In contrast, the rUP1–[³²P]CE1a complex was not retained by the column and only free UP1 eluted at higher salt concentration (Figure 6B, right panels). The specificity of the assay is demonstrated by the fact that the control K⁺ RNA did not co-elute with rA1, and that the RNA column did not retain [³²P]CE1a RNA in the absence of rA1 (left panel). These results indicate that A1 molecules bound to one site can interact specifically with another high-affinity binding site, and that the glycine-rich domain of A1 is essential for simultaneous binding. Because the glycine-rich domain is not required for efficient and specific binding to the RNA (Figure 6A), our results suggest that the glycine-rich domain mediates an interaction between A1 molecules independently bound to high-affinity sites.

CE4 also affects 3' splice site selection

While the effects of CE4 and CE1a on 5' splice site selection were comparable *in vitro*, the removal of CE4 had more impact on the inclusion of exon 7B *in vivo* than the deletion of CE1 (Figure 2). This result suggests that CE4 modulates an additional event in splice site selection. To test whether CE4 could influence 3' splice site selection, we used a model pre-mRNA that contains the 5' splice

site of exon 7 and two competing 3' splice sites (C3' $-/-$; Figure 7A). While C3' $-/-$ RNA was spliced predominantly to the distal 3' splice site, the proximal 3' splice site was also used (Figure 7A, lane 1; and see Figure 8). Notably, insertion of CE4 in between the two 3' splice sites eliminated proximal 3' splice site selection and increased distal splicing (C3' $-/4$; Figure 7A, lane 3). In contrast, insertion of CE1a at the same position, or insertion of CE4 upstream of the proximal 3' splice site, had no effect (C3' $-/1a$ and C3' 4 $-$; lanes 4 and 2, respectively). To address whether CE4 repressed the proximal 3' splice site, we tested a derivative of C3' $-/4$ which lacked the distal 3' splice site (S774; Figure 7B). Compared with a control pre-mRNA carrying the complementary sequence of CE4 at the same position (S774 α), the presence of CE4 was associated with a strong inhibition of splicing (Figure 7B, lanes 1–8). The presence of CE4 in the intron did not inhibit splicing (S747; Figure 7B, lanes 9–12). Because splicing of S747 was not perturbed when S774 RNA was included in the splicing mixture (lanes 13–16), inefficient splicing of S774 RNA was not caused by a general inhibitor in the S774 RNA preparation. Thus, while both CE4 and CE1a modulate 5' splice site selection, only CE4 can repress the 3' splice site of exon 7B in a position-dependent manner. The ability of CE4 to influence both 5' and 3' splice site selection probably explains why the deletion of CE4 promoted more efficient exon 7B inclusion than did the deletion of CE1 (Figure 2).

Fig. 5. Importance of the glycine-rich domain of A1 in 5' splice site selection. A schematic representation of the GST-A1 (rA1) and GST-UP1 (rUP1) proteins is shown at the top. (A) rA1 and rUP1 binding to naked RNA. Small labeled RNA molecules carrying the CE4 element (CE4 RNA) or its complementary sequences (4EC RNA) were incubated in the presence of rA1 or rUP1 bound to glutathione-Sepharose 4B, as described in Chabot *et al.* (1997). The efficiency of RNA recovery is expressed relative to the c.p.m. recovered with the rA1-CE4 RNA combination (absolute value of 100). (B) rUP1 inhibits activation of the distal 5' splice site. C5' 4/4 RNA was incubated in nuclear extracts supplemented with increasing amounts of recombinant rUP1. The lowest concentration of rUP1 tested (1 μg) corresponds to the highest concentration of rA1 tested in Figure 4A. The position of proximal and distal lariat molecules is indicated.

The effect of CE4 on 3' splice site choice is mediated by a trans-acting factor distinct from A1

Because CE1a does not affect 3' splice site selection, A1 binding to CE4 cannot by itself account for the effect of CE4 on 3' splice site usage. Nevertheless, A1 could play a role in repressing the upstream 3' splice site. To address this issue, we tested a CE4 derivative lacking the A1-binding site (CE4m) (Figure 8A). While CE4m was bound considerably less efficiently than CE4 by A1 (Figure 8B), CE4m was as efficient as CE4 at improving selection of the distal 3' splice site (C3' -/4 and C3' -/4m; Figure 8A, lanes 3 and 4, respectively). Insertion of the complementary sequence of CE4m had no effect (C3' -/4mα, lane 2). At concentrations that influence 5' splice site selection, the addition of rA1 did not affect the splicing of either C3' -/- or C3' -/4 (data not shown). These results indicate that A1 binding is dispensable for the effect of CE4 on 3' splice site choice. To determine whether the activity of CE4m in 3' splice site selection requires a *trans*-acting factor, we performed splicing assays in the presence of molar excesses of competitor RNAs. The highest concentration of CE4m RNA shifted 3' splice site selection on the C3' -/4m pre-mRNA from exclusively distal to predominantly proximal (Figure 8C, lanes 13–16). The reduction in splicing efficiency suggests that the excess of CE4m RNA titrates out a constitutive splicing factor. While a similar concentration of CE4m RNA also improved the frequency of proximal 3' splice

site selection on C3' -/- pre-mRNA (Figure 8B, lanes 5–8), the amplitude of the shift was less important than on C3' -/4m RNA. An RNA fragment containing plasmid sequences (K+ RNA), although improving distal 3' splice site selection at low concentrations, did not significantly affect the ratio of 3' splice site utilization at the highest concentration (Figure 8C, lanes 1–4 and 9–12). Our results indicate that the effect of CE4 on 3' splice site selection is mediated by a factor distinct from hnRNP A1 and that nearby binding of A1 is not required for this activity.

Discussion

Modulation of alternative splicing by hnRNP A1

We report the identification and activity of a 24 nucleotide conserved element (CE4) located downstream of alternative exon 7B. CE4 promotes exon 7B skipping *in vivo*, and influences both 5' and 3' splice site selection *in vitro*. The effect of CE4 on 5' splice site choice is indistinguishable from that of the previously characterized CE1a intron element located downstream of exon 7 and bound by hnRNP A1 (Chabot *et al.*, 1997). As evidence for the contribution of hnRNP A1 to the activity of CE4, we have shown that CE4-containing pre-mRNAs are more responsive to an increase in the concentration of A1 than is a pre-mRNA lacking CE4. Despite the fact that recombinant A1 can completely counteract the shift in 5' splice site selection caused by an excess of A1-binding

Fig. 6. A1 bridging activity. (A) A model for the role of A1 in alternative splicing. In the hnRNP A1 pre-mRNA, an interaction between A1 molecules bound to CE1a and CE4 would bring the 5' splice site of exon 7 into closer proximity to the 3' splice site of exon 8, increasing the frequency of exon 7B skipping. Bars indicate the distance separating the distal pair of splice sites. (B) Simultaneous interaction of A1 molecules with two high-affinity binding sites. Cold CE1a RNA was covalently linked to an agarose adipic acid hydrazide resin. A mixture of ³²P-labeled RNAs (CE1a and K⁺) alone or with rA1 or rUP1 were loaded onto the column. The 0.05 M KCl fractions correspond to the flow-through. The column was eluted successively with 0.1, 0.25 and 0.5 M KCl. The recovered RNAs and proteins were fractionated on an 8% acrylamide-8 M urea gel and a 12.5% SDS-polyacrylamide gel, respectively (top and bottom panels, respectively). The input and 0.05 M lanes of the RNA gel contain one-tenth of the initial fractions.

substrate, other factors may be recruited by A1 to modulate 5' splice site selection. The binding of A1 to our model pre-mRNAs is not associated with the assembly of core hnRNP particles that contain the hnRNP C protein (data not shown). Whereas formation of an (A1)₃B2 tetramer remains to be verified, the small size of the pre-mRNAs tested *in vitro* may not be conducive to hnRNP core particle formation since a minimum of 700 nucleotides is required for the assembly of a core 40S hnRNP particle (Conway *et al.*, 1988). Thus, whereas it is possible that the binding of A1 to the natural A1 pre-mRNA promotes the formation of a complex containing additional hnRNP proteins, the participation of hnRNP C in this process does not appear essential *in vitro*.

The mechanism by which A1-binding elements modu-

late 5' splice site selection and exon skipping remains unclear. The presence of CE1 and CE4 is not associated with differences in the binding of U1 snRNP to competing 5' splice sites (Chabot *et al.*, 1997; data not shown), suggesting that these elements do not directly influence 5' splice site recognition. Moreover, the inclusion of CE1 or CE4 in the intron of simple one-intron pre-mRNAs does not affect splicing efficiency (Figure 6B and data not shown), indicating that A1-binding elements do not affect spliceosome assembly. On the other hand, we showed that a derivative of A1 lacking the C-terminal glycine-rich domain (UP1) antagonizes distal 5' splice site selection in a nuclear extract, most likely by competing with endogenous A1 for binding to CE4. This indicates that the glycine-rich domain of A1 performs an essential

Fig. 7. CE4 promotes distal 3' splice site selection. **(A)** *In vitro* splicing assays performed in HeLa nuclear extract with pre-mRNAs containing competing 3' splice sites. Labeled splicing products were fractionated on an 11% acrylamide–8 M urea gel. **(B)** CE4 inhibits upstream intron removal. Labeled pre-mRNAs containing a single pair of splice sites were incubated in a HeLa extract. Aliquots of each splicing reaction were taken at the indicated time and run on an 11% acrylamide–8 M urea gel. A mixture containing both S747 and S774 RNAs was analyzed to rule out the presence of an inhibitor in the S747 transcript preparation (lanes 13–16). A schematic representation of the model pre-mRNAs is presented on top of each panel.

function in 5' splice site selection. The glycine-rich domain is responsible for the self-associating property of A1 and its RNA strand-annealing activity (Kumar and Wilson, 1990; Pontius and Berg, 1990; Munroe and Dong, 1992; Casas-Finet *et al.*, 1993; Portman and Dreyfuss, 1994; Cartegni *et al.*, 1996). We have shown that A1 molecules can interact simultaneously with two high-affinity sites. Because this behavior is not seen with UP1, the glycine-rich domain probably mediates an interaction between A1 molecules bound to distinct sites. Based on these observations, we propose a simple model for the mechanism of action of CE1a and CE4. As depicted in Figure 6A, A1 proteins bound to CE1 and CE4 would interact with one another through their glycine-rich domain. This interaction would change the conformation of the pre-mRNA and increase the frequency of commitment between the U1-bound distal 5' splice site and the 3' splice site of exon 8.

In its simplest version, the model predicts that appropriately positioned A1-binding sites should also affect 3' splice site selection. However, CE1a has no effect on 3' splice site choice (Chabot *et al.*, 1997), and concentrations of A1 that affect 5' splice site selection *in vitro* do not influence 3' splice site choice (Mayeda *et al.*, 1994; data not shown). These results suggest that a 3' splice site positioned in between two strong A1-binding sites hinders the interaction between bound A1 molecules. Consistent with this view, the inclusion of an optimal U2AF⁶⁵-binding site in between two CE1a elements seriously compromises activation of the distal 5' splice site (unpublished results).

The details of the mechanism that leads to this antagonism are currently being investigated.

The above model implies that activation of the distal 5' splice site requires an interaction between A1 molecules bound on each side of the proximal 5' splice site. Given that A1 can bind to a variety of RNA sequences with lower affinity *in vitro* (Burd and Dreyfuss, 1994; Mayeda *et al.*, 1994; Abdul-Manan and Williams, 1996), high concentrations of A1 could lead to appropriately positioned A1 proteins on a pre-mRNA that lacks optimal A1-binding sites. The concentration of A1 in nuclear extracts is probably high enough to allow less discriminate A1 binding because inclusion of only one optimal A1-binding site can stimulate distal 5' splice site selection, albeit less efficiently than when two appropriately positioned A1-binding sites are present. Moreover, increasing the number of A1-binding sites reduces the level of exogenous A1 required to achieve preferential distal 5' splice site utilization. Whether A1 only binds to high-affinity A1-binding sites in nascent pre-mRNAs is unknown. The high abundance of core hnRNP proteins, their association with pre-mRNAs extracted from nuclei and the visualization of hnRNP complexes in active genes has fueled the notion that major hnRNP proteins associate indiscriminately with pre-mRNAs *in vivo* (Beyer and Osheim, 1988; Amero *et al.*, 1992; Matunis *et al.*, 1992; Zu *et al.*, 1998; for a review see McAfee *et al.*, 1997). However, when the distribution of individual hnRNP proteins has been looked at, their position on nascent pre-mRNAs follows a non-random distribution (Pinol-Roma *et al.*, 1989; Matunis

Fig. 8. The activity of CE4 in 3' splice site selection does not require A1. (A) Labeled pre-mRNA substrates carrying CE4 or CE4m (a shortened version of CE4 lacking the A1-binding site) were incubated in a HeLa extract. A schematic representation of the substrates and the inserted sequence is shown. (B) Binding of A1 to CE4 or CE4m. Radiolabeled RNAs were incubated with increasing concentrations of rA1. Each set of gel shifts was performed with 0, 0.016, 0.063, 0.25 and 1.0 μ g of rA1. (C) Inhibition of proximal 3' splice site utilization by CE4m requires a *trans*-acting factor. Splicing of C3' -/- and C3' -/4m RNAs was performed in a HeLa extract pre-incubated with increasing amounts of the competitor CE4m RNA (lanes 5–8 and 13–16) or with a control RNA (lanes 1–4 and 9–12). Each set of competition experiments was performed with 0, 1, 2 and 4 pmol of unlabeled competitor RNA. Splicing products were fractionated on an 11% acrylamide–8 M urea gel.

et al., 1993; Visa *et al.*, 1996; Wurtz *et al.*, 1996), an observation that may reflect the binding specificity of each individual hnRNP protein. Our study shows that the presence of high-affinity A1-binding sites has functional implications in pre-mRNA splicing both *in vitro* and *in vivo*. Thus, while the combined interactions of a constellation of hnRNP proteins may package nascent pre-mRNAs in a form that is primed to undergo many types of maturation events, interactions between individual hnRNP proteins bound at specific sites may modulate pre-mRNA conformation and, in the case of hnRNP A1, promote splice site pairing.

A role for hnRNP A1 in intron definition

A key feature of the model proposed is that the interaction between bound A1 proteins brings into close proximity sequences or factors flanking A1-binding sites. While this interaction has important consequences in splice site selection, it may also be relevant to the constitutive splicing of introns which require commitment between partners often separated by several kilobases. To address the possible role of A1-binding sites in enforcing commit-

ment between distant splice junctions, we examined the distribution of putative A1-binding sites in exon and intron sequences of six genes (*BRCA2*, *RB1*, *NF1*, *c-myc*, troponin T and *abl*) containing a total of 119 introns spanning 501 145 bp with a mean intron length of 4211 bp. The consensus A1-binding site TAG^A/G^T/A represents a combination between the 'winner' sequence determined by Burd and Dreyfuss (1994) and the A1-binding site found in CE1a and CE4. If A1-binding sites are distributed randomly, the theoretical frequency should average one A1-binding site per 1024 bp. The first striking observation concerns the imbalance in the frequency of optimal A1-binding sites between exon and intron sequences. cDNAs contain a low frequency of A1-binding sites (0.32 site per 1000 bp; Table I). This frequency is not biased by the small number of exon sequences analyzed since a similar frequency is observed in an additional 23 822 bp taken from randomly selected cDNAs (0.29 site per 1000 bp; Table I). While the frequency of the UAG triplet is expected to be lower in coding sequences, the prevalence is too low to account for the paucity of A1-binding sites in all reading frames. In contrast, A1-binding sites are

abundant in introns, the overall frequency indicating a prevalence slightly superior to the theoretical frequency (Table I). When the distribution of A1-binding sites is plotted for each gene, a predominance of A1-binding sites near splice sites is noted in several introns (Figure 9A). A compilation of the distribution of A1-binding sites for introns carrying three or more sites (26% of all introns sampled) shows the high prevalence of A1-binding sites near 5' and 3' splice sites (Figure 9B). While this frequency decreases gradually in the internal intron segments, the central portion displays a moderately high occurrence of A1-binding sites. The high prevalence of G triplets near the 5' splice sites of primate introns has been reported previously (Nussinov, 1988, 1989; Engelbrecht *et al.*, 1992; Solov'yev *et al.*, 1994). At least in small introns, G triplets have been shown to enhance splicing efficiency, to affect 5' splice site selection and to play a role in fixing exon-intron borders (McCullough and Berget, 1997). As hnRNP A1 may bind to at least a subset of these G-rich sequences, interactions between A1 molecules bound at different sites may be responsible for the effects observed.

We propose that the prevalence of A1-binding sites in several introns, particularly near 5' and 3' splice junctions, plays a crucial role in intron definition. Interactions between A1 proteins bound in one intron would loop out segments of this intron until distant splicing partners are in close proximity (Figure 9C). This process ultimately resembles the situation in several yeast introns in which the formation of an RNA duplex structure between sequences near a 5' splice site and a branch site enhances splicing efficiency (Newman, 1987; Libri *et al.*, 1995; Charpentier and Rosbash, 1996; Howe and Ares, 1997). Duplex structures can also promote exon or splice site skipping when introduced into mammalian pre-mRNAs (Solnick, 1985a; Eperon *et al.*, 1986). However, because duplex formation between distant but perfectly complementary sequences does not occur efficiently in mammalian cells (Solnick and Lee, 1987; Eperon *et al.*, 1988), the task of bringing distant intron junctions into close proximity in higher eukaryotes may have been taken over by proteins such as hnRNP A1. Given that an optimal U2AF-binding site located in between two A1-binding sites antagonizes the activity of these elements in 5' splice site selection, interactions between A1 molecules bound

in adjacent introns may be prevented by the presence of a strong 3' splice site, thus restricting A1-bridging interactions to within individual introns. In the presence of a suboptimal 3' splice site, the presence of A1-binding sites in flanking introns may set the stage for exon skipping or alternative 3' splice site selection.

A variety of elements contribute to exon 7B alternative splicing

The efficiency of A1-mediated exon 7B skipping will vary depending on the rate of commitment between proximal pairs of splice sites. In other words, CE1a- and CE4-mediated commitment between the distal pair of splice sites (7–8) must occur before each proximal pair (7–7B and 7B–8) is committed. Possibly to delay commitment between the 7B–8 pair, U1 snRNP binding to the 5' splice site of exon 7B is compromised by the formation of a highly stable secondary structure with the downstream CE6 element (Blanchette and Chabot, 1997). A different mechanism operates to prevent commitment between the upstream pair of splice sites (7–7B). In this case, a sequence element within CE4 (CE4m) bound by a factor that remains to be identified represses 3' splice site utilization. Because an optimal U2AF⁶⁵-binding site interferes with the activity of A1-binding elements (unpublished results), repression of 3' splice site recognition may be essential to permit an interaction between A1 proteins bound in neighboring introns.

Elements that repress 3' splice site utilization have been localized on either side of a 3' splice junction. CE4m belongs to a group of negative elements that repress 3' splice site utilization from a downstream position (Caputi *et al.*, 1994; Amendt *et al.*, 1995; Del Gatto and Breathnach, 1995; Staffa and Cochrane, 1995; Del Gatto *et al.*, 1996; Zheng *et al.*, 1996; Si *et al.*, 1997). A portion of CE4m (5'-AGCUAGAUUAGAUCU-3') is almost identical to the silencer element 5'-CUAGACUAGA-3' present in *tat* exon 2 of human immunodeficiency virus (HIV) (Si *et al.*, 1997), and is similar to a related negative element in the HIV *tat-rev* exon 3 (Amendt *et al.*, 1995; Staffa and Cochrane, 1995). Given the sequence similarity between CE4m and the silencer element in *tat* exon 2 of HIV, their activity is likely to be modulated by a common factor. In contrast to a recent study on the *tat* exon 2 of HIV (Del Gatto-Konczak *et al.*, 1999), our results indicate that hnRNP A1 is not implicated in the activity of CE4m; CE4m is not bound efficiently by hnRNP A1, and substituting CE4m for the A1-binding element CE1a does not affect 3' splice site selection. Experiments are in progress to identify the factor which binds to CE4m, and to determine whether CE4m represses the interaction of constitutive splicing factors at the 3' splice site of exon 7B.

In conclusion, processing of the A1 pre-mRNA to produce mRNAs predominantly lacking exon 7B is controlled by the combined action of at least four distinct intron elements (Figure 10). Elements bound by A1 are

Table I. Distribution of A1-binding sites in some mammalian genes

Sequence	Length (bp)	Sites found (No.)	Frequency (x/1000 bp)
Introns ^a	501 145	607	1.21
Exons ^a	25 276	9	0.36
cDNAs ^b	23 822	7	0.29

^aData extracted from *BRCA2*, *NF1*, *c-myc*, *abl*, *RB1* and cTnT genomic sequence.

^bData extracted from 10 other randomly selected mammalian cDNAs.

Fig. 9. Distribution of putative hnRNP A1-binding sites in mammalian pre-mRNAs. (A) Boxes and lines represent exons and introns, respectively. The position of each putative A1-binding site (TAG^A/G^T/A) is indicated below each gene by a thin vertical bar. (B) Compilation of the occurrence of A1-binding sites in introns carrying three or more A1-binding sites. Each histogram indicates the frequency of A1-binding sites in successive 10% portions of the normalized intron. A random distribution was given the arbitrary value of 1. (C) Interactions between bound A1 molecules would bring distant splicing partners into close proximity to help define intron borders. Boxes, thin lines and vertical bars represent exons, introns and putative A1-binding sites, respectively. A1 proteins (filled circles) bound to intron sequences would loop out a large fraction of the intron to bring distant splicing partners into closer proximity.

A

B

C

Fig. 10. Exon 7B skipping requires multiple elements and factors. Boxes and thin lines represent exons and introns, respectively. An interaction between A1 molecules bound to CE1a and CE4 would bring the 5' splice site of exon 7 into closer proximity to the 3' splice site of exon 8, increasing the frequency of exon 7B skipping. Recognition of the 5' splice site of exon 7B is compromised by a secondary structure with the CE6 element (Blanchette and Chabot, 1997), while the 3' splice site of exon 7B is repressed by a factor bound to CE4m.

located on each side of exon 7B, and probably foster a change in pre-mRNA conformation to increase commitment between the 5' splice site of exon 7 and the 3' splice site of exon 8. Possibly to increase the efficiency of this event, the recognition of splice sites flanking exon 7B is undermined by CE6 and CE4, which repress 5' and 3' splice site utilization, respectively. Examining the participation of other conserved intron and exon elements in the control of exon 7B skipping will continue to reveal important insights into the complex mechanisms that influence splice site selection in a simple alternative splicing unit.

Materials and methods

Plasmid constructs

The expression plasmid pmA1 which carries the A1 minigene (exons 5–10), and its derivative carrying a deletion of CE1 (pmA1ΔCE1), have been described previously (Blanchette and Chabot, 1997). To construct pmA1ΔCE4, a 125 bp fragment containing the second half of exon 7B and 15 nucleotides of the downstream intron was amplified by PCR using oligonucleotides A1B2 (TTCCATAACCCTGTCCACCA) and STR-2 (CTACTGGATACC), and cloned into the *EcoRV* site of pBluescript KS+ to produce pK7B. The *BstXI*–*SmaI* fragment from pK7B was then substituted for the *BstXI*–*SmaI* fragment of pmA1. pmA1RCE4 resulted from the reannealing between oligonucleotide CE4 (AGCTAGATTAGACTTCTAGAGTTT) and its complementary sequence inserted into the *PstI* and *EcoRI* sites of pmA1Δ, following Klenow treatment. pmA1ΔΔ was constructed by inserting an *EcoRV* linker into the T4 DNA polymerase-treated pmA1Δ, previously cut with *Bsu36I* and *SphI*. To produce pmA1RCE1a, the CE1a oligonucleotide (GGGTACCTTTAGAGTAGGCCT) reannealed to its complementary sequence was inserted into the *EcoRV* site of pmA1ΔCE1. The structure of the resulting constructs was confirmed by extensive restriction enzyme analysis, and DNA sequencing when appropriate.

pS2, pK7A1 and pKgA1 were described in Chabot *et al.* (1997). pC5' –/– was made by inserting the *BstXI*–*SmaI* fragment from pK7B into pS2 previously cut with *BstXI* and *SmaI*. To generate pC5' 1a/–, pC5' 4/–, pC5' –/1a and pC5' –/4, the reannealed CE1a and CE4 oligonucleotides were inserted into the *SmaI* site downstream of exon 7 and/or the *EcoRI* site downstream of exon 7B of pC5' –/–. *BstXI*–*SacI* fragments from C5' –/4 were inserted into the *BstXI* and *SacI* sites of C5' 1a/– and C5' 4/– to produce pC5' 1a/1a, pC5' 4/4, pC5' 1a/4 and pC5' 4/1a.

pC3' –/– was constructed from the stepwise assembly of various restriction fragments. First, the *SphI*–*XbaI* fragment of pKgA1 was inserted into pUC18 to yield pUC7Ba. pUC7Ba was digested with *HindIII* and filled-in with Klenow prior to digestion with *XbaI*. The *HindIII* filled–*XbaI* fragment was then ligated into the *SmaI*–*XbaI* site of pK7A1 to yield pK77BΔ. Ligation of a *PstI*–*SmaI*–*BamHI* adaptor (CCCGGGATCCTGCA) into the *PstI* site of pK77BΔ generated pK77BΔ+. A *HindIII*–*HincII* fragment containing the 3' splice site of adenovirus L2 exon from pSPAd (Solnick, 1985b) was inserted into pBluescript II KS+ to generate pK3'Ad. Finally, the Klenow-filled *EcoRI*–*XhoI* fragment from pK3'Ad was inserted into the *BstXI* site of pK77BΔ+ blunted by treatment with T4 DNA polymerase to generate pC3' –/–. The reannealed CE1a or CE4 oligonucleotides were inserted into the *SmaI* site downstream of exon 7 or into the *EcoRV* site downstream of exon 7B of pC3' –/– to generate pC3' 1a/– and pC3' 4/–

or pC3' –/1a and pC3' –/4, respectively. Insertion of the reannealed CE4m oligonucleotide (AGCTAGATTAGATCTTC) in the sense or antisense orientation into the *EcoRV* site downstream of exon 7B in pC3' –/– generated pC3' –/4m and pC3' –/4mα, respectively. The reannealed CE4, CE1a or CE4m oligonucleotides were inserted into the *EcoRV* site of pBluescript KS+ to produce pKCE4 and pKCE4m.

Transfection assays

Transfections in HeLa cells were performed using the standard calcium phosphate co-precipitation procedure. At 48 h post-transfection, total RNA was extracted and RT-PCR was performed as described previously using oligonucleotides A1E9 (GTGGTTTAGCAAAGTACTGG) and E1a-569 (ATTATCTGCCACGGAGGTGT). Markers for the A1 and A1^B amplification products were obtained by direct PCR amplification of A1 and A1^B cDNAs inserted into expression vectors.

Transcription and splicing assays

Splicing substrates were produced from plasmids linearized with *ScaI*, except for S774, S774α and S747 which were obtained by cutting with *HindIII*, and transcribed with T3 RNA polymerase (Pharmacia Biotech) in the presence of cap analog and [α -³²P]UTP (Amersham). CE1a, CE4 and CE4m RNAs were produced from pKCE1a, pKCE4 and pKCE4m linearized with *EcoRI* and transcribed as above. RNA purification was performed as described in Chabot (1994). HeLa nuclear extracts were prepared (Dignam *et al.*, 1983) and used in splicing reactions as described previously (Chabot *et al.*, 1997). The identity of lariat molecules was confirmed by debranching reactions in S100 extract and by comparing the migration of debranched products with known standards.

Gel-shift assays

rA1 or rUP1 were incubated for 5 min on ice in splicing conditions prior to addition of loading dye [1× agarose dye, 1 mg/ml heparin (final concentration)]. The reactions were run on 6% native acrylamide gel [29:1 acrylamide:bis-acrylamide, 5% glycerol, 50 mM Tris (pH 8.8), 50 mM glycine] in Tris–glycine running buffer [50 mM Tris (pH 8.8), 50 mM glycine].

RNA affinity column assay

Transcription to produce 'cold' CE1a was carried out in the absence of the cap analog, in the presence of 1 mM rNTPs and of a small amount of [α -³²P]UTP to aid in the purification and quantitation of CE1a RNA. A 400 pmol aliquot of CE1a RNA was coupled to 500 μl of agarose adipic acid hydrazide according to the manufacturer's recommendation (Pharmacia-Amersham). Microcolumns were set up with 25 μl of CE1a beads packed into a 200 μl pipetman tip. A mixture containing 2.5 fmol of ³²P-labeled CE1a and K+ RNAs (produced from standard transcriptions) and either rA1 or rUP1 (5 pmol) was applied to the column. This mixture was set up in buffer R50 [20 mM HEPES pH 7.9, 0.5 mM EDTA, 50 mM KCl, 3 mM MgCl₂, 1 mM dithiothreitol, 0.05% NP-40 and 10% glycerol]. The column was washed with 200 μl of buffer R50 and the flow-through was concentrated to 50 μl on a Microcon-30 column (Millipore). Stepwise elutions were accomplished with 50 μl of buffers R100, R250 and R500 (each one identical to buffer R50 except that the KCl concentration was 100, 250 and 500 mM, respectively). Input, flow-through and eluted fractions were aliquoted in two sets. One set was extracted with phenol/chloroform/isoamylalcohol, ethanol precipitated and loaded onto an 8% acrylamide–8 M urea gel for RNA visualization. Samples from the other set were loaded onto a 12.5% SDS–polyacrylamide gel and the fractionated proteins were silver stained.

Acknowledgements

We thank J.Toutant for performing transfections, and S.AbouElela for critical comments on the manuscript. This work was supported by a grant from the Medical Research Council of Canada (MRCC). M.B. is the recipient of a studentship from the MRCC. B.C. is Chercheur Boursier Senior from the FRSC.

References

- Abdul-Manan,N. and Williams,K.R. (1996) hnRNP A1 binds promiscuously to oligoribonucleotides: utilization of random and homo-oligonucleotides to discriminate sequence from base-specific binding. *Nucleic Acids Res.*, **24**, 4063–4070.
- Abovich,N. and Rosbash,M. (1997) Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell*, **89**, 403–412.
- Achsel,T. and Shimura,Y. (1996) Factors involved in the activation of pre-mRNA splicing from downstream splicing enhancers. *J. Biochem.*, **120**, 53–60.
- Amendt,B.A., Si,Z.H. and Stoltzfus,C.M. (1995) Presence of exon splicing silencers within human immunodeficiency virus type 1 *tat* exon 2 and *tat-rev* exon 3: evidence for inhibition mediated by cellular factors. *Mol. Cell. Biol.*, **15**, 6480.
- Amero,S.A., Raychaudhuri,G., Cass,C.L., van Venrooij,W.J., Habets,W.J., Krainer,A.R. and Beyer,A.L. (1992) Independent deposition of heterogeneous nuclear ribonucleoproteins and small nuclear ribonucleoprotein particles at sites of transcription. *Proc. Natl Acad. Sci. USA*, **89**, 8409–8413.
- Ashiya,M. and Grabowski,P.J. (1997) A neuron-specific splicing switch mediated by an array of pre-mRNA repressor sites: evidence of a regulatory role for the polypyrimidine tract binding protein and a brain-specific PTB counterpart. *RNA*, **3**, 996–1015.
- Beyer,A.L. and Osheim,Y.N. (1988) Splice site selection, rate of splicing and alternative splicing on nascent transcripts. *Genes Dev.*, **2**, 754–765.
- Black,D.L. (1995) Finding splice sites within a wilderness of RNA. *RNA*, **1**, 763–771.
- Blanchette,M. and Chabot,B. (1997) A highly stable duplex structure sequesters the 5' splice site region of hnRNP A1 alternative exon 7B. *RNA*, **3**, 405–419.
- Burd,C.G. and Dreyfuss,G. (1994) RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing. *EMBO J.*, **13**, 1197–1204.
- Caceres,J.F., Stamm,S., Helfman,D.M. and Krainer,A.R. (1994) Regulation of alternative splicing *in vivo* by overexpression of antagonistic splicing factors. *Science*, **265**, 1706–1709.
- Caputi,M., Casari,G., Guenzi,S., Tagliabue,R., Sidoli,A., Melo,C.A. and Baralle,F.E. (1994) A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. *Nucleic Acids Res.*, **22**, 1018–1022.
- Cartegni,L., Maconi,M., Morandi,E., Cobianchi,F., Riva,S. and Biamonti,G. (1996) hnRNP A1 selectively interacts through its Gly-rich domain with different RNA-binding proteins. *J. Mol. Biol.*, **259**, 337–348.
- Casas-Finet,J.R., Smith,J.D., Jr, Kumar,A., Kim,J.G., Wilson,S.H. and Karpel,R.L. (1993) Mammalian heterogeneous ribonucleoprotein A1 and its constituent domains. Nucleic acid interaction, structural stability and self-association. *J. Mol. Biol.*, **229**, 873–889.
- Chabot,B. (1994) Synthesis and purification of RNA substrates. In Hames,D. and Higgins,S. (eds), *RNA Processing*. Oxford University Press, Oxford, UK, Vol. 1, pp. 1–29.
- Chabot,B. (1996) Directing alternative splicing: cast and scenarios. *Trends Genet.*, **12**, 472–478.
- Chabot,B., Blanchette,M., Lapierre,I. and La Branche,H. (1997) An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. *Mol. Cell. Biol.*, **17**, 1776–1786.
- Chan,R.C. and Black,D.L. (1997a) Conserved intron elements repress splicing of a neuron-specific *c-src* exon *in vitro*. *Mol. Cell. Biol.*, **17**, 2970.
- Chan,R.C. and Black,D.L. (1997b) The polypyrimidine tract binding protein binds upstream of neural cell-specific *c-src* exon N1 to repress the splicing of the intron downstream. *Mol. Cell. Biol.*, **17**, 4667–4676.
- Charpentier,B. and Rosbash,M. (1996) Intramolecular structure in yeast introns aids the early steps of *in vitro* spliceosome assembly. *RNA*, **2**, 509–522.
- Conway,G., Wooley,J., Bibring,T. and LeSturgeon,W.M. (1988) Ribonucleoproteins package 700 nucleotides of pre-mRNA into a repeating array of regular particles. *Mol. Cell. Biol.*, **8**, 2884–2895.
- Del Gatto,F. and Breathnach,R. (1995) Exon and intron sequences, respectively, repress and activate splicing of a fibroblast growth factor receptor 2 alternative exon. *Mol. Cell. Biol.*, **15**, 4825–4834.
- Del Gatto,F., Gesnel,M.C. and Breathnach,R. (1996) The exon sequence TAGG can inhibit splicing. *Nucleic Acids Res.*, **24**, 2017–2021.
- Del Gatto-Konczak,F., Olive,M., Gesnel,M.C. and Breathnach,R. (1999) hnRNP A1 recruited to an exon *in vivo* can function as an exon splicing silencer. *Mol. Cell. Biol.*, **19**, 251–260.
- Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.*, **11**, 1475–1489.
- Engelbrecht,J., Knudsen,S. and Brunak,S. (1992) G+C-rich tract in 5' end of human introns. *J. Mol. Biol.*, **227**, 108–113.
- Eperon,I.C., Ireland,D.C., Smith,R.A., Mayeda,A. and Krainer,A.R. (1993) Pathways for selection of 5' splice sites by U1 snRNPs and SF2/ASF. *EMBO J.*, **12**, 3607–3617.
- Eperon,L.P., Estibeiro,J.P. and Eperon,I.C. (1986) The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. *Nature*, **324**, 280–282.
- Eperon,L.P., Graham,I.R., Griffiths,A.D. and Eperon,I.C. (1988) Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell*, **54**, 393–401.
- Fu,X.D. (1993) Specific commitment of different pre-mRNAs to splicing by single SR proteins. *Nature*, **365**, 82–85.
- Gallego,M.E., Gattoni,R., Stevenin,J., Marie,J. and Expert-Bezancon,A. (1997) The SR splicing factors ASF/SF2 and SC35 have antagonistic effects on intronic enhancer-dependent splicing of the β -tropomyosin alternative exon 6A. *EMBO J.*, **16**, 1772–1784.
- Heinrichs,V. and Baker,B.S. (1995) The *Drosophila* SR protein RBP1 contributes to the regulation of doublesex alternative splicing by recognizing RBP1 RNA target sequences. *EMBO J.*, **14**, 3987–4000.
- Howe,K.J. and Ares,M., Jr (1997) Intron self-complementarity enforces exon inclusion in a yeast pre-mRNA. *Proc. Natl Acad. Sci. USA*, **94**, 12467–12472.
- Humphrey,M.B., Bryan,J., Cooper,T.A. and Berget,S.M. (1995) A 32-nucleotide exon-splicing enhancer regulates usage of competing 5' splice sites in a differential internal exon. *Mol. Cell. Biol.*, **15**, 3979–3988.
- Kanopka,A., Muhlemann,O. and Akusjarvi,G. (1996) Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature*, **381**, 535–538.
- Kohtz,J.D., Jamison,S.F., Will,C.L., Zuo,P., Luhrmann,R., Garcia-Blanco,M.A. and Manley,J.L. (1994) Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature*, **368**, 119–124.
- Kumar,A. and Wilson,S.H. (1990) Studies of the strand-annealing activity of mammalian hnRNP complex protein A1. *Biochemistry*, **29**, 10717–10722.
- Lavigne,A., La Branche,H., Kornblihtt,A.R. and Chabot,B. (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.
- Libri,D., Stutz,F., McCarthy,T. and Rosbash,M. (1995) RNA structural patterns and splicing: molecular basis for an RNA-based enhancer. *RNA*, **1**, 425–436.
- Lin,C.H. and Patton,J.G. (1995) Regulation of alternative 3' splice site selection by constitutive splicing factors. *RNA*, **1**, 234–245.
- Lynch,K.W. and Maniatis,T. (1995) Synergistic interactions between two distinct elements of a regulated splicing enhancer. *Genes Dev.*, **9**, 284–293.
- Matunis,E.L., Matunis,M.J. and Dreyfuss,G. (1993) Association of individual hnRNP proteins and snRNPs with nascent transcripts. *J. Cell Biol.*, **121**, 219–228.
- Matunis,M.J., Matunis,E.L. and Dreyfuss,G. (1992) Isolation of hnRNP complexes from *Drosophila melanogaster*. *J. Cell Biol.*, **116**, 245–255.
- Mayeda,A. and Krainer,A.R. (1992) Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell*, **68**, 365–375.
- Mayeda,A., Munroe,S.H., Caceres,J.F. and Krainer,A.R. (1994) Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *EMBO J.*, **13**, 5483–5495.
- McAfee,J.G., Huang,M., Soltaninassab,S., Rech,J.E., Iyengar,S. and LeSturgeon,W.N. (1997) The packaging of pre-mRNA. In Krainer,A.R. (ed.), *Eukaryotic mRNA Processing*. Oxford University Press, Oxford, UK, pp. 68–102.

- McCullough,A.J. and Berget,S.M. (1997) G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection. *Mol. Cell. Biol.*, **17**, 4562–4571.
- Min,H., Turck,C.W., Nikolic,J.M. and Black,D.L. (1997) A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer. *Genes Dev.*, **11**, 1023–1036.
- Munroe,S.H. and Dong,X.F. (1992) Heterogeneous nuclear ribonucleoprotein A1 catalyzes RNA.RNA annealing. *Proc. Natl Acad. Sci. USA*, **89**, 895–899.
- Newman,A. (1987) Specific accessory sequences in *Saccharomyces cerevisiae* introns control assembly of pre-mRNAs into spliceosomes. *EMBO J.*, **6**, 3833–3839.
- Nussinov,R. (1988) Conserved quartets near 5' intron junctions in primate nuclear pre-mRNA. *J. Theor. Biol.*, **133**, 73–84.
- Nussinov,R. (1989) Conserved signals around the 5' splice sites in eukaryotic nuclear precursor mRNAs: G-runs are frequent in the introns and C in the exons near both 5' and 3' splice sites. *J. Biomol. Struct. Dynam.*, **6**, 985–1000.
- Parker,R. and Patterson,B. (1987) Architecture of fungal introns: implications for spliceosome assembly. In Dudock,B. and Inouye,M. (eds), *Molecular Biology of RNA. New Perspectives*. Academic Press, New York, NY, pp. 133–150.
- Pinol-Roma,S., Swanson,M.S., Gall,J.G. and Dreyfuss,G. (1989) A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts. *J. Cell Biol.*, **109**, 2575–2587.
- Pontius,B.W. and Berg,P. (1990) Renaturation of complementary DNA strands mediated by purified mammalian heterogeneous nuclear ribonucleoprotein A1 protein: implications for a mechanism for rapid molecular assembly. *Proc. Natl Acad. Sci. USA*, **87**, 8403–8407.
- Portman,D.S. and Dreyfuss,G. (1994) RNA annealing activities in HeLa nuclei. *EMBO J.*, **13**, 213–221.
- Ramchatesingh,J., Zahler,A.M., Neugebauer,K.M., Roth,M.B. and Cooper,T.A. (1995) A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol. Cell. Biol.*, **15**, 4898–4907.
- Si,Z., Amendt,B.A. and Stoltzfus,C.M. (1997) 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. *Nucleic Acids Res.*, **25**, 861–867.
- Siebel,C.W., Admon,A. and Rio,D.C. (1995) Soma-specific expression and cloning of PSI, a negative regulator of P element pre-mRNA splicing. *Genes Dev.*, **9**, 269–283.
- Singh,R., Valcarcel,J. and Green,M.R. (1995) Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science*, **268**, 1173–1176.
- Solnick,D. (1985a) Alternative splicing caused by RNA secondary structure. *Cell*, **43**, 667–676.
- Solnick,D. (1985b) Trans splicing of mRNA precursors. *Cell*, **42**, 157–164.
- Solnick,D. and Lee,S.I. (1987) Amount of RNA secondary structure required to induce an alternative splice. *Mol. Cell. Biol.*, **7**, 3194–3198.
- Solovyev,V.V., Salamov,A.A. and Lawrence,C.B. (1994) Predicting internal exons by oligonucleotide composition and discriminant analysis of spliceable open reading frames. *Nucleic Acids Res.*, **22**, 5156–5163.
- Staffa,A. and Cochrane,A. (1995) Identification of positive and negative splicing regulatory elements within the terminal *tat*-*rev* exon of human immunodeficiency virus type 1. *Mol. Cell. Biol.*, **15**, 4597–4605.
- Staknis,D. and Reed,R. (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.
- Sun,Q., Mayeda,A., Hampson,R.K., Krainer,A.R. and Rottman,F.M. (1993) General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. *Genes Dev.*, **7**, 2598–2608.
- Tian,M. and Maniatis,T. (1992) Positive control of pre-mRNA splicing *in vitro*. *Science*, **256**, 237–240.
- Valcarcel,J., Singh,R., Zamore,P.D. and Green,M.R. (1993) The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature*, **362**, 171–175.
- Visa,N., Alzhanova-Ericsson,A.T., Sun,X., Kiseleva,E., Bjorkroth,B., Wurtz,T. and Daneholt,B. (1996) A pre-mRNA-binding protein accompanies the RNA from the gene through the nuclear pores and into polysomes. *Cell*, **84**, 253–264.
- Wang,Z., Hoffmann,H.M. and Grabowski,P.J. (1995) Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA*, **1**, 21–35.
- Wu,J.Y. and Maniatis,T. (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*, **75**, 1061–1070.
- Wurtz,T., Kiseleva,E., Nacheva,G., Alzhanova-Ericsson,A., Rosen,A. and Daneholt,B. (1996) Identification of two RNA-binding proteins in Balbiani ring premessenger ribonucleoprotein granules and presence of these proteins in specific subsets of heterogeneous nuclear ribonucleoprotein particles. *Mol. Cell. Biol.*, **16**, 1425–1435.
- Yang,X., Bani,M.R., Lu,S.J., Rowan,S., Ben-David,Y. and Chabot,B. (1994) The A1 and A1B proteins of heterogeneous nuclear ribonucleoproteins modulate 5' splice site selection *in vivo*. *Proc. Natl Acad. Sci. USA*, **91**, 6924–6928.
- Zheng,Z.M., He,P. and Baker,C.C. (1996) Selection of the bovine papillomavirus type 1 nucleotide 3225 3' splice site is regulated through an exonic splicing enhancer and its juxtaposed exonic splicing suppressor. *J. Virol.*, **70**, 4691–4699.
- Zu,K., Sikes,M.L. and Beyer,A.L. (1998) Separable roles *in vivo* for the two RNA binding domains of *Drosophila* A1-hnRNP homolog. *RNA*, **4**, 1585–1598.
- Zuo,P. and Maniatis,T. (1996) The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. *Genes Dev.*, **10**, 1356–1368.

Received May 25, 1998; revised and accepted February 5, 1999