Natural base-pairing interactions between 5' splice site and branch site sequences affect mammalian 5' splice site selection

J. Cote and B. Chabot

RNA 1997 3: 1248-1261

RNA (1997), 3:1248–1261. Cambridge University Press. Printed in the USA.

Natural base-pairing interactions between 5′ splice site and branch site sequences affect mammalian 5′ splice site selection

JOCELYN CÔTÉ and BENOIT CHABOT

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

ABSTRACT

In the murine gene encoding the neuronal cell adhesion molecule (NCAM), the integrity of the 5′ splice site of exon 18 (E18) is essential for regulation of alternative splicing. To further study the contribution of 5′ splice site sequences, we used a simple NCAM pre-mRNA containing a portion of E18 fused to E19 and separated by a shortened intron. This RNA is spliced in vitro to produce five sets of lariat intermediates and products, the most abundant set displaying aberrant migration in acrylamide/urea gels. Base pairing interactions between positions +5 and +8 of the intron and positions –3 and –6 from the branch point were responsible for the faster migration of this set of lariat molecules. To test whether the duplex structure forms earlier and contributes to 5′ splice site selection, we used NCAM substrates carrying the 5′ splice sites of E17 and E18 in competition for the 3′ splice site of E19. Mutations upstream of the major branch site improve E18/E19 splicing in NIH3T3 extracts, whereas compensatory mutations at positions +7 and +8 neutralize the effect of branch site mutations and curtail E18/E19 splicing. Our data suggest that duplex formation occurs early and interferes with the assembly of complexes initiated on the 5′ splice site of NCAM E18. This novel type of intron interaction may exist in the introns of other mammalian pre-mRNAs.

Keywords: alternative splicing; NCAM; RNA structure; spliceosome

INTRODUCTION

The exquisite precision that characterizes the splicing of nuclear pre-mRNAs is orchestrated within the spliceosome, a large multicomponent complex assembled sequentially on the pre-mRNA. The splicing reaction requires the coordinated action of small nuclear ribonucleoproteins (snRNPs) and a variety of non-snRNP splicing factors (Moore et al., 1993). The 5′ splice site region is recognized initially by the U1 snRNP, whereas the branch site region is recognized by the U2 snRNP following binding of $U2AF^{65}$ to the downstream polypyrimidine tract-AG at the 3′ splice junction. Other spliceosomal factors include the SR proteins (Fu, 1995; Chabot, 1996; Manley & Tacke, 1996) which, in addition to their ability to bind to 5′ splice site sequences directly (Kohtz et al., 1994; Zuo & Manley, 1994), can promote commitment complex formation through simultaneous interaction with the U1 snRNP 70K and the $U2AF^{35}$ proteins (Wu & Maniatis, 1993; Staknis & Reed, 1994; Zuo & Maniatis, 1996). The subsequent entry of the [U4/U6.U5] tri-snRNP particle leads to major rearrangements in snRNP interactions and allows U6 snRNA to interact with 5′ splice site sequences (Nilsen, 1994; Sharp, 1994). Following cleavage at the 5′ splice site and concomitant branch formation, other rearrangements occur and the second step of splicing then releases the lariat intron and the ligated exons.

Although the biochemical events that lead to intron removal are now becoming better understood, the parameters that dictate recognition and pairing of appropriate pairs of splice sites on complex pre-mRNAs are just beginning to be deciphered (Black, 1995; Reed, 1996). The situation is further complicated when alternative splicing is considered, for splicing signals that are recognized normally and used efficiently in a given cell-type or tissue can be ignored completely in another cellular environment or developmental stage (Smith et al., 1989; McKeown, 1992). Clearly, splice site sequences alone cannot fully account for their frequency of utilization in complex pre-mRNAs. Splicing

Reprint requests to: Benoit Chabot, Département de Microbiologie et d'Infectiologie, Faculté de Médecine, Université de Sherbrooke, 3001 12e Avenue Nord, Sherbrooke, Québec J1H 5N4, Canada; e-mail: b.chabot@courrier.usherb.ca.

signals in alternative splicing units are often suboptimal, possibly to accommodate the contribution of other elements that influence splice site selection. Intron sequences near the 3′ splice site are often tailored to allow the interaction of factors that modulate the binding of U2AF^{65} and U2 snRNP (Valcárcel et al., 1993; Lin & Patton, 1995; Kanopka et al., 1996).

Sequences flanking a splice site often influence its frequency of utilization. Splicing enhancers that improve the use of an upstream 3′ splice site have been found in many vertebrate alternative and constitutive exons (reviewed in Black, 1995). In one case, a purine-rich enhancer activates splicing at an upstream 5′ splice site (Humphrey et al., 1995). Several of these positive elements are recognized by the SR family of splicing factors (reviewed in Fu, 1995; Chabot, 1996; Manley & Tacke, 1996), which promote a more efficient interaction of U2AF^{65} with the 3′ splice site region (Lavigueur et al., 1993; Wang et al., 1995). Likewise, U1 snRNP binding to a 5′ splice site can facilitate the binding of U2AF^{65} to an upstream or a downstream 3′ splice site through a network of interactions that requires the participation of SR proteins (Hoffman & Grabowski, 1992; Staknis & Reed, 1994; Côté et al., 1995). An alternative array of cross-intron bridging interactions involving SF1 and a new U1 snRNP-associated protein was also recently described (Abovich & Rosbash, 1997). Positively acting sequences have been uncovered in introns, but the details of the mechanism of stimulation remain to be understood (Huh & Hynes, 1994; Zhao et al., 1994; Del Gatto & Breathnach, 1995; Min et al., 1995; Sirand-Pugnet et al., 1995a; Carlo et al., 1996). A variety of exon and intron elements that affect the recognition of nearby splice sites negatively have been reported (Watakabe et al., 1991; Gontarek et al., 1993; Amendt et al., 1994; Caputi et al., 1994; Siebel et al., 1994, 1995; Chan & Black, 1995; Staffa & Cochrane, 1995; Adams et al., 1997). Finally, RNA secondary structure can also influence splice site selection either negatively, by sequestering splice sites (Solnick, 1985; Eperon et al., 1986; Solnick & Lee, 1987; Halfter & Gallwitz, 1988; Helfman et al., 1990; Clouet d'Orval et al., 1991a, 1991b; Deshler & Rossi, 1991; Libri et al., 1991; Estes et al., 1992; Goguel et al., 1993; Liu et al., 1995; Sirand-Pugnet et al., 1995b; Blanchette & Chabot, 1997), or positively, by decreasing the effective distance between splicing signals (Chebli et al., 1989; Goguel & Rosbash, 1993; Libri et al., 1995; Charpentier & Rosbash, 1996), or by facilitating the presentation of splice sites (Newman, 1987; Kister et al., 1993). Structure formation in the yeast RPL32 pre-mRNA allows formation of a stable complex with U1 snRNP and the L32 protein, preventing the entry of U2 snRNP into the spliceosome (Vilardell & Warner, 1994).

Regulation of gene expression through alternative splicing is a recurrent theme in the mammalian nervous system (Burke et al., 1992; Stamm et al., 1994). For example, in the mouse c-*src* gene, the 18 nt-long N1 exon is included in neurons, but skipped in all other tissues. In non-neuronal cells, negative elements located on both sides of the alternate exon promote exon skipping (Chan & Black, 1995). In neuronal cells, the block to N1 inclusion is relieved through the positive control of an intronic element bound by several proteins, including hnRNP F and a new regulatory protein, KSRP (Min et al., 1995, 1997). The pre-mRNA encoding the larger forms of the mouse neuronal cell adhesion molecule (NCAM) is alternatively spliced to yield two mRNAs differing by the incorporation of an 801-nt exon (exon 18 or E18) (Barbas et al., 1988). The 5′ splice site of E18 is important in setting the regulated splicing profile because replacing it with a corresponding sequence from the α-globin gene promotes skipping of E18 in neuronal cells (Tacke & Goridis, 1991). In a previous study, we reported that a point mutation at position +6 of the 5′ splice site of E18 promoted more efficient U2AF^{65} binding to the 3′ splice site of E19 (Côté et al., 1995). Because this mutation improves the complementarity with U1 snRNA, stronger U1 snRNP binding may drive the assembly of commitment complexes to stabilize U2AF^{65} binding. However, in the present study, we report that this mutation also promotes a change in the migration of naked lariat molecules in acrylamide gels. We document the existence of base pairing interactions between the 5′ splice site of E18 and sequences upstream of the major branch site in the lariat intron. Because destabilizing mutations improve the frequency with which the 5′ splice site of E18 is used in vitro, our work suggests that duplex formation occurs early and contributes to the selection of 5′ splice sites in this NCAM alternative splicing unit. We propose that this novel type of interaction between 5′ splice site and branch site regions exists in other mammalian introns.

RESULTS

Aberrant migration of lariat molecules

As part of a study aimed at investigating the effects of NCAM 5′ splice site sequences on U2AF^{65} binding, we had used four model pre-mRNA substrates (Côté et al., 1995; Fig. 1A). E RNA contains the 3′ portion of exon 18 (E18) with its donor splice site fused to the 5′ portion of exon 19 (E19) and its acceptor site. M RNA is identical to E RNA except for a G to U mutation at position +6 of the intron, converting the 5′ splice site of E18 into that of exon 17 (E17). W RNA was produced by converting the 5′ splice site of E18 into that of the adenovirus major late L1 exon. D RNA contains the 3′ portion of E17 with its respective donor site fused to the 5′ portion of E19 and its acceptor site. Using these model pre-mRNAs, we showed that U2AF^{65} bound less readily to E than to the D, M, and

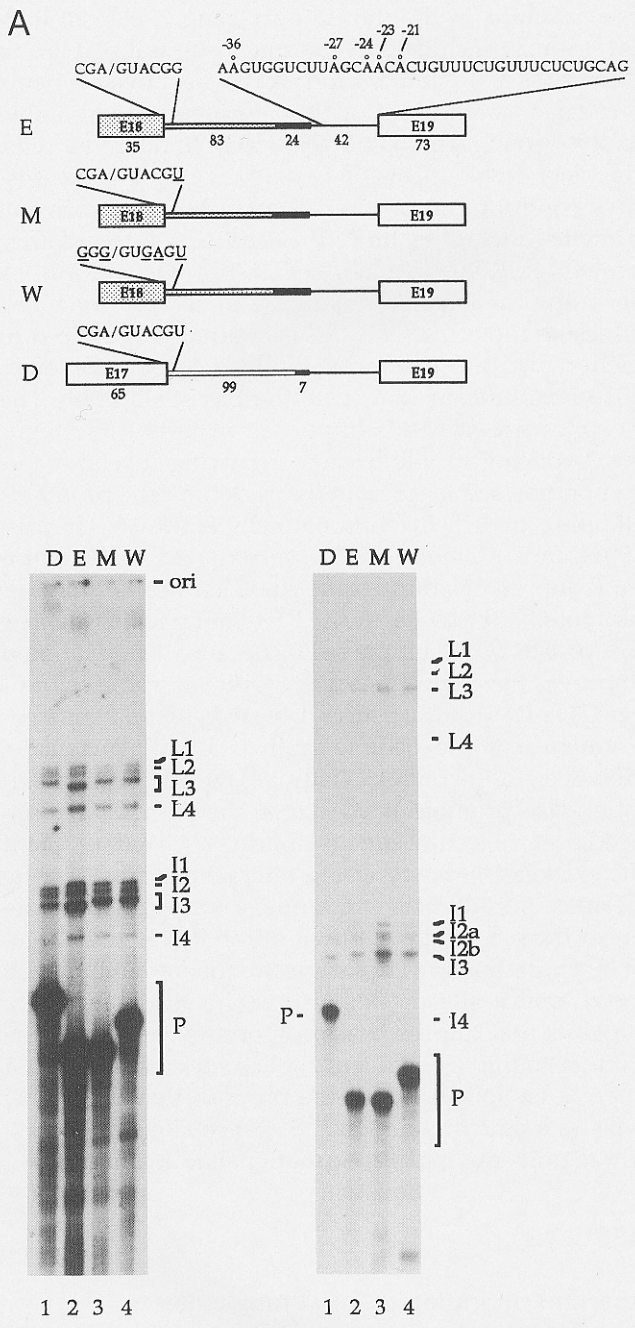

FIGURE 1. In vitro splicing of simple N-CAM pre-mRNAs. **A:** Structure of the RNA substrates. Larger boxes indicate exons. Thinner boxes and lines indicate intron sequences downstream of E18 and upstream of E19, respectively. The black area in the intron indicates a plasmid-derived polylinker region. The length (in nucleotides) of each exon and intron region is indicated. The 5′ splice site sequence of each RNA is shown and the mutated nucleotides are underlined. The 3′ splice site sequence common to all RNAs is shown and the branch sites used are indicated by a circle above the nucleotide. **B:** ^{32}P-labeled transcripts were incubated for 2 h in HeLa nuclear extracts under splicing conditions. Splicing products were separated on a 9% acrylamide/7 M urea gel. Positions of pre-mRNAs (P), lariat intermediates (L1–L4), and lariat introns (I1–I4) are indicated. Note that the slower migration of W RNA is caused by a cloning artifact that increased the size of the first exon (data not shown). **C:** Splicing products derived from incubations performed as in B were separated in a 9% acrylamide/8 M urea/30% formamide gel. To improve resolution of lariat species, gels were run extensively accounting for the absence of the mRNAs and 5′ exons in the gels shown in B and C.

W RNAs (Côté et al., 1995). To address whether the mutations affected splicing efficiency, we incubated the RNAs in HeLa nuclear extracts. Splicing products were isolated and separated on a 9% polyacrylamide/7 M urea gel. D, E, M, and W RNAs were all spliced with equivalent efficiencies (Fig. 1B, lanes 1–4), as judged by the levels of lariat intermediates and products migrating above the pre-mRNAs. This result was not unexpected because the effect of mutations is not always reflected in the rate of splicing of simple pre-mRNAs (Reed & Maniatis, 1986). Inspection of the splicing profiles revealed that four distinct sets of lariat intermediates (L1–L4) and products (I1–I4) were generated for each RNA. In gels run longer, I2 lariat molecules were resolved into two species (data not shown). The L1:I1, L2:I2, and L4:I4 lariat species derived from M and W RNAs co-migrated exactly with those produced from E RNA. Notably, the most abundant set of lariat molecules generated from M and W RNAs ($L3^M$:$I3^M$ and $L3^W$:$I3^W$, respectively) migrated significantly slower than the corresponding lariat species derived from E RNA ($L3^E$:$I3^E$). Thus, a single G to U mutation at position +6 of the 5′ splice site of E18 shifted the migration of the most abundant set of lariat molecules. A similar difference is noted when the migration of lariat molecules produced from E and D RNAs is compared. The size of the D intron is 1 nt shorter than the E intron; lariat species $L3^D$:$I3^D$ and $L3^E$:$I3^E$ displayed a difference in migration that was inconsistent with the 1-nt difference in size (Fig. 1B, compare lane 1 with lane 2).

Separation of the lariat products in 9% polyacrylamide/8 M urea/30% formamide gels yielded $L3^E$:$I3^E$ lariat species that now co-migrated with the $L3^M$:$I3^M$ and $L3^W$:$I3^W$ molecules (Fig. 1C, lanes 2–4). In such strongly denaturing conditions, the difference in migration between $L3^D$:$I3^D$ and $L3^E$:$I3^E$ was also reduced to a level consistent with a 1-nt difference in intron length (Fig. 1C, lane 1). These results suggest that the $L3^E$:$I3^E$ molecules differed from $L3^M$:$I3^M$ and $L3^W$:$I3^W$ by harboring a more compact structure, possibly because of the presence of more stable base-stacking interactions.

Because the difference in migration was detected with only one of the five sets of lariat molecules, we set out to identify the splice junctions and branch sites used during the formation of $I3^E$ and $I3^M$. The equivalent size of $I3^E$ and $I3^M$ was confirmed by performing a debranching reaction on purified lariat molecules in a HeLa S100 extract followed by separation of the products in a denaturing gel (Fig. 2A). Linear 149 nt-long introns were produced for both $I3^E$ and $I3^M$ in conditions that allowed detection of a 1-nt difference in size (Fig. 2A, lanes 5–7). The I1, I2, and I4 lariat species derived from either E or M RNA also yielded co-migrating 149 nt-long products upon debranching (data not shown).

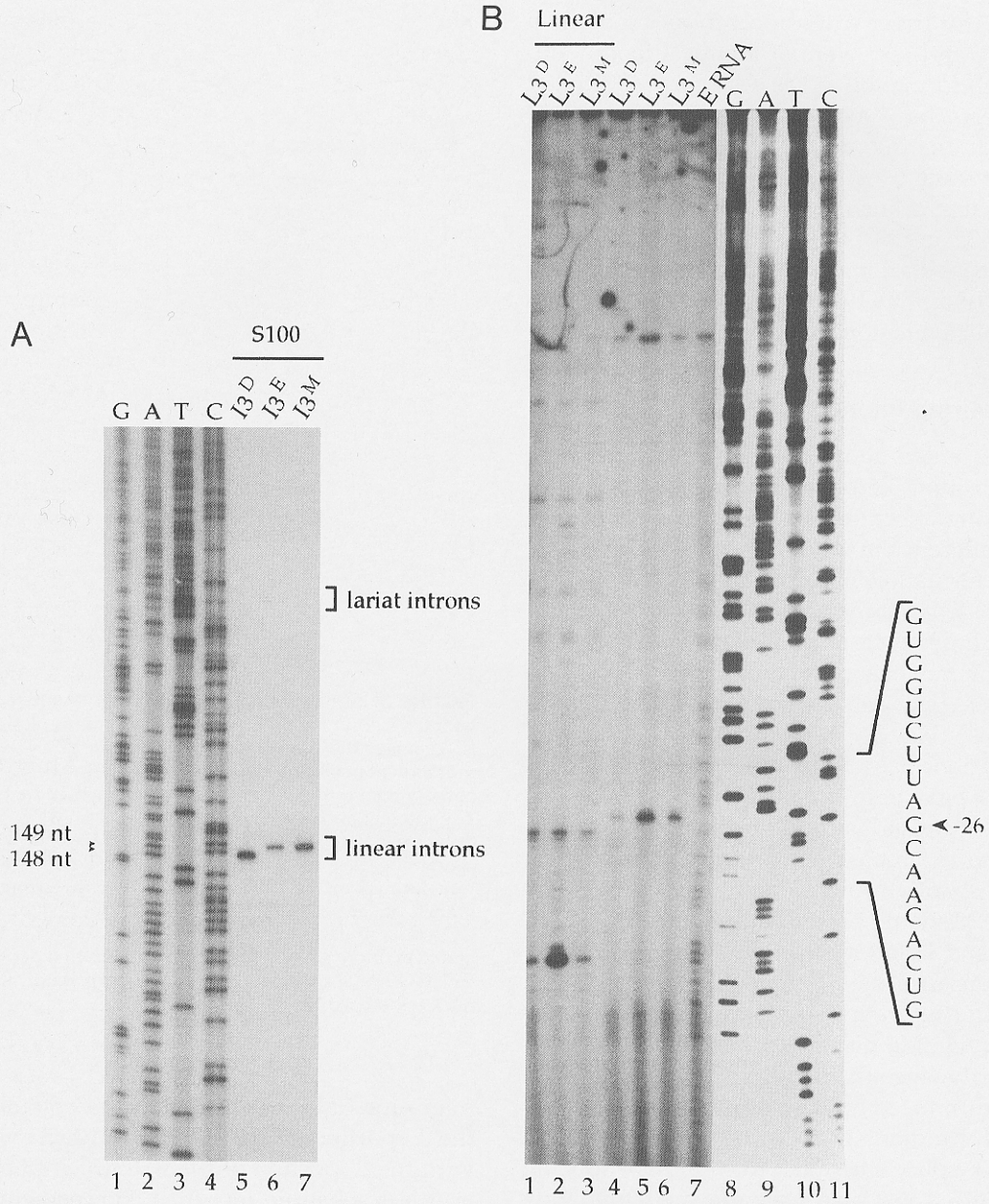

FIGURE 2. Characterization of lariat molecules produced from D, E, and M RNA splicing. **A:** Size determination of the $I3^D$, $I3^E$, and $I3^M$ intron molecules. A debranching reaction in a HeLa S100 extract was performed with each I3 lariat intron following its purification from a denaturing gel. Debranching reactions were loaded onto an 8% acrylamide/7 M urea gel. The position of the lariat and linear introns is indicated relative to the position of nt 148 and 149 on a DNA sequencing ladder (lanes 1–4). **B:** Branch site mapping was performed with each purified lariat intermediate species by reverse transcription with an oligonucleotide mapping in exon 19 (lanes 4–6). As controls, primer extension assays were performed with debranched intermediates (lanes 1–3) and intact E RNA (lane 7). Primer extension products were run in a 9% sequencing gel. The RNA sequence of the intron and the block to reverse transcriptase obtained with the L3 molecules are shown. The sequence is complementary to the DNA sequence of pSPE generated with the same oligonucleotide (lanes 8–11). The presence of a branch makes primer extension products terminate 1 nt before the actual branch point.

Because the difference in migration was also seen between lariat intermediates $L3^E$ and $L3^M$ (Fig. 1B, lanes 2 and 3), the $I3^E$ and $I3^M$ species were unlikely to result from the use of a different 3′ splice site. The position of the 5′ splice junction used to produce lariat introns was mapped by primer extension using an oligonucleotide complementary to the intron loop. The result indicated that all lariat species were produced from the use of the same 5′ splice junction (data not shown). Finally, the position of branch points on $L3^E$ and $L3^M$ lariat intermediates was mapped by primer extension using an oligonucleotide complementary to E19. The L1, L2a, L2b, and L4 lariats gave primer extension products that mapped the branch points at adenosines located, respectively, at positions −21, −23, −24, and −36 from the 3′ end of the intron (data not

shown; Fig. 1A). Primer extension products obtained from $L3^E$, $L3^M$, and $L3^D$ were of identical length and mapped branch formation at the adenosine located at position −27 from the 3′ splice junction (Fig. 2B, lanes 4–6). We conclude that $I3^E$ and $I3^M$ are generated from the use of the same 5′ splice site, the same 3′ splice site, and the same branch site. Thus, the distinct migration of the $I3^E$ and $I3^M$ lariat introns is likely due to a difference in base-stacking interactions associated with the formation of a branched structure 27-nt upstream from the 3′ splice junction.

Intron base pairing interactions

The results presented above suggest that the G to U mutation at position +6 of the E intron has a significant destabilizing effect on base-stacking interactions. Notably, the difference in migration was detected only when branch formation occurred at position −27 to generate the L3:I3 lariat molecules. Inspection of the 5′ splice site and branch site sequences revealed possible base pairing interactions involving positions +5 to +8 from the 5′ splice site and positions −30 to −33 from the 3′ splice site (Fig. 3A; bottom). In this configuration, the mutation in M RNA at position +6 would abolish a U·G base pair, which may destabilize the secondary structure and account for the slower migration of $L3^M$ and $I3^M$ molecules. If this structure exists, introducing disruptive mutations at positions +5 to +8 or −30 to −33 should produce more opened, slower-migrating L3 and I3 lariat molecules. Moreover, combining mutations that restore complementarity should regenerate faster-migrating lariat molecules in urea-containing acrylamide gels, provided that the $\Delta G°$ is similar. To verify this structural model, random mutations were introduced at positions +7 and +8 of the intron (E\78NN mutants), and at positions −33 and −32 from the 3′ splice site (E\L3XX mutants). Splicing of mutated RNA substrates was analyzed following incubation in HeLa extracts. All mutations retarded the migration of L3 and I3 molecules without affecting the migration of L1:I1 and L2:I2 species (Fig. 4, lanes 2–7 and 10–13). The E\78**GG** and E\78**GC** RNAs (bold characters indicate mutated nucleotides) produced L4 and I4 molecules with altered mobility (Fig. 4, lane 3 on longer exposures, and lane 7, respectively). This shift in mobility may be due to the formation of a similar structure between the mutated 5′ splice site and sequences upstream of the branch site used to generate the L4:I4 species (Fig. 3B). The mutation in E\L3**CC** was associated with a decrease in the production of L3:I3 molecules (Fig. 4, lane 12), suggesting that a C at position −32 reduces the use of the branch site at position −27. Our results indicate that mutations at positions +7, +8, −33, and −32 destabilized the structure of L3:I3 lariat molecules. Mutations in the central portion of the intron did not affect the

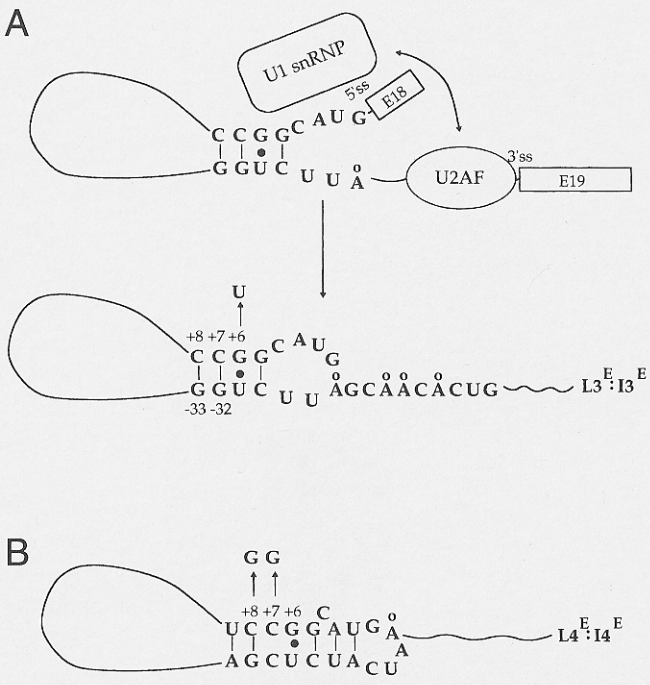

FIGURE 3. Structural model. **A:** Bottom: Putative base pairing interactions occurring in the $L3^E$:$I3^E$ lariat molecules produced from E RNA when the branch site at position −27 is used. The G to U mutation at position +6 in M RNA is indicated. Top: Representation of early commitment complex interactions on the E RNA substrate. The interaction between U1 snRNP and $U2AF^{65}$ at this stage may bring in proximity the 5′ splice site and branch site sequences, thereby promoting base pairing interactions. **B:** Potential secondary structure formed on $L4^E$:$I4^E$ molecules resulting from the use of the branch site at position −36. This structure may be stabilized by the CC to GG mutation at position +7 and +8 in the E\78**GG** mutant, and even more so by a single G to C mutation at position +7 in the E\78GC mutant (see Fig. 4A and text for details). Å indicates the nucleotide used as a branch point.

migration of lariat molecules (data not shown). When the 5′ portion of the E\78**GG** mutant was combined with the 3′ portion of the E\L3**CC** mutant to produce an RNA substrate in which complementarity was restored, the L3:I3 lariat molecules migrated more rapidly (Fig. 4, lane 18). Control combinations that should not reform the secondary structure displayed the expected slower migration of L3:I3 lariat species (data not shown).

Duplex formation affects 5′ splice site selection

Our results indicate that lariat molecules produced from E18/E19 splicing contain a highly stable structure that accounts for their faster migration in denaturing gels. If this structure forms only in lariat molecules, it may not play an important role in splicing. However, given that the 5′ splice site and branch site sequences are brought in close proximity at the commitment stage of spliceosome assembly (Fig. 3A), base pairing interactions may occur before branch formation and could

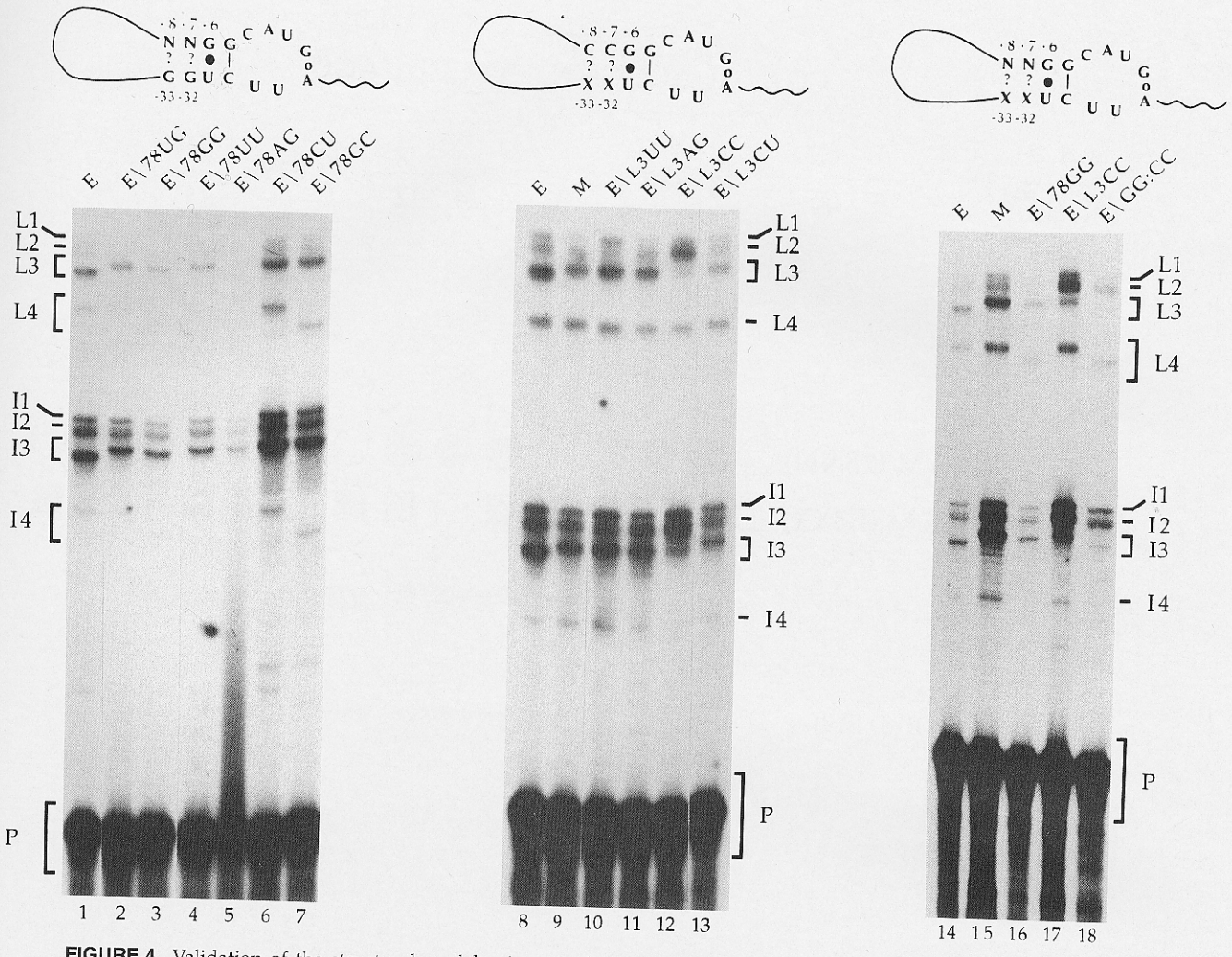

FIGURE 4. Validation of the structural model using site-directed mutagenesis. Random mutations were introduced at positions +7 and/or +8 of the intron (E\78NN mutants) and at positions −32 and −33 relative to the 3′ splice site of E19 (E\L3XX mutants). The 78**GG** mutation was combined with the L3**CC** mutation to generate a complementary pair (lane 18). The wild-type E RNA and mutated RNA substrates were incubated in HeLa nuclear extracts for 2 h at 30 °C. Splicing intermediates and products were resolved onto 9% polyacrylamide/7 M urea gels. The position of pre-mRNA (P), lariat intermediates (L1–L4), and lariat introns (I1–I4) is indicated. Note that the 78GC mutation promoted the formation of aberrantly migrating L4:I4 species (lane 7). Although not seen from this exposure of the gel, the 78**GG** mutation (lane 3) also yielded L4:I4 species migrating faster than wild-type but slower than L4:I4 molecules produced from E\78GC RNA. A diagram illustrating the position and nature of the mutations is shown above each gel.

affect spliceosome assembly. To address whether these base pairing interactions play a role in 5′ splice site selection, we tested the effect of various mutations in the context of C RNA, a substrate that contains the 5′ splice sites of NCAM E17 and E18 in competition for the 3′ splice site of E19 (Fig. 5A). Following incubation in NIH3T3 extracts, we used an RT-PCR assay to measure the relative abundance of the NCAM mRNA species produced in vitro. In this assay, C RNA splicing occurred preferentially to the distal E17 5′ splice site (Fig. 5B, lanes 3 and 4). In contrast, the G to U mutation at position +6 of the E18 5′ splice site strongly favored E18/E19 splicing (C\6**U** RNA; Fig. 5B, lane 5). Because the 6**U** mutation increases the match of the proximal 5′ splice site to the consensus, more efficient E18/E19 splicing may be due, in part or completely, to improved U1 snRNP binding. To evaluate the contribution of intron secondary structure to 5′ splice site selection, we tested C RNA derivatives carrying mutations at positions +7/+8 or −33/−32 of the E intron as well as combinations that restored complementarity between the E18 5′ splice site and the branch site regions. Strikingly, all mutations in the branch site region at positions −32 (C\L3**AG** and C\L3**GC**; Fig. 5B, lanes 12 and 18, respectively) or −33/−32 (C\L3**CU** and C\L3**CC**; lanes 8 and 15, respectively) increased the frequency of proximal E18 5′ splice site utilization. The amplitude of the shift obtained with the L3**CU**, L3**AG**, and L3**CC** mutations was as dramatic as the effect seen with the 6**U** mutation. This result indicates that changes in the sequence bordering a branch site can affect 5′ splice site selection. The specific increase

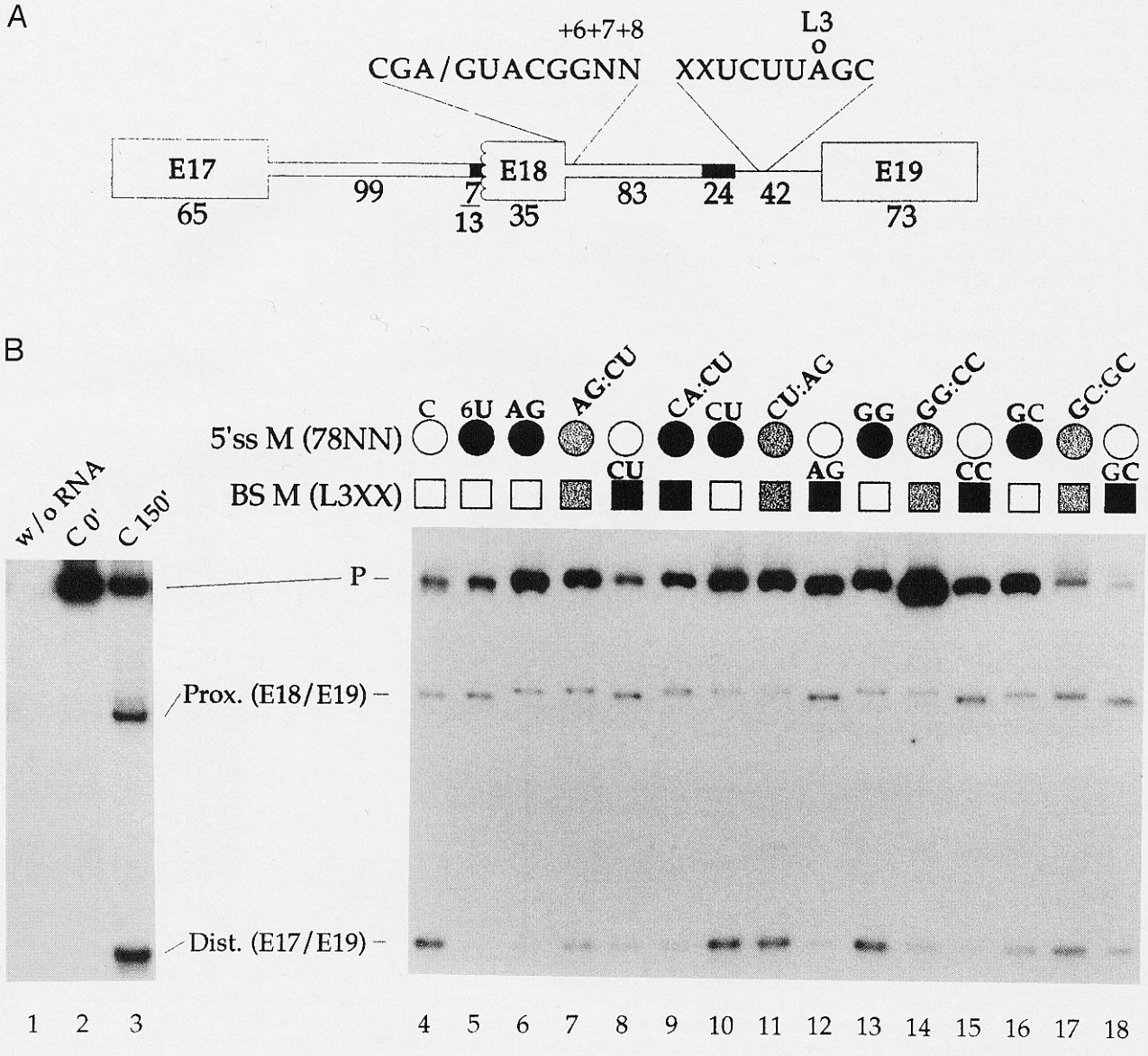

FIGURE 5. Duplex formation affects 5′ splice site selection. **A:** Structure of the C RNA substrate. C RNA corresponds to E RNA with an additional upstream portion containing part of exon 17 and flanking intron sequences (170 nt). Note that the cloning procedure yielded C\78NN and C\NN:XX substrates 6 nt longer than the C\L3XX RNAs. The additional nucleotides are located just upstream of the E18 exon sequences. **B:** Splicing in NIH3T3 extracts incubated with "cold" pre-mRNA substrates was analyzed by RT-PCR. The position of the fragments amplified from the pre-mRNA (P), or from mRNAs generated from the proximal (E18) or the distal (E17) 5′ splice site is indicated. The left panel represents incubations performed in the absence of C RNA (lane 1) or with C RNA incubated 0 min (lane 2) or 150 min (lane 3) in NIH3T3 extracts. A white circle or square above the lanes denotes the presence of the wild-type 5′ splice site (5′ss) or branch site (BS) sequence, respectively. 5′ Splice site and branch site mutations are indicated by a black circle and a black square above corresponding lanes. Gray circles and squares indicate compensatory mutations. **C:** Histogram representing the quantitation of several RT-PCR assays and standard splicing gels for several mutants. Values express the percentage of proximal E18/E19 splicing relative to total splicing products. For splicing gels, the levels of proximal and distal lariat molecules were used for quantitation. Gels were scanned using Adobe Photoshop and the amounts of proximal and distal spliced products were determined using the NIH Image program. Standard deviations are indicated when values represent the average of three independent experiments. Other values represent the average of two independent experiments.

in E18/E19 splicing promoted by all branch site mutations is therefore consistent with the notion that duplex formation specifically impairs E18/E19 splicing. This panel of branch site mutations was combined with mutations at positions +7/+8 to restore complementarity between the E18 5′ splice site and the branch site regions. In all cases, the +7/+8 mutations antagonized the effect of the −33/−32 mutations to reduce proximal 5′ splice site selection (Fig. 5B, lanes 7, 11, 14, and 17, respectively). The control **CA**:**CU** noncomplementary pair did not neutralize the effect of the L3**CU** mutation (Fig. 5B, compare lane 9 with lane 8). We noted that the effects of the individual +7/+8 mutations varied for the mutations tested; although the 78**AG** mutation (Fig. 5B, lane 6) had an effect nearly as dramatic as the L3**CU** or 6**U** mutation, the 78GC (lane 16) had an intermediate effect, and the 78**CU** (lane 10) and 78**GG** (lane 13) had no effect. The reason for these differences is unclear. Inspection of the sequence suggests that 78**CU**, 78**GG**, and 78GC mutations may allow alternative base pairing interactions to occur with the natural GU-rich branch site region (not shown). Although these alternative duplex structures may not be stable enough to resist the denaturing conditions of a urea gel, their formation may affect selection of the E18 5′ splice site negatively in splicing extracts. The splicing results presented above were confirmed in several independent experiments and by direct analysis of proximal and distal lariat molecules in standard splicing gels (Fig. 5C and data not shown). Although total splicing activity varied between different RT-PCR experiments and between different samples (possibly reflecting differences in the quality of each RNA preparation), the relative distribution of proximal and distal products was reproducible. For these reasons, we favor the interpretation that mutations did not affect overall splicing activity, but rather the selection of competing 5′ splice sites. Thus, branch site mutations that disrupt base pairing always improved E18/E19 splicing, whereas including compensatory mutations at positions +7/+8 impaired E18/E19 splicing. Our results suggest that natural base pairing interactions between the 5′ splice site of NCAM E18 and downstream branch site sequences specifically decrease the efficiency of E18/E19 splicing.

Duplex formation reduces the assembly of complexes on the 5′ splice site of E18

Because mutations that alter lariat migration also affected 5′ splice site selection, duplex formation may occur early to influence splice site recognition or the assembly of splicing complexes. We used the oligonucleotide-targeted RNase H protection assay (Eperon et al., 1993; Chabot et al., 1997) to monitor the assembly of complexes at the 5′ splice site of E18. RNase H and oligonucleotides complementary to the 5′ splice sites of E17 and E18 were added following various periods of incubation (0, 15, 30 min) at 30 °C in a standard HeLa splicing reaction (Fig. 6). Compared to the wild-type C RNA (Fig. 6; lanes 1–3), protection at the 5′ splice site of E18 was more efficient with the C\6**U** RNA (lanes 4–6), consistent with an improved match of the E18 5′ splice site to the consensus. Notably, the branch site L3**CU** mutation, which disrupts structure formation and improves E18 5′ splice site utilization, also stimulated protection at the E18 5′ splice site (Fig. 6, lanes 7–9). The protection observed with all pre-mRNAs was U1 snRNP-dependent, but not ATP-dependent (data not shown). Our results suggest that structure formation interferes with the assembly of early splicing complexes at the 5′ splice site of E18.

Duplex formation reduces U2AF^{65} binding

The 5′ splice site protections monitored in the RNase H assay likely reflect interactions taking place in assembling spliceosomes. The intron base pairing interactions may be occurring as early as the commitment stage of spliceosome assembly when 5′ splice site and branch site sequences are in close proximity (Fig. 3A, top). If base pairing occurs at this stage, the resulting structure could hinder the stability of commitment complexes or prevent efficient progression into more advanced stages of spliceosome assembly. We have shown previously that U2AF^{65} binds inefficiently to E RNA and that the G to U mutation at position +6 improves U2AF^{65} binding (Côté et al., 1995). To verify whether other mutations that disrupt the duplex structure also improve U2AF^{65} binding, simple RNA substrates were incubated in standard HeLa splicing reactions and subjected to UV irradiation. Compared to E RNA (Fig. 7, lanes 1, 4, and 8), U2AF^{65} crosslinking was stimulated by mutations at positions +7/+8 (Fig. 7, lanes 2 and 5) as well as by the branch site L3**CU** mutation (Fig. 7, lane 9). In contrast, the efficiency of U2AF^{65} crosslinking to an E RNA derivative carrying compensatory mutations became equivalent to the level observed with E RNA (Fig. 7, compare lanes 6 and 7 with lane 4). Similar effects were seen with other 5′ splice site and branch site mutations (data not shown). Note that the identity of the band corresponding to U2AF^{65} was verified by co-migration with purified U2AF crosslinked to D RNA (lane 3), and by transferring the labeled proteins on nitrocellulose followed by detection with an anti-U2AF antibody (data not shown, but see Côté et al., 1995). Thus, the crosslinking results suggest that base pairing between 5′ splice site sequences and the branch site region affects an early step of spliceosome assembly, namely the interaction of U2AF^{65} with the 3′ splice site of E19.

Secondary structure formation in another pre-mRNA substrate

To address whether some of the features that characterize NCAM splicing could be observed in other splic-

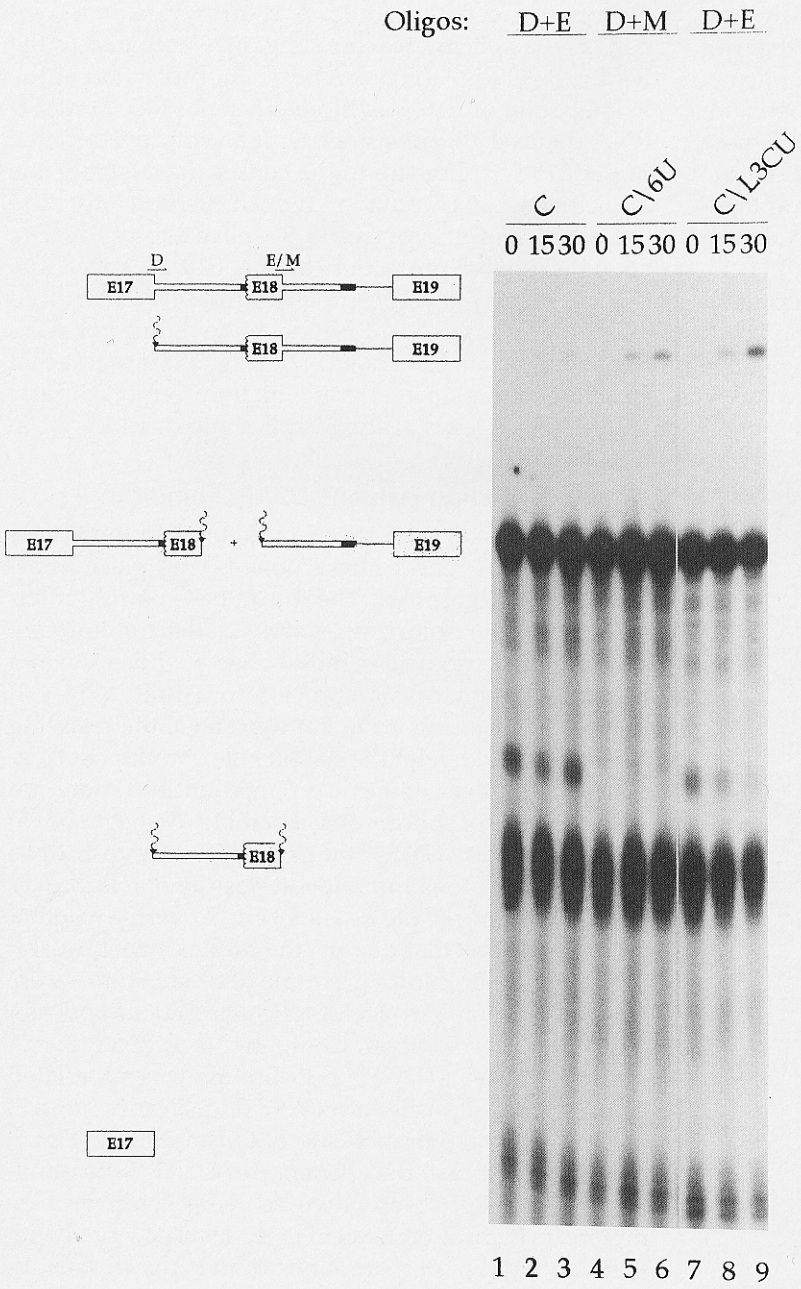

FIGURE 6. Duplex formation reduces the assembly of U1-dependent complexes on the 5′ splice site of E18. Each RNA substrate was incubated in HeLa nuclear extract under splicing conditions for the time indicated (in min) (lanes 1–9). Following incubation, oligonucleotides complementary to each 5′ splice site were added and the mixture was incubated for 15 min in the presence of RNase H. The position and the identity of the cleavage products are indicated on the left. The combination of oligonucleotides (D, E, or M) used for each substrate is indicated above the panels. D is directed against the 5′ splice site of E17, E is complementary to the wild-type 5′ splice site of E18, and M to the 6U mutated version of the E18 5′ splice site. The position of the oligos is indicated above the fully protected pre-mRNA.

ing units, we tested an α-globin pre-mRNA (G RNA). In this substrate, sequences at positions +5 to +9 can potentially base pair with sequences upstream of the branch site to yield five contiguous G:C base pairs (Fig. 8A). A derivative of G RNA (GM RNA) contains mutations that reduce the match to the consensus 5′ splice site and disrupt the complementarity with branch site sequences. Strikingly, incubation of G and GM RNAs in HeLa extracts led to the production of lariat molecules that did not co-migrate even when formamide and urea were present in the gel (Fig. 8B). Characterization of the lariat molecules indicated that the introns released from the two substrates were of identical length and shared a common branch site (Fig. 8C and data not shown). These results suggest that a highly stable duplex structure is formed in lariat molecules derived from G RNA. Because mutations that retarded lariat migration improved U2AF^{65} binding (Fig. 8D), duplex formation in the globin intron may also occur early to reduce the efficiency of U2AF^{65} binding.

DISCUSSION

A simple model pre-mRNA (E RNA) containing part of NCAM exons E18 and E19 and separated by a shortened intron was spliced in vitro through the use of five distinct branch sites located at positions −21, −23, −24, −27, and −36 from the 3′ splice junction. The

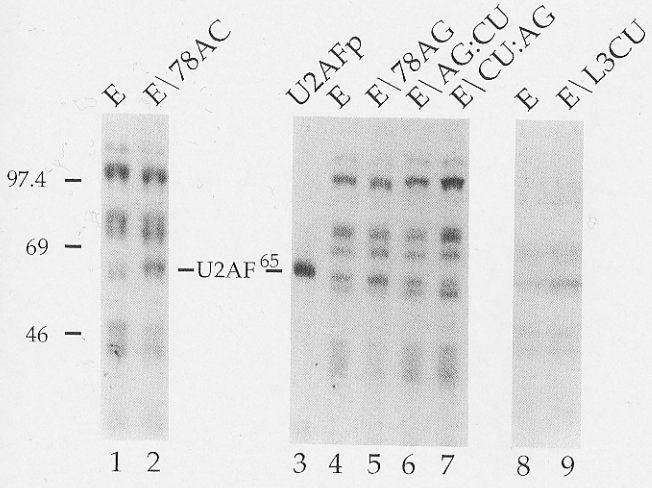

FIGURE 7. Structure formation promotes inefficient U2AF^{65} binding. UV crosslinking to E RNA and various mutated RNAs was performed following a 4-min incubation at 30 °C in a HeLa nuclear extract. Crosslinking was also performed with E RNA and a fraction containing purified U2AF (lane 8). The position of U2AF^{65} and molecular weight markers (in kDa) is indicated. Note that lanes 1–2, lanes 3–7, and lanes 8–9 represent three different crosslinking assays.

branch site at position −27 yielded L3 and I3, the most abundant lariat intermediate and product, respectively. A single mutation at position +6, converting the 5′ splice site of E18 into that of E17 (M RNA), changed the migration of the L3 and I3 molecules to slower-migrating forms in polyacrylamide/urea gels. The migration of the other four sets of lariat species remained unaffected by the mutation. The difference in migration between E and M lariat molecules was not due to the use of different splicing signals. Instead, base pairing interactions between sequences at intron positions +5 to +8 and sequences at positions −30 to −33 from the 3′ splice site were responsible for the faster migration of the L3 and I3 lariat molecules produced from E RNA splicing.

The existence of a stable structure in lariat molecules raised the possibility that the base pairing interactions occur earlier when 5′ splice site and branch site sequences are brought in close proximity during spliceosome assembly. If so, duplex formation may affect E18/E19 splicing. When tested in the context of a model pre-mRNA that contains the 5′ splice site of E17 and E18 in competition for the 3′ splice site of E19, all branch site mutations that disrupted complementarity improved the selection of the E18 5′ splice site. Although the branch site region has been shown to influence 5′ splice site choice in yeast (Goguel & Rosbash, 1993), our results indicate that mutations in the branch site region can also affect mammalian 5′ splice site selection. Combining branch site mutations with 5′ splice site mutations to restore complementarity always counteracted the effect the individual branch site mutations and decreased the relative efficiency of E18/E19 splicing. Whether the mutations similarly affect NCAM splicing in vivo is being investigated currently. Our in vitro results therefore suggest that the base pairing interactions are most likely established early to compromise E18/E19 splicing.

The results of the RNase H protection assay suggest that duplex formation interferes with the assembly of U1-dependent complexes on the 5′ splice site of E18. Mutations at positions +7/+8 or at positions −32/−33 also improved U2AF^{65} binding. In contrast, combining mutations to restore complementarity reduced U2AF^{65} binding. Because U2AF binding is normally destabilized during the transition from the E to the B complex (Bennett et al., 1992; Staknis & Reed, 1994), our data suggest that duplex formation occurs early, possibly as soon as the 5′ splice site and branch site regions are brought in proximity during commitment complex formation (Fig. 3A, top). Duplex formation may destabilize U1 snRNP binding through the appropriation of sequences normally paired with U1 RNA. This in turn would promote the dissociation of U2AF^{65}, which binds only weakly to the 3′ splice site of E19 in the absence of U1 snRNP (Côté et al., 1995). Duplex formation would therefore compromise or delay further spliceosome assembly. Because a reduction in E18/E19 splicing was detected only in the context of competing 5′ splice sites, duplex formation may favor kinetically spliceosome assembly with the E17 5′ splice site, resulting in more efficient E17/E19 splicing.

Our model NCAM C RNA substrate was spliced efficiently to the E17 distal 5′ splice site in NIH3T3 extracts. In contrast, the proximal 5′ splice site of E18 is used almost exclusively in mouse and human neuronal cell extracts (J. Côté & B. Chabot, unpubl. results). Because this effect is specific to the NCAM pre-mRNA substrate, neuronal extracts therefore appear to be considerably less sensitive to the inhibitory effect of the intron structure. The ability of neuronal extracts to overcome the restriction imposed by the intramolecular base pairing interactions may be due to the presence in these extracts of an activity that prevents or transiently destabilizes duplex formation. Alternatively, neuronal extracts may contain specific factors that allow efficient and stable splicing complex formation despite the presence of the secondary structure. Characterizing the neuronal factors responsible for promoting E18/E19 splicing in spite of an interfering secondary structure should lead to new insights into the mechanisms controlling alternative splicing.

A number of studies in vertebrate and yeast systems have indicated a role for pre-mRNA structures in splice site selection (see Introduction). Base pairing interactions between yeast intron sequences downstream from the 5′ splice site and sequences upstream from the branch site have been documented (Goguel & Rosbash, 1993), and putative helix-forming sequences are found at these positions in several yeast introns (Newman, 1987; Parker & Patterson, 1987). Such intron base

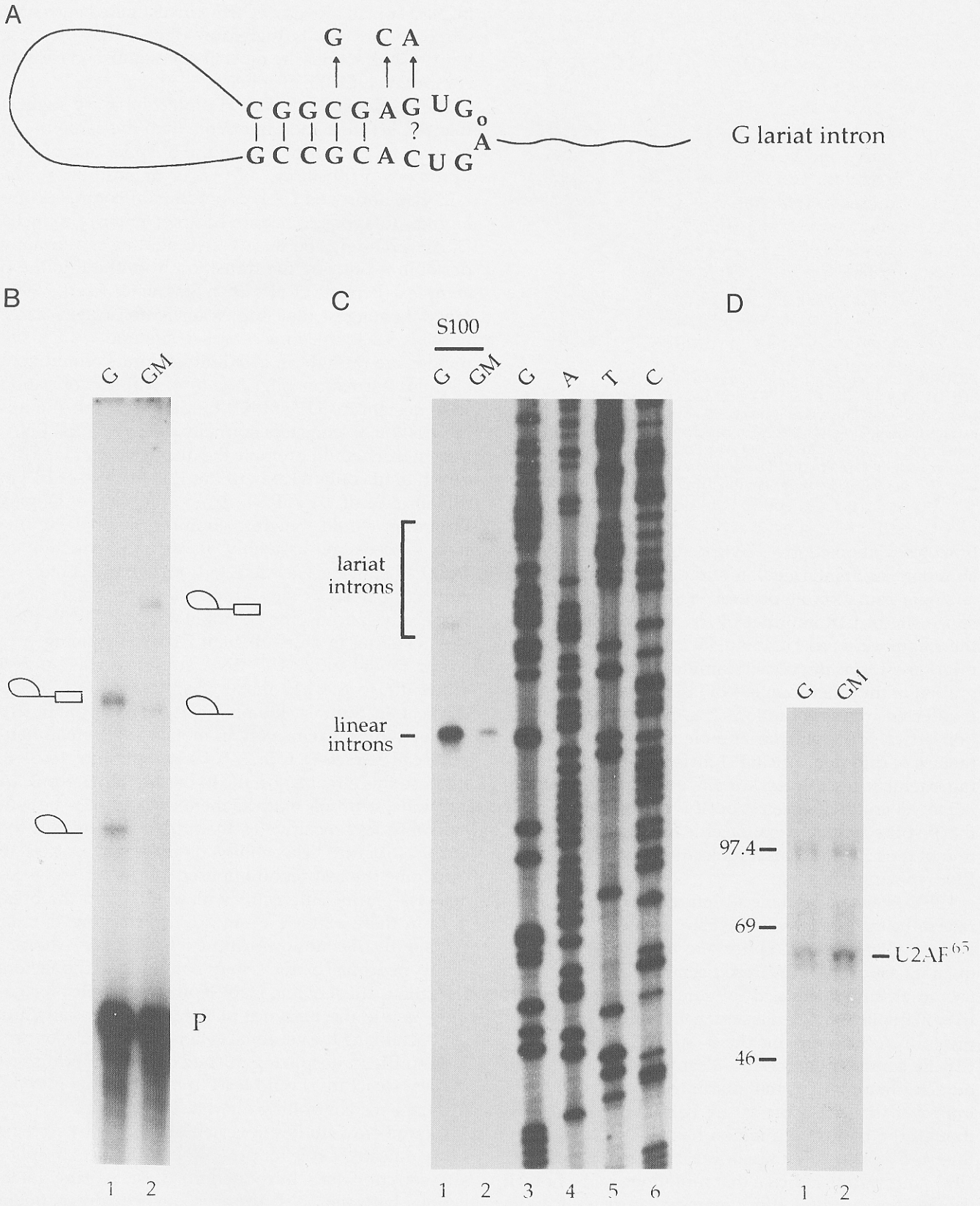

FIGURE 8. Structure formation in an α-globin pre-mRNA substrate. **A:** Putative base pairing interactions occurring in lariat molecules produced from G RNA. The mutations in GM RNA are indicated. **B:** ^{32}P-labeled transcripts were incubated for 2 h in HeLa nuclear extracts under splicing conditions. Splicing products were separated on a 13% acrylamide/8 M urea/20% formamide gel. Note the distinct mobility of the respective lariat intermediates and products above the pre-mRNA. **C:** Size determination of the G and GM intron molecules. A debranching reaction in a HeLa S100 extract was performed with each lariat intron following its purification from a denaturing gel. Debranching reactions were loaded onto a 6% acrylamide/8 M urea gel. The position of the lariat and linear introns is indicated relative to a DNA sequencing ladder (lanes 3–6). **D:** Structure formation is associated with inefficient U2AF^{65} crosslinking. UV crosslinking to G and GM RNA was performed in a HeLa nuclear extract as described previously.

pairing interactions promote more efficient splicing by reducing the effective distance that separates splice sites (Libri et al., 1995; Charpentier & Rosbash, 1996), or by improving the presentation of the 5′ splice site (Newman, 1987). The 5′ splice site/branch site interactions in the NCAM intron represent a unique example of negative modulation of 5′ splice selection by duplex formation. Commitment complex formation may be required to bring the interacting nucleotides in proximity but, once established, the structure would affect the assembly of productive spliceosomes.

The 5′ splice site and branch site sequences may have a general propensity to associate due to their proximity in splicing complexes before and following branch formation. Although only duplex structures with exceptional stability may be detected in the usual denaturing conditions of splicing gels, base pairing interactions between 5′ splice site and branch site sequences may turn out to be a feature of several mammalian introns. Interestingly, mutations in the 5′ splice site region of an α-globin substrate also led to slower-migrating lariat molecules that had used the same branch and splice sites. Base pairing interactions between the 5′ splice site sequences and sequences upstream of the branch site can also explain the migration behavior of α-globin lariat molecules. As seen with the NCAM intron, mutations that destabilized the globin structure improved U2AF^{65} binding.

It is notable that positions +7 to +10 in identical introns of different species are often conserved. Because sequences at these positions are not conserved between different introns and are not part of the consensus 5′ splice site (Mount, 1982), it is possible that the evolutionary pressure that led to their conservation was driven by their contribution to base pairing interactions with sequences near branch sites. A preliminary survey of known 5′ splice site and branch site sequences revealed that similar duplex structures may exist in other alternative splicing units. Future efforts will be directed at assessing the prevalence of such structures and their contribution to splice site selection. Although we have documented the negative effect that duplex interactions can have, it is possible that, in some cases, similar interactions in mammalian pre-mRNAs provide additional stability to splicing complexes. In either case, the existence of this type of interaction may contribute toward clarifying the notion of context and hierarchical compatibility between splice sites (Gallego & Nadal-Ginard, 1990; Lear et al., 1990).

MATERIALS AND METHODS

Plasmids, mutagenesis, and RNA substrates

pSPM, pSPW, and pSPD were described previously (Côté et al., 1995). pSPE contains a *Pst* I–*Taq* I fragment containing the 5′ splice site region of E18 followed by a *Sac* I–*Eco*R I fragment containing part of E19 and the upstream intron. pSPC and pSPC\6U were generated by inserting a *Pst* I fragment containing 65 nt of E17 and its 5′ splice site in pSPE and pSPM, respectively. Site-directed mutagenesis was achieved by overlap extension using PCR (Ho et al., 1989). The overlapping mutagenic oligonucleotides used for creating the E\78NN and E\L3XX mutants were (5′ to 3′) E78Na, ACCGAGTACGGNNTCTCTTTGT; L3Na, AGTNNTCTTAG CAACACTG; and the corresponding complements (E78Nb and L3Nb). The flanking oligonucleotides were SP6, TTGTC GTTAGAACGCGGCTA; and Nf1, TTGCTTGGTACCCATC ATGC. The mutagenized E\78NN and E\L3XX fragments were substituted into pSPE at the *Pst* I site. The compensatory mutants were obtained by substituting *Pst* I fragments harboring E\L3XX mutations into the complementary E\78NN mutants. *Hin*d III fragments containing E\78NN mutations and *Sac* I–*Eco*R I fragments containing E\L3XX mutations were substituted into pSPC to generate C RNA derivatives. The E\L3XX fragments were also substituted into C\78NN mutants to generate compensatory C derivatives. G and GM substrates were kindly provided by Roland Tacke. The identity of all mutated recombinant molecules was confirmed by DNA sequencing. All plasmids were linearized at the *Ava* II site prior to in vitro transcription, except G and GM, for which *Eco*R I was used.

In vitro transcription and splicing reactions

Splicing substrates were synthesized using the SP6 RNA polymerase from corresponding linear templates, in the presence of CAP analogue and [α-^{32}P] UTP (Chabot, 1994). Full-length transcripts were gel-purified before use. HeLa and NIH3T3 cell nuclear extracts were prepared as described (Dignam et al., 1992). Splicing reactions were set up and processed as described previously (Lavigueur et al., 1993), except that NIH3T3 extracts were complemented with 1 U of creatine kinase. Labeled RNA molecules were either separated on polyacrylamide/urea gels (acrylamide:bis-acrylamide, 38:2) directly or submitted to RT-PCR amplification. RT-PCR was done using a one-tube procedure. The lower layer contained 15 pmol of each primer (Nd2: 5′-CCAAACCATGATG GGGGGAA-3′; Nf2: 5′-ACTTCAGTTGGCGCTGGCTT-3′), 1.38 mM $MgCl_2$, 2.9 μg BSA (RNase/DNase free; Pharmacia), 1× GeneAmp® PCR buffer II (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 3 μCi [α-^{32}P]dCTP (Amersham), and 5 U of Taq DNA polymerase (Pharmacia) completed to 80 μL with H_2O. The upper layer contained 4 mM DTT, 2 mM $MgCl_2$, 1.25 mM dNTPs, 1× GeneAmp® PCR buffer II, 8 U of RNAguard (Pharmacia), 15 pmol of Nf2 primer, 200 U of Mo-MuLV Reverse Transcriptase (Promega), and spliced RNA completed to 20 μL with H_2O. Layers were separated using AmpliWax™ PCR Gem 100 (Perkin Elmer). The reverse transcriptase step was done at 42 °C for 45 min, then, following a 3-min incubation at 95 °C, a 2-step PCR amplification (94 °C for 30 s and 65 °C for 1 min) was performed for 30 cycles.

Debranching and primer extension assays

Gel-purified lariat intron (0.5 μL) was incubated in 25 μL of buffer D (20 mM Hepes-KOH, pH 8.0, 20% glycerol, 100 mM KCl, 0.5 mM DTT) containing 40% (v/v) HeLa S100 extract and 8 mM Na_2EDTA. Incubation was at 30 °C for 90 min (Ruskin & Green, 1990).

To map the sites of branch formation, lariat intermediates were gel-purified and subjected to primer extension analysis. The Nf2 oligonucleotide complementary to E19 was 5′-end labeled with [γ-^{32}P]-ATP and purified on a 15% polyacrylamide/8 M urea gel. Annealing of $\approx 1 \times 10^6$ cpm (Cerenkov) of primer and $\approx$600 cpm ($\approx$15 pmol) of lariat intermediate was performed at 50 °C for 3 min followed by 30 °C for 30 min in 8 μL total of buffer H containing 300 mM NaCl, 40 mM Tris-HCl, pH 8.0, and 0.1 mM Na_2EDTA. Reactions were then initiated by adding an equal volume of buffer to give final concentrations of 100 mM Tris-HCl, pH 8.0, 12 mM $MgCl_2$, 10 mM DTT, 0.75 mM of each dNTP, 15 units of RNase inhibitor (Pharmacia), and 7.5 units of avian myeloblastosis virus reverse transcriptase (Promega). After 45 min at 42 °C, the reactions were stopped by adding 0.7 volume of 12 mM Na_2EDTA/5 M ammonium acetate prior to ethanol precipitation. The RNA was hydrolyzed by redissolving the pellet in 0.5 M NaOH/0.1 mM Na_2EDTA and incubation at 65 °C for 10 min, followed by neutralization with 0.7 volume of 5 M ammonium acetate. Extension products were then precipitated with ethanol, redissolved, and loaded on an 8% polyacrylamide/8 M urea gel.

RNase H protection

The RNA substrates (1 fmol) were incubated for the time indicated in 10 μL HeLa splicing reactions supplemented with 0.44 U of RNase H (Pharmacia). Oligonucleotides (100 pmol) complementary to the 5′ splice sites of E17 and E18 (D, 5′-CCACGTACTCGGGT-3′ and E, 5′-GGCCGTACTCG GTC-3′, respectively, and M, 5′-GGACGTACTCGGTC-3′ for the C\6U RNA) were then added and the incubation continued at 30 °C for 15 min. Labeled RNA fragments were resolved on 8% polyacrylamide/8 M urea gels.

U2AF^{65} crosslinking

For UV crosslinking assays, RNA substrates were synthesized in the presence of 5-BrUdr (Boehringer-Mannheim) and incubated in a standard splicing reaction without RNase inhibitors, but in the presence of 2.5 mM EDTA. A 5-μL aliquot was removed from each reaction and irradiated 20 min with UV and digested with RNase A as described (Côté et al., 1995). Crosslinking products were analyzed by electrophoresis on 9% SDS-polyacrylamide gels.

ACKNOWLEDGMENTS

We thank R. Tacke for the α-globin plasmids, and R. Reed for helpful comments on the manuscript. We thank J. Toutant for help in the preparation of NIH3T3 extracts. This work was supported by a grant to B.C. from the Medical Research Council of Canada. B.C. is a Chercheur-Boursier Senior of the FRSQ. J.C. is the recipient of a studentship from the FCAR.

Received June 27, 1997; returned for revision July 25, 1997; revised manuscript received August 12, 1997

REFERENCES

Abovich N, Rosbash M. 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell 89*:403–412.

Adams MD, Tarng RS, Rio DC. 1997. The alternative splicing factor PSI regulates P-element third intron splicing in vivo. *Genes & Dev 11*:129–138.

Amendt BA, Hesslein D, Chang LJ, Stoltzfus CM. 1994. Presence of negative and positive cis-acting RNA splicing elements within and flanking the first tat coding exon of human immunodeficiency virus type 1. *Mol Cell Biol 14*:3960–3970.

Barbas JA, Chaix JC, Steinmetz M, Goridis C. 1988. Differential splicing and alternative polyadenylation generates distinct NCAM transcripts and proteins in the mouse. *EMBO J 7*:625–632.

Bennett M, Michaud S, Kingston J, Reed R. 1992. Protein components specifically associated with prespliceosome and spliceosome complexes. *Genes & Dev 6*:1986–2000.

Black DL. 1995. Finding splice sites within a wilderness of RNA. *RNA 1*:763–771.

Blanchette M, Chabot B. 1997. A highly stable duplex structure sequesters the 5′ splice site region of hnRNP A1 alternative exon 7B. *RNA 3*:405–419.

Burke JF, Bright KE, Kellett E, Benjamin PR, Saunders SE. 1992. Alternative mRNA splicing in the nervous system. *Prog Brain Res 92*:115–125.

Caputi M, Casari G, Guenzi S, Tagliabue R, Sidoli A, Melo CA, Baralle FE. 1994. A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. *Nucleic Acids Res 22*:1018–1022.

Carlo T, Sterner DA, Berget SM. 1996. An intron splicing enhancer containing a G-rich repeat facilitates inclusion of a vertebrate micro-exon. *RNA 2*:342–353.

Chabot B. 1994. Synthesis and purification of RNA substrates. In: Hames D, Higgins S, eds. *RNA processing, vol I*. Oxford: Oxford University Press. pp 1–29.

Chabot B. 1996. Directing alternative splicing: Cast and scenarios. *Trends Genet 12*:472–478.

Chabot B, Blanchette M, Lapierre I, 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 RC, Black DL. 1995. Conserved intron elements repress splicing of a neuron-specific c-src exon in vitro. *Mol Cell Biol 15*:6377–6385.

Charpentier B, Rosbash M. 1996. Intramolecular structure in yeast introns aids the early steps of in vitro spliceosome assembly. *RNA 2*:509–522.

Chebli K, Gattoni R, Schmitt P, Hildwein G, Stevenin J. 1989. The 216-nucleotide intron of the E1A pre-mRNA contains a hairpin structure that permits utilization of unusually distant branch acceptors. *Mol Cell Biol 9*:4852–4861.

Clouet O, d'Aubenton Carafa Y, Sirand-Pugnet P, Gallego M, Brody E, Marie J. 1991a. RNA secondary structure repression of a muscle-specific exon in HeLa cell nuclear extracts. *Science 252*:1823–1828.

Clouet O, d'Aubenton-Carafa Y, Brody JM, Brody E. 1991b. Determination of an RNA structure involved in splicing inhibition of a muscle-specific exon. *J Mol Biol 221*:837–856.

Côté J, Beaudoin J, Tacke R, Chabot B. 1995. The U1 small nuclear ribonucleoprotein/5′ splice site interaction affects U2AF65 binding to the downstream 3′ splice site. *J Biol Chem 270*:4031–4036.

Del Gatto F, 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.

Deshler JO, Rossi JJ. 1991. Unexpected point mutations activate cryptic 3′ splice sites by perturbing a natural secondary structure within a yeast intron. *Genes & Dev 5*:1252–1263.

Dignam JD, Lebovitz RM, Roeder RG. 1992. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res 11*:1475–1489.

Eperon IC, Ireland DC, Smith RA, Mayeda A, Krainer AR. 1993. Pathways for selection of 5′ splice sites by U1 snRNPs and SF2/ASF. *EMBO J 12*:3607–3617.

Eperon LP, Estibeiro JP, Eperon IC. 1986. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. *Nature 324*:280–282.

Estes PA, Cooke NE, Liebhaber SA. 1992. A native RNA secondary structure controls alternative splice-site selection and generates two human growth hormone isoforms. *J Biol Chem 267*:14902–14908.
Fu XD. 1995. The superfamily of arginine/serine-rich splicing factors. *RNA 1*:663–680.
Gallego ME, Nadal-Ginard B. 1990. Myosin light-chain 1/3 gene alternative splicing: Cis regulation is based upon a hierarchical compatibility between splice sites. *Mol Cell Biol 10*:2133–2144.
Goguel V, Rosbash M. 1993. Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. *Cell 72*:893–901.
Goguel V, Wang Y, Rosbash M. 1993. Short artificial hairpins sequester splicing signals and inhibit yeast pre-mRNA splicing. *Mol Cell Biol 13*:6841–6848.
Gontarek RR, McNally MT, Beemon K. 1993. Mutation of an RSV intronic element abolishes both U11/U12 snRNP binding and negative regulation of splicing. *Genes & Dev 7*:1926–1936.
Halfter H, Gallwitz D. 1988. Impairment of yeast pre-mRNA splicing by potential secondary structure-forming sequences near the conserved branchpoint sequence. *Nucleic Acids Res 16*:10413–10423.
Helfman DM, Roscigno RF, Mulligan GJ, Finn LA, Weber KS. 1990. Identification of two distinct intron elements involved in alternative splicing of β-tropomyosin pre-mRNA. *Genes & Dev 4*:98–110.
Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene 77*:51–59.
Hoffman BE, Grabowski PJ. 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3′ splice site by a network of interactions spanning the exon. *Genes & Dev 6*:2554–2568.
Huh GS, Hynes RO. 1994. Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. *Genes & Dev 8*:1561–1574.
Humphrey MB, Bryan J, Cooper TA, Berget SM. 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, Mühlemann O, Akusjärvi G. 1996. Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature 381*:535–538.
Kister L, Domenjoud L, Gallinaro H, Monique J. 1993. A cis-acting selector of a 5′ splice site. *J Biol Chem 268*:21955–21961.
Kohtz JD, Jamison SF, Will CL, Zuo P, Lührmann R, Garcia-Blanco MA, Manley JL. 1994. Protein–protein interactions and 5′-splice-site recognition in mammalian mRNA precursors. *Nature 368*:119–124.
Lavigueur A, La Branche H, Kornblihtt AR, 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.
Lear AL, Eperon LP, Wheatley IM, Eperon IC. 1990. Hierarchy for 5′ splice site preference determined in vivo. *J Mol Biol 211*:103–115.
Libri D, Piseri A, Fiszman MY. 1991. Tissue-specific splicing in vivo of the β-tropomyosin gene: Dependence on an RNA secondary structure. *Science 252*:1842–1845.
Libri D, Stutz F, McCarthy T, Rosbash M. 1995. RNA structural patterns and splicing: Molecular basis for an RNA-based enhancer. *RNA 1*:425–436.
Lin CH, Patton JG. 1995. Regulation of alternative 3′ splice site selection by constitutive splicing factors. *RNA 1*:234–245.
Liu HX, Goodall GJ, Kole R, Filipowicz W. 1995. Effects of secondary structure on pre-mRNA splicing: Hairpins sequestering the 5′ but not the 3′ splice site inhibit intron processing in *Nicotiana plumbaginifolia*. *EMBO J 14*:377–388.
Manley JL, Tacke R. 1996. SR proteins and splicing control. *Genes & Dev 10*:1569–1579.
McKeown M. 1992. Alternative mRNA splicing. *Annu Rev Cell Biol 8*:133–155.
Min H, Chan RC, Black DL. 1995. The generally expressed hnRNP F is involved in a neural-specific pre-mRNA splicing event. *Genes & Dev 9*:2659–2671.
Min H, Turck CW, Nikolic JM, Black DL. 1997. A new regulatory protein, KSRP, mediates exon inclusion through an intronic splicing enhancer. *Genes & Dev 11*:1023–1036.
Moore MJ, Query CC, Sharp PA. 1993. Splicing of precursors to messenger RNAs by the spliceosome. In: Gesteland R, Atkins J, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 1–30.
Mount SM. 1982. A catalogue of splice junction sequences. *Nucleic Acids Res 10*:459–472.
Newman A. 1987. Specific accessory sequences in *Saccharomyces cerevisiae* introns control assembly of pre-mRNAs into spliceosomes. *EMBO J 6*:3833–3839.
Nilsen TW. 1994. RNA–RNA interactions in the spliceosome: Unraveling the ties that bind. *Cell 78*:1–4.
Parker R, Patterson B. 1987. Architecture of fungal introns: Implications for spliceosome assembly. In: Dudock B, Inouye M, eds. *Molecular biology of RNA. New perspectives*. New York: Academic Press. pp 133–150.
Reed R. 1996. Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr Opin Genet Dev 6*:215–220.
Reed R, Maniatis T. 1986. A role for exon sequences and splice-site proximity in splice-site selection. *Cell 46*:681–690.
Ruskin B, Green MR. 1990. RNA lariat debranching enzyme as tool for analyzing RNA structure. *Methods Enzymol 181*:180–188.
Sharp PA. 1994. Split genes and RNA splicing. *Cell 77*:805–815.
Siebel CW, Admon A, Rio DC. 1995. Soma-specific expression and cloning of PSI, a negative regulator of P element pre-mRNA splicing. *Genes & Dev 9*:269–283.
Siebel CW, Kanaar R, Rio DC. 1994. Regulation of tissue-specific P-element pre-mRNA splicing requires the RNA-binding protein PSI. *Genes & Dev 8*:1713–1725.
Sirand-Pugnet P, Durosay P, Brody E, Marie J. 1995a. An intronic (A/U)GGG repeat enhances the splicing of an alternative intron of the chicken β-tropomyosin pre-mRNA. *Nucleic Acids Res 23*:3501–3507.
Sirand-Pugnet P, Durosay P, Clouet d'Orval B, Brody E, Marie J. 1995b. β-Tropomyosin pre-mRNA folding around a muscle-specific exon interferes with several steps of spliceosome assembly. *J Mol Biol 251*:591–602.
Smith CW, Patton JG, Nadal-Ginard B. 1989. Alternative splicing in the control of gene expression. *Annu Rev Genet 23*:527–577.
Solnick D. 1985. Alternative splicing caused by RNA secondary structure. *Cell 43*:667–676.
Solnick D, Lee SI. 1987. Amount of RNA secondary structure required to induce an alternative splice. *Mol Cell Biol 7*:3194–3198.
Staffa A, 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, 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.
Stamm S, Zhang MQ, Marr TG, Helfman DM. 1994. A sequence compilation and comparison of exons that are alternatively spliced in neurons. *Nucleic Acids Res 22*:1515–1526.
Tacke R, Goridis C. 1991. Alternative splicing in the neural cell adhesion molecule pre-mRNA: Regulation of exon 18 skipping depends on the 5′-splice site. *Genes & Dev 5*:1416–1429.
Valcárcel J, Singh R, Zamore PD, Green MR. 1993. The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature 362*:171–175.
Vilardell J, Warner JR. 1994. Regulation of splicing at an intermediate step in the formation of the spliceosome. *Genes & Dev 8*:211–220.
Wang Z, Hoffmann HM, Grabowski PJ. 1995. Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA 1*:21–35.
Watakabe A, Sakamoto H, Shimura Y. 1991. Repositioning of an alternative exon sequence of mouse IgM pre-mRNA activates splicing of the preceding intron. *Gene Expr 1*:175–184.
Wu JY, Maniatis T. 1993. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell 75*:1061–1070.
Zhao Q, Schoborg RV, Pintel DJ. 1994. Alternative splicing of pre-mRNAs encoding the nonstructural proteins of minute virus of mice is facilitated by sequences within the downstream intron. *J Virol 68*:2849–2859.
Zuo P, Maniatis T. 1996. The splicing factor U2AF35 mediates critical protein–protein interactions in constitutive and enhancer-dependent splicing. *Genes & Dev 10*:1356–1368.
Zuo P, Manley JL. 1994. The human splicing factor ASF/SF2 can specifically recognize pre-mRNA 5′ splice sites. *Proc Natl Acad Sci USA 91*:3363–3367.