MOLECULAR AND CELLULAR BIOLOGY, May 1993, p. 2677–2687
0270-7306/93/052677-11$02.00/0

Vol. 13, No. 5

In Vivo Recognition of a Vertebrate Mini-Exon as an Exon-Intron-Exon Unit

DEBORAH A. STERNER AND SUSAN M. BERGET*

Verna and Marrs McClean Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030

Received 24 July 1992/Returned for modification 5 October 1992/Accepted 26 January 1993

Very small vertebrate exons are problematic for RNA splicing because of the proximity of their 3′ and 5′ splice sites. In this study, we investigated the recognition of a constitutive 7-nucleotide mini-exon from the troponin I gene that resides quite close to the adjacent upstream exon. The mini-exon failed to be included in spliced RNA when placed in a heterologous gene unless accompanied by the upstream exon. The requirement for the upstream exon disappeared when the mini-exon was internally expanded, suggesting that the splice sites bordering the mini-exon are compatible with those of other constitutive vertebrate exons and that the small size of the exon impaired inclusion. Mutation of the 5′ splice site of the natural upstream exon did not result in either exon skipping or activation of a cryptic 5′ splice site, the normal vertebrate phenotypes for such mutants. Instead, a spliced RNA accumulated that still contained the upstream intron. In vitro, the mini-exon failed to assemble into spliceosome complexes unless either internally expanded or accompanied by the upstream exon. Thus, impaired usage of the mini-exon in vivo was accompanied by impaired recognition in vitro, and recognition of the mini-exon was facilitated by the presence of the upstream exon in vivo and in vitro. Cumulatively, the atypical in vivo and in vitro properties of the troponin exons suggest a mechanism for the recognition of this mini-exon in which initial recognition of an exon-intron-exon unit is followed by subsequent recognition of the intron.

All intron-containing pre-mRNAs contain specific conserved sequence elements at the 5′ and 3′ splice sites that are essential for accurate and efficient splicing. In vertebrates, the branch point, pyrimidine tract, 3′ splice site, and 5′ splice site are important determinants of splicing efficiency (recently reviewed in reference 27). Although each can diverge from consensus, the degree to which each matches the consensus is related to the efficiency of splicing of the adjacent exon (19, 21, 25, 35, 44, 48, 50, 52, 71, 73–76). In certain differentially spliced genes, additional elements located upstream of the branch point (36, 58) or downstream of the 5′ splice site (7) have been shown to influence recognition of the neighboring exon. In addition, splicing efficiency in a number of systems is influenced by exon sequences (5, 6, 11, 16, 18, 20, 22, 23, 28, 29, 36, 37, 40, 51, 59, 63–65, 72). Thus, a region encompassing shortly upstream of the branch point to shortly downstream of the 5′ splice site usually contains all of the sequences necessary for recognition of an exon. We have recently suggested that the exon is the unit of splice site recognition in vertebrates (53). This perspective suggests that successful exon recognition occurs when the balance of all of these intron and exon elements is sufficient to support the definition process.

The length of the exon is also an important parameter in its recognition. The average size of vertebrate internal exons is 137 nucleotides (30). Experiments manipulating exon size have suggested that shortening an internal constitutive exon to less than 51 nucleotides results in skipping of the exon, suggesting that placement of a 3′ and 5′ splice site too close together impairs recognition (17). However, a significant number of constitutive exons are shorter than this experimentally derived exon minimum length (30). The mechanism whereby short exons are recognized is unclear. For one of these, the alternatively included, neural-specific N1 *src* exon (18 nucleotides), expansion of the exon causes loss of differential splicing and results in constitutive inclusion of the N1 exon (6).

Interestingly, there are naturally occurring constitutive exons which are extremely short (less than 12 nucleotides). Such mini-exons are frequent in genes coding for contractile proteins, including the troponin I family of muscle-specific myofibrillar proteins involved in calcium regulation of cardiac and skeletal muscle contraction (70). The chicken fast skeletal troponin I (sTNI) gene is a single-copy gene of approximately 4.5 kb which is organized into eight exons, the first of which is untranslated (45). The third exon of the sTNI gene (TGAAGAG, encoding only two complete amino acid residues) is only 7 nucleotides in length. Analysis of the extent of homology between genomic DNA encoding sTNI and cDNAs isolated from muscle RNA revealed that all cDNAs contained the mini-exon (45). Thus, despite the apparent lack of space for splicing factors to associate simultaneously with both ends of such a small exon, the mini-exon is included in a constitutive fashion in the wild-type mRNA in vivo.

We were interested in defining the mechanism by which naturally occurring vertebrate mini-exons are recognized by the cellular splicing machinery and correctly processed. We suspected the existence of essential *cis*-acting sequences that facilitate recognition of mini-exons, perhaps by extending the recognition domain. Here we report that the sTNI mini-exon is constitutively included in in vivo-spliced product RNA when inserted into a heterologous gene, but only when accompanied by the wild-type, naturally positioned upstream exon. In vitro, the mini-exon does not assemble into detectable spliceosomes unless accompanied by the upstream exon. Furthermore, mutation of the 5′ splice site terminating the upstream exon produces atypical in vivo and in vitro phenotypes consistent with the presence of an unusual recognition unit. We suggest that the sTNI 7-nucle-

* Corresponding author.

otide exon 3 is recognized as a unique exon-intron-exon unit encompassing the mini-exon and its nearby upstream exon.

MATERIALS AND METHODS

Plasmid construction. All sTNI sequences originated from plasmid RTA1, containing chicken sTNI exons 1 to 5, provided generously by Thomas Cooper. The wild-type minigenes were constructed by inserting either a 417-nucleotide *Nhe*I-*Bst*BI sTNI DNA fragment containing exons 2 and 3 (nucleotides 904 to 1320; minigene E2E3) or a 291-nucleotide *Pvu*II-*Bst*BI DNA fragment containing exon 3 flanked by 111 and 173 nucleotides of introns 2 and 3, respectively (nucleotides 1029 to 1320; minigene E3), into a modified mouse metallothionein-II (mt) gene. The recipient vector (mt1mt3) was modified prior to insertion of any sTNI DNA sequences by removing the second of three mt exons by digestion with *Sma*I while retaining mt exon 1-intron 1 (161 nucleotides)/intron 2 (67 nucleotides)-mt exon 3. Transcription is driven by a Rous sarcoma virus long terminal repeat (RSV LTR) promoter element in the recipient vector, and polyadenylation occurs at the authentic mt site in exon 3. Verification of these plasmid constructions as well as all subsequent ones was done by extensive restriction enzyme mapping and by primer extension DNA sequencing when required.

Several mutants were created by using a three-piece polymerase chain reaction (PCR) cloning strategy. In general, specific sTNI DNA fragments were amplified by using unique internal primers (X and Y) which contain the specific sequence modification and common external primers (A [5′-ttgaccattcaccacattggtgtgc-3′] and B [5′-attaggtgacactata gaatacaa-3′]). The amplified DNA fragments were then digested either with *Spe*I (mt exon 1) and *Xho*I or with *Xho*I and *Hin*dIII (pGEM-1), while the vector DNA was digested with *Spe*I and *Hin*dIII. The only exception to this procedure was mutant E3d96, in which the A-X-amplified DNA fragment was digested with *Spe*I and *Bam*HI and the Y-B-amplified DNA fragment was digested with *Bam*HI and *Hin*dIII. A three-piece cohesive end ligation was then set up overnight in the presence of T4 DNA ligase. In mutant E2E3.Xho, three nucleotides of the authentic sTNI exon 3 sequence were changed in order to generate an *Xho*I site (TGAAGAG to TCTCGAG). This was accomplished using primers X (5′-gtgg**ctcgag**actatgcagggagag-3′) and Y (5′-gtggt **ctcgag**gtaagtgtgtcctg-3′). Clone E2E3.300 has a unique *Xho*I site inserted at sTNI nucleotide 1029 by using the primers X (5′-gtgg**ctcgag**ctgaacaaaaggttg-3′) and Y (5′-gtgg**ctcgag**ctgc caacttttac-3′). In mutant E2.5′E3, the exon 2 5′ splice site has been changed from GA:**GT**AAGT to GA:**CG**AAGT by using primers X (5′-gtggctcgagggaactt**cg**tcagac3′) and Y (5′-tgttctcgagcattttctgtctccac-3′). Mutant E2E3.5′ contains an exon 3 5′ splice site mutation (**AG**:GTAAGT to **GC**:GTAAGT) and was constructed using primers X (5′-gtggctc gagctatgcagggagagt-3′) and Y (5′-gtggctcgag**cg**taagtgtgtcctg-3′).

Further truncation of the sTNI intron 2 to just 5 nucleotides upstream of the putative exon 3 branch point sequence was accomplished by deleting a 162-nucleotide *Spe*I-*Bam*HI DNA fragment from minigene E2E3 (mutant E3d96). The *Spe*I site is located 30 nucleotides upstream of the mt intron 1/sTNI nucleotide 904 in minigene E2E3 within mt intron 1. To allow truncation just upstream of the putative exon 3 branch point, a *Bam*HI site was engineered in minigene E2E3 at sTNI nucleotides 1070 and 1071, using the three-piece PCR cloning technique designed above (X [5′-gtgt**ggatcc**attggaaacaagcag-3′] and Y [5′-gtgt**ggatcc**gaagat gctaactgc-3′]).

Several sTNI sequence replacement mutants were generated for this study. Clone E3.rep is the E3 minigene with a 122-nucleotide *Kpn*I-*Bam*HI DNA fragment of calcitonin intron 3 inserted at the site of sTNI intron 2 truncation (sTNI nucleotide 1029). The exon replacement clone mtE2E3 was constructed by (i) deleting a 161-nucleotide *Spe*I-*Xho*I DNA fragment from the E2E3 minigene which spans sTNI exon 2 and (ii) substituting a 184-nucleotide *Sma*I DNA fragment of metallothionein that spans mt exon 2. Clone E2E3.74 was generated by inserting a 63-nucleotide adenosine deaminase cDNA fragment into the *Xho*I site engineered into exon 3 in clone E2E3.Xho. Clone E3.74 was constructed from E2E3.74 by deleting a 156-nucleotide *Spe*I-*Pvu*II DNA fragment, filling in with Klenow enzyme, and ligating with T4 DNA polymerase. Clone E2E3.323 was constructed by inserting a 323-nucleotide *Bam*HI DNA fragment of calcitonin intron 2 into the *Xho*I site engineered at sTNI nucleotide 1079 in the E2E3 minigene. The final replacement clone, E2E3d40, represents clone E3d96 in which the natural sTNI exon 2 was replaced to rescue inclusion of exon 3. This was done by deleting a 51-nucleotide *Xho*I-to-*Bam*HI DNA fragment from the modified E2E3 minigene containing unique *Xho*I and *Bam*HI sites at sTNI nucleotides 1029 and 1070, respectively. Finally, E3.APRT was generated by inserting the 291-nucleotide *Pvu*II-*Bst*BI sTNI DNA fragment into a unique *Eco*RI site located in intron 2 of the human adenine phosphoribosyltransferase (APRT) gene.

Primer extension DNA sequencing. The original sTNI plasmid, RTA1, was sequenced to verify its nucleotide content in comparison to that of the published sTNI genomic sequence (45). Confirmation of all plasmid constructions was also accomplished by using this sequencing protocol. Specifically, sequencing was done by the dideoxy-chain termination method on double-stranded DNA, using the Sequenase protocol as provided by the manufacturer (United States Biochemical Corp.). Verification of PCR cDNA products was done by a modified Sequenase protocol (13).

Cell culture. NIH 3T3 fibroblast cells were grown in Dulbecco's modified Eagle medium (DMEM) containing 10% bovine serum. APRT$^-$ CHO cells (cell line U1S36 was provided by John Wilson) were maintained in DMEM supplemented with 2× nonessential amino acids and 10% bovine serum.

DNA transfection. NIH 3T3 cells were transfected by the calcium phosphate precipitation method (26, 47). One T-75 flask containing cells at approximately 80% confluency was transfected with 20 μg of plasmid DNA. Total cellular RNA was isolated 48 h after transfection, using the RNAzol B method as provided by the manufacturer (Biotecx Laboratories, Inc.). APRT$^-$ cells were transfected by lipofection, using the Lipofectin reagent and protocol provided by GIBCO/BRL. One T-75 flask containing cells at approximately 60% confluency was cotransfected with 10 μg of plasmid DNA and 1 μg of plasmid pSV2neo. Drug-containing medium (400 μg of Geneticin [GIBCO/BRL] per ml in DMEM–10% bovine serum) was added to the cells 48 h after transfection. Resulting drug-resistant colonies were pooled and maintained in drug-free DMEM supplemented with 10% bovine serum. Total cellular RNA was isolated by the RNAzol B method.

RT-PCR. The standard reverse transcription (RT) protocol used was performed in the presence of actinomycin D. Briefly, 5 μg of total cellular RNA was denatured at 95°C for 5 min in presence of 1 μg of oligonucleotide complementary

to sequences in mt exon 3 (primer B [5′-tatggagaacgagtc agggttgtac-3′]). Annealing was then allowed to occur in the presence of KCl at 55°C for 20 min; this procedure was followed by the addition of an RT mix containing 5× RT buffer, RNAsin, 0.83 mM deoxynucleoside triphosphates (dNTPs), 0.5 μg of actinomycin D, and 200 U of Moloney murine leukemia virus reverse transcriptase and extension at 37°C for 1 h. At this point, the cDNA samples were base denatured, boiled for 3 min, and precipitated in the presence of HCl and sodium acetate.

After resuspension of the cDNAs in 50 μl of Tris-EDTA, 5 μl of cDNA was used in 100-μl PCRs containing 50 mM KCl, 10 mM Tris-HCl (pH 9.0 at 25°C), 0.1% Triton X-100, 1.5 mM $MgCl_2$, 1.25 μM each dNTP, 1 μg each of primers A and B, approximately 40,000 cpm of [γ-^{32}P]dATP-labeled primer A, and 2.5 U of *Taq* DNA polymerase (Promega). Primer A is homologous to the R region of the RSV LTR (5′-accattcaccacattggtgtgcacc-3′). PCR conditions were as follows: an initial denaturation cycle at 94°C for 5 min, followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 2 min. PCR cDNA products were digested with *Pvu*II and visualized on 10% nondenaturing polyacrylamide gels prior to autoradiography. In some instances, nonradioactive primers were used, and the gels were silver stained. In Fig. 6B, APRT-specific primers homologous to nucleotides 409 to 428 of exon 1 (primer C [5′-atggcggaatctgagttgca-3′]) and 2288 to 2297 of exon 5 (primer D [5′-agaatggtactgat cctagc-3′]) were used. In this case, the RT-PCR protocol used was exactly the same as discussed above except that 30 PCR cycles were done at 94°C for 1 min, 60°C for 1 min, and 72°C for 2 min. The resulting cDNA samples were visualized on silver-stained 10% nondenaturing polyacrylamide gels.

In vitro analysis. In vitro splicing templates were constructed as follows: E3 and the expanded exon 3, a genomic sTNI fragment beginning with a *Pvu*II site 111 nucleotides upstream of exon 3 and terminating at a *Bst*BI site 172 nucleotides downstream of exon 3 (RNA was transcribed from template DNA truncated at a *Sac*I site 35 nucleotides downstream of exon 3); E2E3 wild type and 5′ splice site mutant, a genomic fragment beginning with an *Eco*RI site 99 nucleotides upstream of exon 2 and terminating with a *Sac* site 35 nucleotides downstream of exon 3 (RNA was transcribed from template DNA truncated at a *Bam*HI site located just downstream of the insert in the vector polylinker) cloned into pGEM transcribable vectors. In vitro splicing reactions were performed as described previously (53). Assembly reactions were analyzed by using reactions supplemented with 1% polyethylene glycol; splicing reactions were supplemented with 4% polyethylene glycol. Assembled complexes were analyzed on native gels consisting of 3.5% polyacrylamide (80:1) and 1% agarose as described previously (77).

RESULTS

sTNI gene organization and experimental design of minigenes. We initially suspected that the sTNI mini-exon might require elements distal to the exon itself for maximal recognition and inclusion in spliced RNA. Our first goal, therefore, was to define a minimum cassette of sTNI information competent to direct inclusion of the mini-exon when transferred to a heterologous gene. The intron upstream of the mini-exon (exon 3) is only 166 nucleotides, and the total length of the portion of the sTNI gene encompassing exons 2 and 3 is only 203 nucleotides, within the size limit for vertebrate internal exons (53). From a mechanistic standpoint, it seemed possible that the relatively small size of intron 2 might have a functional significance in recognition of exon 3. Therefore, our first constructs included both exons 2 and 3.

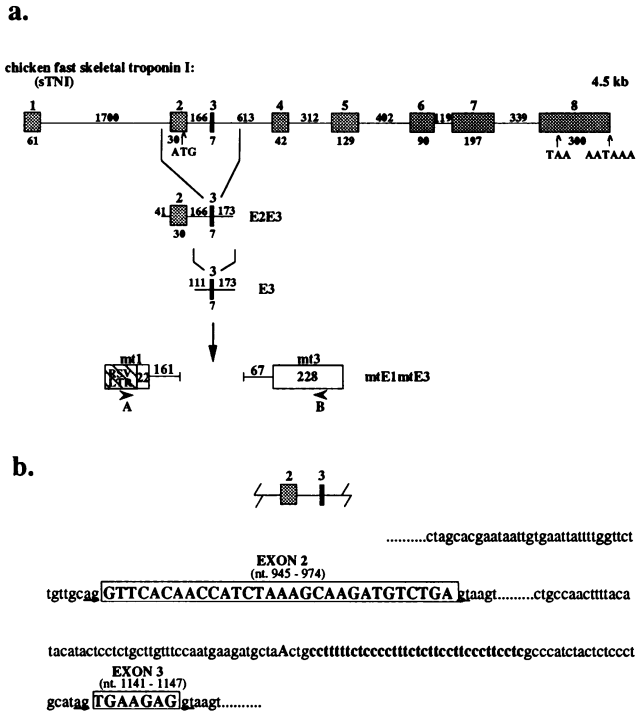

FIG. 1. Schematic diagram of the chicken sTNI gene and construction of the parental sTNI minigenes. (a) The entire eight-exon sTNI gene. The E2E3 minigene was constructed by inserting a 417-nucleotide *Nhe*I-*Bst*BI sTNI DNA fragment, which encompasses exons 2 and 3, into a modified mouse mt gene (mtE1mtE3). Prior to insertion of any sTNI sequence, the mtE1mtE3 gene was digested with *Sma*I to remove the second of three mt exons. The E3 minigene was generated by inserting a 291-nucleotide *Pvu*II-*Bst*BI sTNI DNA fragment into the mtE1mtE3 gene. In all figures, sTNI exon sequences are shown as filled boxes and mt sequences are shown as white boxes. Expression of the corresponding pre-mRNA is driven by an RSV LTR promoter element. (b) The region of the sTNI gene (nucleotides [nt.] 904 to 1320) encompassing exon 2-intron 2-exon 3. Exons are boxed, conserved splicing consensus sequences are underlined, the 34-nucleotide uninterrupted polypyrimidine tract is shown in boldface type, and the putative exon 3 branch point is indicated by an **A**.

Figure 1a depicts a schematic of the sTNI gene, including the DNA fragments used to construct the minigenes E2E3 and E3 containing sTNI exons 2 and 3 and just exon 3, respectively. The E2E3 minigene contains sTNI sequences encompassing the last half of intron 1 through the first half of intron 3. The E3 minigene contains exon 3 flanked by 111 and 173 nucleotides of introns 2 and 3, respectively. In both cases, the sTNI DNA fragments were inserted into a modified mouse mt gene in which the second of three mt exons was replaced with sTNI sequences. Figure 1b displays the primary sequence of the region of the sTNI gene that is present in the E2E3 minigene. Included are the sequences and nucleotide locations of pertinent splicing consensus signals flanking both exons 2 and 3. Especially notable is the 34-nucleotide pyrimidine tract upstream of the mini-exon.

Inclusion of the mini-exon requires upstream sTNI sequences. The splicing phenotypes of the E2E3 and E3

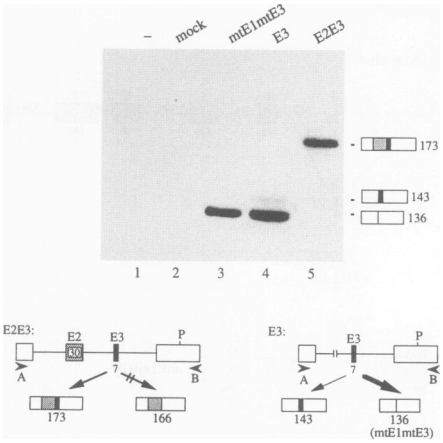

FIG. 2. Evidence that inclusion of the mini-exon requires upstream sTNI sequences. Total RNA (5 μg) isolated from 3T3 cells 48 h posttransfection with the indicated minigenes was assayed for inclusion of exon 3 by RT-PCR, using primer B for cDNA synthesis and primers A (radiolabeled) and B for PCR amplification. The resulting products were digested with *Pvu*II and visualized on a 10% nondenaturing polyacrylamide gel, which was subsequently subjected to autoradiography. Utilized minigenes and potential spliced RNA products are illustrated at the bottom; arrows indicate utilized splicing pathways. Nucleotide numbers indicate the size of the expected RT-PCR product after *Pvu*II digestion. Lanes: 1, no RNA; 2, RNA from untransfected cells; 3 to 5, RNA from transfections of the indicated minigenes.

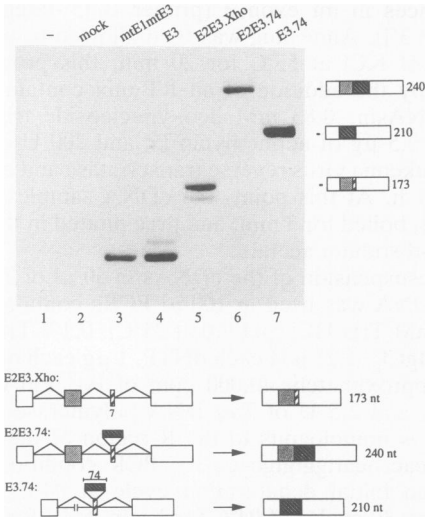

FIG. 3. Evidence that an expanded mini-exon does not require upstream sTNI sequences. Total RNA from cells transfected with the indicated minigenes was assayed for inclusion of exon 3 by RT-PCR as described for Fig. 2. Observed products are illustrated at the bottom. Lanes: 1, no RNA; 2, RNA from untransfected cells; 3 to 7, RNA from transfections of the indicated minigenes.

minigenes were determined by using an RT-PCR assay of total RNA obtained from transiently transfected NIH 3T3 cells (Fig. 2). In this assay, the final PCR product was subjected to restriction nuclease digestion before gel visualization to maximize the ability to distinguish products that differ by seven nucleotides, resulting from exclusion or inclusion of exon 3. Analysis of the splicing phenotype of the E2E3 minigene indicated that in this sequence context, both exons 2 and 3 were correctly recognized and included in the processed mRNA. Inclusion was demonstrated by the presence of a single 173-nucleotide amplified fragment containing sTNI exons 2 and 3, having a sequence consistent with correct splicing of exons 2 and 3 (Fig. 2, lane 5). No PCR product consistent with RNA resulting from skipping of exon 3 was detected. Furthermore, no PCR product RNA that contained intron sequences was observed, indicating that the introns between and flanking the sTNI exons were efficiently spliced in the transgene. Because the splicing pattern of the E2E3 minigene was identical to that observed with the wild-type gene in vivo (i.e., constitutive inclusion of the mini-exon), we believe that the E2E3 mini-gene contains all of the sTNI sequence information necessary for proper inclusion of the mini-exon in spliced RNA.

In contrast, analysis of the splicing phenotype of the E3 minigene indicated that exon 3 was not recognized in this context. PCR analysis revealed a prominent 136-nucleotide amplified fragment diagnostic of exon 3 skipping and only trace amounts of a 143-nucleotide amplified fragment diagnostic of exon 3 inclusion (Fig. 2; compare lanes 3 and 4). Absence of exon 3 in the produced RNA was confirmed by sequencing the amplified band. These data indicated that the sequence context of exon 3 in the E3 minigene is not sufficient to support efficient inclusion of the 7-nucleotide exon. No PCR product that contained intron sequences was observed, indicating that the presence of the mini-exon in the primary transcript did not cause the accumulation of a partially processed RNA species in total cell RNA.

An expanded mini-exon does not require upstream sTNI sequences. The phenotypes of the E3 and E2E3 minigenes suggested that exon 3 alone was incapable of being efficiently recognized in vivo. We suspected that the extremely small size of exon 3 contributed to the exclusion phenotype of exon 3 within the context of the E3 minigene. To test this hypothesis, a mini-exon containing an *Xho*I restriction site within exon 3 was created by mutation of nucleotides 2 to 4 of the exon (TGAAGAG to TCTCGAG). The modified exon was still included in mRNA when part of an E2E3 minigene (Fig. 3, lane 5). A small amount of skipping of the modified exon was detected by RT-PCR, suggesting that sequences within the mini-exon, although not necessary for recognition of exon 3, may facilitate maximal recognition.

The modified exon 3 was expanded at the *Xho*I site with a 63-nucleotide fragment of adenosine deaminase cDNA, increasing the size of exon 3 from 7 to 74 nucleotides. The effect of expanding the exon on splicing was examined in minigenes with or without exon 2. The expanded exon 3 was included in spliced RNA when exon 2 was present (Fig. 3, lane 6). More importantly, the expanded exon 3 was also included when exon 2 was absent (Fig. 3, lane 7). This phenotype suggests that expansion of exon 3 alleviated steric problems associated with the size of the natural exon and allowed exon 3 to be processed independently of upstream sTNI sequences. It also indicated that the splice sites bordering the mini-exon could be spliced to the splice sites of the transgene mt exons when the sTNI sites were placed further apart. Therefore, the failure of the natural mini-exon to be included in the E3 product RNA is probably not due to an incompatibility of splice sites between the sTNI and mt exons but rather is caused by some other enabling effect of the upstream sTNI exon 2.

Mutation of the upstream sTNI exon 2 alters recognition of the mini-exon. To determine whether a specific sTNI exon 2 sequence was essential for efficient recognition of exon 3,

sTNI exon 2 was inactivated in the context of the E2E3 minigene. If exon 2 represents an essential *cis*-acting sequence for recognition of exon 3, inactivation of the exon 2 5′ splice site should affect recognition and subsequent inclusion of exon 3. To test this prediction, the 5′ splice site of exon 2 was altered from GA:**GT**AAGT to GA:**CG**AAGT in the E2E3 minigene. In vertebrates, mutation of internal 5′ splice sites results in either activation of a nearby cryptic 5′ splice site or skipping of the mutated exon. Therefore, were the mini-exon a normal exon, the predicted product RNA would be expected to contain exon 3 but not exon 2. Given the previously observed dependence of exon 3 on exon 2, however, a likely phenotype was skipping of both exons 2 and 3.

Upon transfection of the mutant minigene, however, a novel phenotype was observed: exon 3 was neither properly included nor skipped. Instead, a spliced RNA was observed from which only introns 1 and 3 had been removed but intron 2 was retained. The identity of the RNA product was confirmed by sequencing the PCR product band. There was no indication of a spliced species resulting from skipping of exon 2, exon 3, or both exons 2 and 3 (Fig. 4a, lanes 4 and 5). Instead, inactivation of the exon 2 5′ splice site effectively inhibited recognition of the 3′ splice site preceding exon 3, resulting in a product that retained intron 2. This product was observed only in the mutant construct; it was never observed in wild-type constructs, indicating that this intron is removed with normal efficiency when its splice sites are wild type.

The presence of a product RNA containing an included intron is strikingly unusual among the vertebrate 5′ splice site mutants that have been analyzed to date (2, 8, 9, 14, 24, 31–34, 38, 42, 46, 56, 57, 60, 62, 66–69). The observation of skipping of exons bearing 5′ splice site mutations was one of the cornerstones of our previous suggestion that exons are the units of vertebrate splice site recognition (53, 60). Such skipping has been reported even in cases in which the affected intron is quite short (14), so the inclusion of the intron upstream of the mini-exon is presumably not the result of the small size of the intron but instead may reflect the mechanism of mini-exon recognition.

Although exon 3 required exon 2 for recognition, the converse was not observed. To test the requirement for exon 3 inclusion, we inactivated exon 3 in the E2E3 minigene by mutation of the exon 3 5′ splice site from **AG**:GTAAGT to **GC**:GTAAGT. In this case, the predominant splicing pattern observed was inclusion of exon 2 and exclusion of exon 3 (Fig. 4b; compare lanes 3 and 4). Hence, the dependence of one exon on the other is not bidirectional; only exon 3 is dependent on an adjacent sTNI exon.

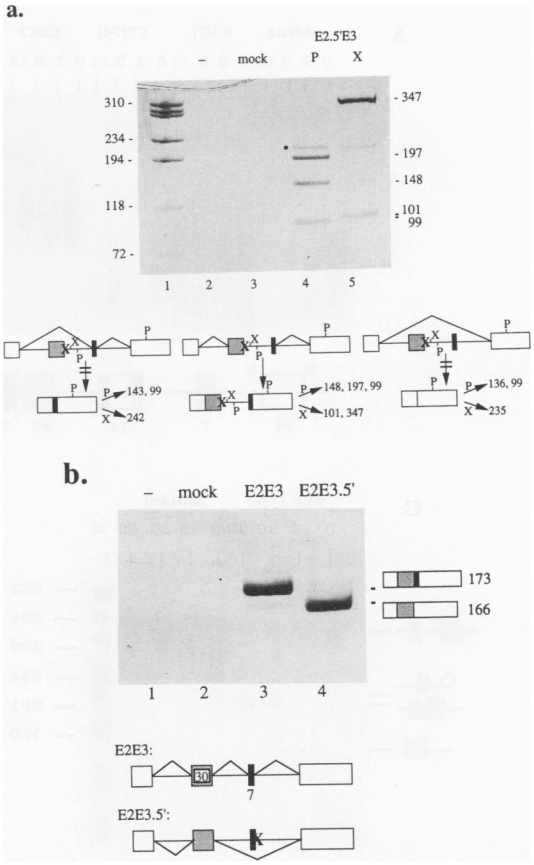

FIG. 4. Evidence that mutation of the upstream sTNI exon 2 alters recognition of the mini-exon. (a) Mutation of exon 2. Total RNA from cells transfected with the indicated minigenes was assayed for inclusion of exon 3 by RT-PCR as described for Fig. 2 except that the products were visualized by silver staining. Prior to loading of the gel, the amplified product was digested with either *Pvu*II or *Xho*I. Utilized minigenes and potential amplification and digestion products are illustrated at the bottom. Lanes: 1, ϕX174 DNA digested with *Hae*III; 2, no RNA; 3, RNA from untransfected cells; 4 and 5, RNA from transfections of the indicated minigene digested with the indicated enzyme. The minor band denoted by a dot in lanes 4 and 5 is of unknown origin. (b) Mutation of exon 3. Total RNA from cells transfected with the indicated minigenes was assayed for inclusion of exon 3 by RT-PCR as described for Fig. 2. Observed products are illustrated at the bottom. Lanes: 1, no RNA; 2, RNA from untransfected cells; 3 and 4, RNA from transfections of the indicated minigenes.

In vitro assembly of the mini-exon requires the upstream exon. The dependence of recognition of the mini-exon upon the presence of the upstream exon was also observed in in vitro assembly and splicing assays. The first set of experiments examined the ability of precursor RNAs containing a single exon to direct assembly of spliceosomal complexes. A precursor containing the mini-exon did not direct the assembly of noticeable amounts of complex compared with the assembly of a standard adenovirus single-exon precursor RNA when reactions were analyzed on native gels (Fig. 5A). A precursor containing the expanded mini-exon did direct the assembly of detectable complex, although not to the levels of the adenovirus transcript. A precursor containing the upstream exon 2 also directed low levels of complex formation. Thus, both exon 2 and an expanded exon 3 were recognized better in in vitro splicing extracts than was the natural mini-exon.

We also examined a precursor RNA containing both exons 2 and 3. This precursor directed high levels of complex formation. At early times of assembly, a single complex with mobility less than that of the reference adenovirus complex was formed (Fig. 5A). Thus, a precursor containing both exons 2 and 3 assembled considerably better than a precursor containing either exon alone. The requirement for more than one exon for maximal initial early assembly is not typical behavior for vertebrate precursor RNAs in which a single internal exon normally directs identical levels of assembly as a two-exon precursor RNA (53, 77).

Figures 5B and C analyze the assembly and splicing activity of two-exon sTNI pre-mRNAs in which the 5′ splice site terminating the first exon was wild type or mutant. The

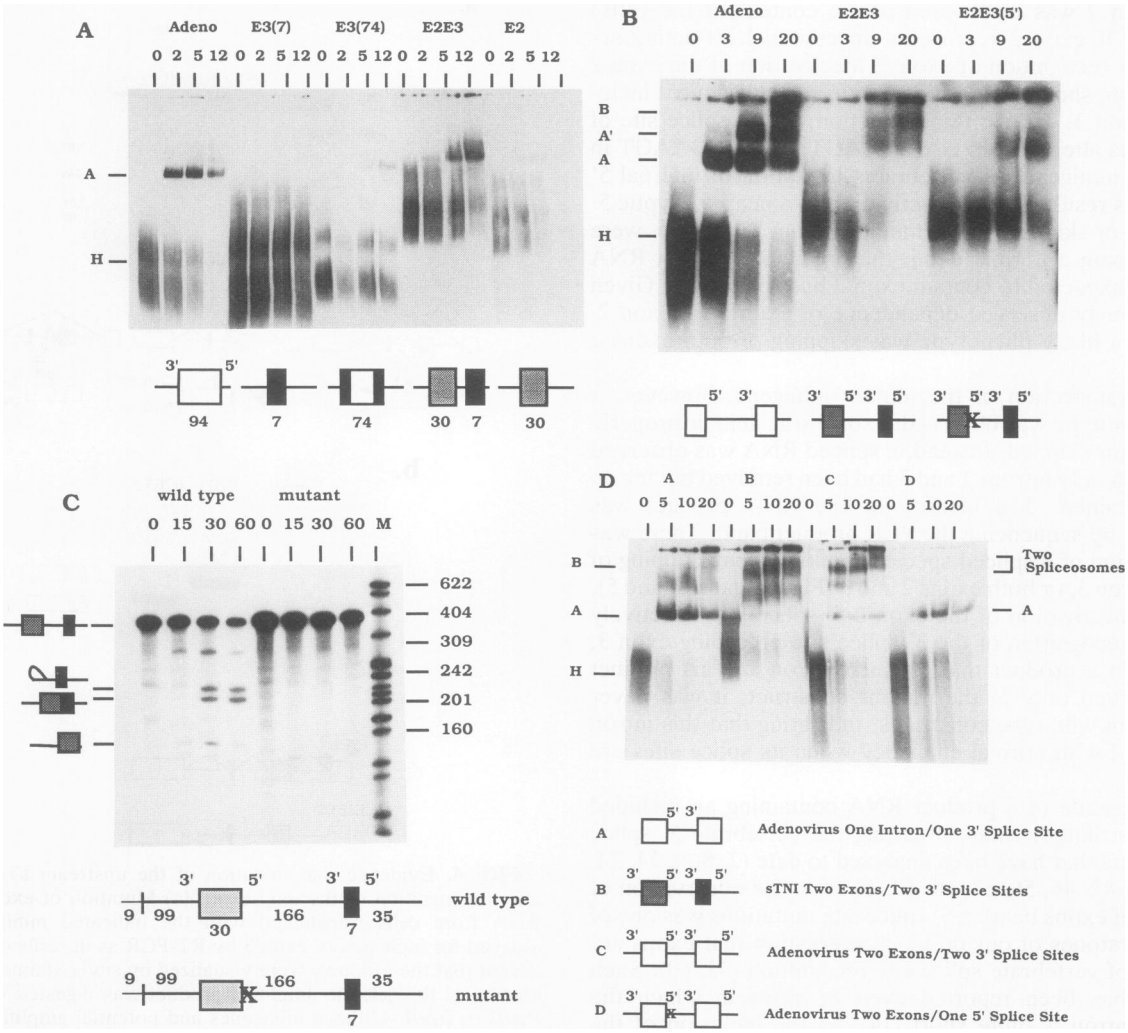

FIG. 5. Evidence that in vitro recognition of the mini-exon also requires the upstream exon. In vitro splicing reactions of the diagrammed precursor RNAs were examined for complex formation on nondenaturing gels (A, B, and D) or RNA product formation (C) on denaturing gels at the indicated times. The adenovirus reference precursor RNAs have been described previously (53). The adenovirus exons in all precursors are identical. (A) Assembly of pre-mRNAs consisting of a single exon and flanking intron sequences. Adeno, an adenovirus internal exon; E3(7), the natural sTNI exon 3; E3(74), an expanded exon 3 with a total length of 74 nucleotides; E2E3, the natural sTNI exons 2 and 3 with flanking introns; E2, the natural sTNI exon 2. (B) Assembly of sTNI pre-mRNAs consisting of two internal exons. Adeno, a single intron adenovirus precursor RNA with one 5′ and one 3′ splice site; E2E3, the natural sTNI exons 2 and 3 with two 5′ and two 3′ splice sites; E2E3(5′), a mutant sTNI exon 2 with an altered 5′ splice site and the natural exon 3 with one 5′ splice site and two 3′ splice sites. (C) RNA products from the reactions depicted in panel B, analyzed on a denaturing acrylamide gel. Splicing product and intermediate RNAs are indicated. (D) Assembly of precursor RNAs with two complete internal exons. Precursor RNAs A to D are diagrammed below the gel. All exons represented by white boxes are from adenovirus exon 2; B is the natural sTNI exons 2 and 3. Precursor RNA A can direct the formation of only one spliceosome; precursor C can direct the assembly of two spliceosomes; precursor D can direct the assembly of only a single spliceosome.

wild-type sTNI construct directed the assembly of at least two complexes, both of which had gel migration properties noticeably different from that of a standard two-exon adenovirus construct. This difference could be due to inherent differences in the factors associating with the two precursor RNAs or to the presence of two 3′ splice sites in the sTNI constructs and only one in the adenovirus constructs (see below). The important point from Fig. 5B is that mutation of the 5′ splice site of exon 2 resulted in observation of only one complex of migration identical to that of the first complex seen with the wild-type sTNI RNA. Notice that this single complex is noticeably different in gel migration from the complex (complex A) formed by precursor RNAs containing a single adenovirus exon or those formed by precursors containing only sTNI exon 2 (Fig. 5A) or an expanded sTNI exon 3 (Fig. 5A). This difference suggests that the two-exon sTNI complex represents an unusual assembly requiring sequence elements from both exons 2 and 3.

The pattern of assembly in Fig. 5B reflects the assembly of an active complex. Analysis of the RNA from such an assembly reaction indicates that the two-exon sTNI pre-mRNA was active for splicing (Fig. 5C). Thus, not only was the sTNI mini-exon recognized for assembly, it was also recognized sufficiently well to permit splicing.

The complexes formed with the two-exon sTNI precursor RNAs differed in gel mobility from those formed with

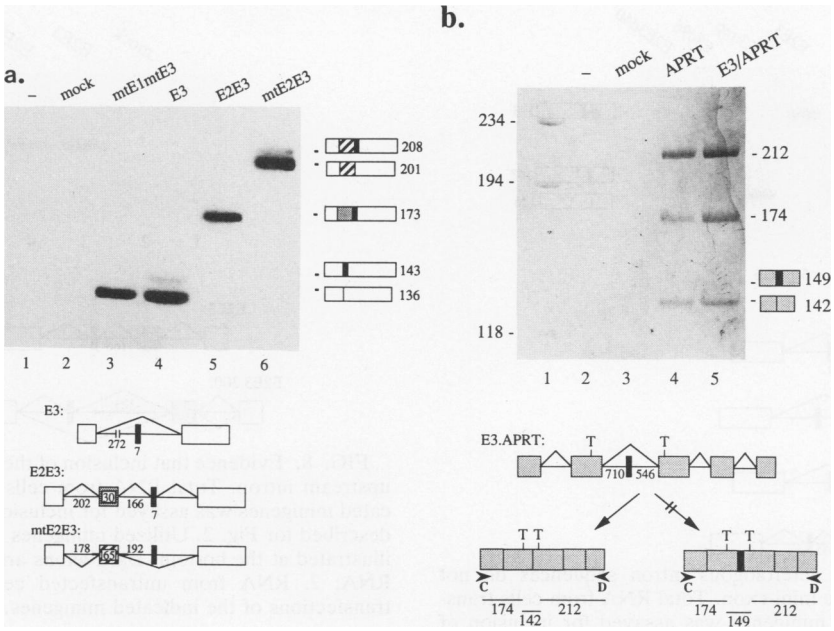

FIG. 6. Evidence that heterologous exons do not facilitate recognition of the mini-exon. (a) mt exon 2. Total RNA from cells transfected with the indicated minigenes was assayed for inclusion of exon 3 by RT-PCR as described for Fig. 2. Utilized minigenes and potential products are illustrated at the bottom. Black exon, sTNI mini-exon 3; stippled exon, sTNI exon 2; striped exon, mt exon 2. Lanes: 1, no RNA; 2, RNA from untransfected cells; 3 to 6, RNA from transfections of the indicated minigenes. (b) APRT exon 2. The E3.APRT minigene containing the sTNI mini-exon 3 inserted into intron 2 of the hamster APRT gene was cotransfected with pSV2neo into an APRT$^-$ cell line. Total RNA isolated from drug-resistant colonies was assayed for inclusion of exon 3 by RT-PCR as described for Fig. 2 (modified to a PCR profile of 94°C for 1 min, 60°C for 1 min, and 72°C for 2 min), using oligonucleotides complementary to exons 1 and 5 of the APRT gene. The resulting products were visualized by silver staining. Prior to loading of the gel, the amplified product was digested with *Taq*I. The utilized minigene and potential amplification and digestion products are illustrated at the bottom. Lanes: 1, ϕX174 DNA digested with *Hae*III; 2, no RNA; 3, RNA from untransfected APRT$^-$ cells; 4, endogenous APRT RNA from an APRT$^+$ cell line; 5, RNA from the APRT$^-$ cell line containing the E3.APRT minigene.

standard adenovirus precursor RNAs. Figure 5D compares the assembly gel profiles of wild-type and mutant two-exon precursor RNAs containing adenovirus or sTNI sequences. The adenovirus precursor RNA consisting of two internal exons rapidly assembled into a large, slowly migrating complex of mobility consistent with the presence of two spliceosomes on the pre-mRNA. With time, larger, more slowly migrating complexes appeared, consistent with appearance of the active spliceosome and reaction products. The migration of these complexes was noticeably different from that of similarly sized and constructed sTNI precursor RNAs. The sTNI two-exon pre-mRNA directed the assembly of a set of complexes intermediate in mobility between the single and double spliceosome complexes formed on the reference adenovirus pre-mRNAs. Thus, this RNA behaved like neither a single nor a double exon but more like something in between. This behavior is consistent with a two-step mode of recognition for sTNI exons 2 and 3 in which first the two-exon unit is recognized as a single exon and then the intervening intron is recognized.

Heterologous exons do not facilitate recognition of the mini-exon. To determine whether the requirement for an upstream sequence element was specific for an sTNI exon, we substituted two heterologous exons for the authentic sTNI exon 2. The first construction (mtE2E3) utilized the mouse mt exon 2 (65 nucleotides) and 43 and 76 nucleotides of the flanking upstream and downstream introns, respectively. As shown in Fig. 6a, the presence of an upstream mt exon did not facilitate inclusion of exon 3, in contrast to the complete inclusion pattern observed when the natural sTNI exon 2 was present (Fig. 6a; compare lanes 5 and 6). This result suggested that simply providing an internal exon of comparable size in a position equivalent to that of the natural gene was not adequate to restore efficient inclusion of exon 3. Further support for this observation comes from a second exon swap experiment in which the isolated sTNI exon 3, flanked by 111 and 173 nucleotides of introns 2 and 3, respectively, was inserted into the second intron of the human APRT gene (the E3.APRT minigene). In this sequence context, APRT exon 2 (111 nucleotides) serves as the upstream exon for the 7-nucleotide exon 3. Analysis of the splicing pattern of E3.APRT revealed that exon 3 was completely skipped in this context, indicating that APRT exon 2 also failed to rescue inclusion of exon 3 (Fig. 6b, lane 5). Inclusion of exon 3 was restored when both exons 2 and 3 were inserted into the APRT transgene (data not shown); therefore, the exclusion seen in transgenes containing only exon 3 was the result of the absence of exon 2.

Heterologous intron sequences do not facilitate recognition of the mini-exon. Collectively, the data presented thus far strongly suggest that efficient inclusion of exon 3 in spliced product RNA requires the presence of the normal upstream exon 2. To determine the role, if any, that intron length played in the mechanism of inclusion of exon 3, we constructed appropriate intronic replacement mutants. The difference in length between the E2E3 and E3 minigenes is 126 nucleotides. Accordingly, we inserted additional vertebrate intron sequence in the E3 minigene upstream of exon 3. Analysis of the splicing phenotype of the mutant E3.rep revealed that addition of heterologous intron sequence did

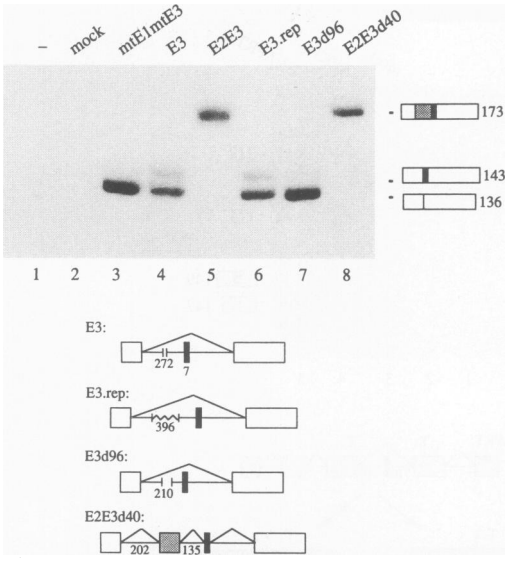

FIG. 7. Evidence that heterologous intron sequences do not facilitate recognition of the mini-exon. Total RNA from cells transfected with the indicated minigenes was assayed for inclusion of exon 3 by RT-PCR as described for Fig. 2. Utilized minigenes and observed products are illustrated at the bottom. Dark exons are from sTNI. Lanes: 1, no RNA; 2, RNA from untransfected cells; 3 to 8, RNA from transfections of the indicated minigenes.

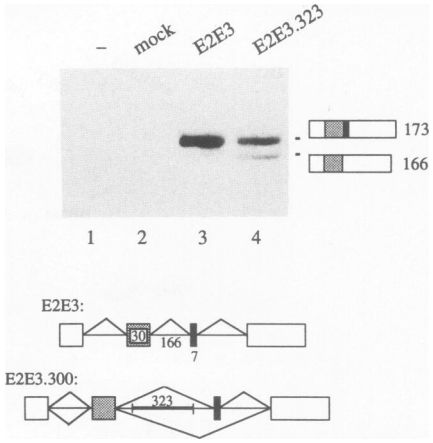

FIG. 8. Evidence that inclusion of the mini-exon requires a small upstream intron. Total RNA from cells transfected with the indicated minigenes was assayed for inclusion of exon 3 by RT-PCR as described for Fig. 2. Utilized minigenes and observed products are illustrated at the bottom. Dark exons are from sTNI. Lanes: 1, no RNA; 2, RNA from untransfected cells; 3 and 4, RNA from transfections of the indicated minigenes.

not improve the level of exon 3 inclusion compared with that observed with the E3 minigene (Fig. 7; compare lanes 4 and 6). This observation further supported a role for a specific sTNI upstream sequence element in recognition of exon 3. The various tested E3 minigenes all demonstrated minor RNAs resulting from inclusion of exon 3. Further truncation of the sTNI intron 2 preceding exon 3 from 111 to 70 nucleotides (mutant E3d96) abolished these trace levels of inclusion of exon 3 (Fig. 7; compare lanes 4 and 7). Interestingly, we were able to rescue inclusion of exon 3, despite the deletion of intron 2 sequences, by adding back exon 2 and its bordering splice sites (mutant E2E3d40; Fig. 7, lane 8). Taken together, these data indicate that the majority of sTNI intron 2 is dispensable for exon 3 recognition, since E2E3d40 is missing the middle one-third of intron 2.

Maximal inclusion of the mini-exon requires a small upstream intron. In the natural sTNI gene, intron 2 is only 166 nucleotides in length. This fact, in conjunction with the observed requirement for the natural upstream exon, suggested that exon 3 may be recognized initially as an exon-intron-exon unit. In terms of size, the length of sTNI exon 2-intron 2-exon 3 (203 nucleotides) is within the size limits of natural internal vertebrate exons (30, 53). This hypothesis predicts that expansion of the natural sTNI intron 2 should depress recognition of exon 3. To test this prediction, intron 2 was expanded with 323 nucleotides of vertebrate intron sequences within the context of the E2E3 minigene. Intronic expansion reduced, but did not eliminate, inclusion of exon 3 (Fig. 8; compare lanes 3 and 4). This observation suggests that maximal inclusion of exon 3 requires both exon 2 and an adjacent small intron. The relationship between exon 2 and intron 2 is somewhat unclear. When exon 2 was replaced with a constitutive exon, exon 3 was skipped with intron lengths ranging from 192 to 710 nucleotides. Furthermore, the shorter the intervening sequence in the intron replacement series, the greater the skipping. In the presence of exon 2, exon 3 was included when the intron was 135 or 166 nucleotides but began to be skipped when it was enlarged to 489 nucleotides. Therefore, when exon 2 is present, shorter introns seem preferable for inclusion; when exon 2 is absent, longer introns seem more beneficial.

DISCUSSION

Vertebrate internal exons are normally short, some extremely so. A number of mini-exons whose length is less than 12 nucleotides have been identified (1, 3, 4, 10, 15, 41, 43, 45, 49, 54, 55, 61). Such short exons seem problematic for splicing because it is difficult to envision simultaneous recognition of splice sites separated by only a few nucleotides. In this study, we have investigated the mechanism of in vivo recognition of a vertebrate mini-exon. The mini-exon under consideration was the 7-nucleotide exon 3 of the chicken sTNI gene. Despite its extremely small size, this mini-exon is constitutively included in sTNI mRNA in vivo (45). When transferred to a heterologous gene, however, exon 3 was included in the final mRNA only when preceded by the natural upstream exon and intron. The upstream exon was required only when exon 3 was its natural small size; expansion of exon 3 to 74 bases permitted it to be included in spliced RNA in the absence of exon 2. The ability of the expanded exon to be included in product RNA suggests that the splice sites flanking the mini-exon are capable of being spliced to neighboring exons when they are further apart. Thus, it seems likely that it is the small size of exon 3, rather than some fundamental incompatibility of splice sites between the mini-exon and other vertebrate constitutive exons, that prevents its recognition in a transgene.

Mutation of the 5′ splice site of exon 2 depressed inclusion of exon 3. Instead, an unusual vertebrate phenotype resulted in which all introns except the mutated one were removed. Normally, mutation of the 5′ splice site of an internal vertebrate exon results in either exon skipping or the recognition of a cryptic 5′ splice site, not intron inclusion. The exon skipping phenotype has been observed even when the affected intron is small (14). Therefore, intron inclusion in the sTNI mutant likely arises from unique requirements for

the recognition of this mini-exon. We suggest that the troponin I mini-exon is initially recognized as an exon-intron-exon unit, followed by subsequent recognition of the small intervening sequence between them. From this perspective, mutation of the 5′ splice site of this intron permits initial recognition of the exon-intron-exon unit but prohibits later intron recognition, resulting in a product RNA with an included intron.

Vertebrate internal exons have a natural maximum (30); few exons exceed 300 nucleotides, and artificial expansion of exons inactivates their recognition in vitro (53) and in vivo (59a). The intron separating the 30-nucleotide exon 2 and the 7-nucleotide exon 3 is only 166 nucleotides, making a potential exon 2-intron 2-exon 3 unit of only 203 nucleotides, well within the range of vertebrate internal exons. The 166-nucleotide intron 2 has very strong splicing signals compared with the average intron. Most striking is the presence of an uninterrupted stretch of 34 pyrimidines adjacent to the putative branch point of exon 3. If recognition of this region of the sTNI pre-mRNA actually occurs by a two-step mechanism, strong signals within the contained intron may be required for its recognition during the second step. Interestingly, these same strong signals are not sufficient to permit inclusion of exon 3 when it is separated from its natural upstream exon, suggesting that recognition of the mini-exon requires additional elements, presumably because of its extremely small size.

Although recognition of the mini-exon required the presence of a nearby upstream internal exon, other constitutive exons could not replace the natural upstream sTNI exon. Visual inspection of the relative strengths of the utilized heterologous upstream exons revealed that both possess longer, contiguous pyrimidine tracts located immediately adjacent to the 3′ splice site AG, in comparison with the natural sTNI exon 2 in which the polypyrimidine tract is shorter and punctuated with purines. This difference suggests that efficient recognition of the mini-exon requires the nearby upstream exon to be relatively weak. From a mechanistic viewpoint, this observation suggests that a weak upstream exon is necessary for initial recognition of an exon 2-intron 2-exon 3 unit. The presence of a strong upstream exon would likely promote skipping of a downstream mini-exon; only with a relatively weak upstream exon could the splicing machinery be capable of recognizing the larger exon-intron-exon unit instead of just the upstream exon. However, we do not yet know what precisely about this exon is responsible for suppression of its own 5′ splice site such as to lead to recognition of exons 2 and 3 as a unit. How does it know to skip over the strong 5′ splice site bordering exon 2 and reach further downstream to the 5′ splice site bordering exon 3 during initial recognition? Presently, we are subjecting this exon to further mutagenesis to determine whether nonconserved sequences within exon 2 are responsible for suppression of its 5′ splice site.

The sTNI exon 3 is just one example of a naturally occurring vertebrate mini-exon (Table 1). The limited number of known mini-exons can be grouped into two categories: (i) mini-exons with a short adjoining intervening sequence (i.e., cardiac troponin T exon 3, and fast skeletal troponin T exon 7) and (ii) mini-exons with long adjacent introns (i.e., neural cell adhesion molecule exon 12d, cardiac troponin T exon 17, renin exon 6, pro-alpha 2 (I) collagen exon 2, and AMP deaminase exon 2). Obviously, mechanisms other than the one proposed here for sTNI must operate for mini-exons bordered by large introns. It seems likely that mini-exons lying within large introns also require an enabling element to ensure efficient inclusion, but that in these cases the element will be located within a flanking intron.

TABLE 1. Vertebrate mini-exons

Gene	Exon no.	Exon length (nucleotides)	Short adjoining intron
Chicken/murine NCAM[a]	12d	3	−
Chicken cardiac troponin T	17	6	−
Chicken/quail sTNI	3	7	+
Human renin	6	9	−
Chicken cardiac troponin T	3	11	+
Chicken pro-alpha 2 (I) collagen	2	11	−
Chicken/rat fast skeletal troponin T	7	12	+
Human/rat AMP deaminase	2	12	−

[a] NCAM, neural cell adhesion molecule.

Whereas only a handful of mini-exons with lengths of 12 nucleotides or less have been identified, there are a significant number of identified exons within the 20- to 30-nucleotides size range (30). These exons are probably also somewhat stressed for recognition in that constitutive exons shortened below 51 nucleotides are skipped unless the strength of their splicing signals is increased (17). Computer correlation of exon length and splice site strength indicates that as exon length deviates from the average value of 130 to 150 nucleotides, splice site strength increases (12). Therefore, short exons seem to require stronger than normal splice sites to be recognized.

Finally, some differentially utilized exons are small. At least one of these has been shown to become constitutively recognized when lengthened (6). Others seem to have special elements that are required for recognition but do not participate in differential recognition (14a). In general, differential exons frequently have weaker splicing signals than do constitutive exons. Our results would suggest that one way to weaken an exon and make it dependent upon special sequences and their cognate factors is to shorten the exon. As a general model for the recognition of small vertebrate exons, we propose the following: (i) exons in the 20- to 30-nucleotide size range require strong splicing signals for inclusion; (ii) exons in the 20- to 30-nucleotide size range with suboptimal splicing signals require an additional enabling element for inclusion; and (iii) mini-exons of 12 nucleotides or less require both strong splicing signals as well as an enabling element for inclusion.

REFERENCES

1. **Aho, S., V. Tate, and H. Boedtker.** 1984. Location of the 11 bp exon in the chicken pro alpha 2(1) collagen gene. Nucleic Acids Res. **12**:6117–6125.
2. **Bakhshi, A., P. Guglielmi, U. Siebenlist, J. V. Ravetch, J. P. Jensen, and S. J. Korsmeyer.** 1986. A DNA insertion/deletion necessitates an aberrant RNA splice accounting for a mu heavy chain disease protein. Proc. Natl. Acad. Sci. USA **83**:2689–2693.
3. **Baldwin, A. S., E. L. Kittler, and C. P. Emerson.** 1985. Structure, evolution, and regulation of a fast skeletal muscle troponin I gene. Proc. Natl. Acad. Sci. USA **82**:8080–8084.
4. **Barthels, D., M.-J. Santoni, W. Wille, C. Ruppert, J.-C. Chaix, M.-R. Hirsch, J.-C. Fontecilla-Camps, and C. Goridis.** 1987. Isolation and nucleotide sequence of mouse N-CAM cDNA that codes for a Mr 79,000 polypeptide without a membrane-spanning region. EMBO J. **6**:907–914.
5. **Bhat, B. M., H. A. Brady, M. H. Pursley, and W. S. M. Wold.**

1986. Deletion mutants that alter differential RNA processing in the E3 complex transcription unit of adenovirus. J. Mol. Biol. **190:**543–557.

6. **Black, D. L.** 1991. Does steric interference between splice sites block the splicing of a short c-src neuron-specific exon in non-neuronal cells? Genes Dev. **5:**389–402.
7. **Black, D. L.** 1992. Activation of c-Src neuron specific splicing by an unusual RNA element in vivo and in vitro. Cell **69:**795–807.
8. **Bonadio, J., F. Rameriz, and M. Barr.** 1990. An intron mutation in the human a1(I)collagen gene alters the efficiency of pre-mRNA splicing and is associated with osteogenesis imperfecta type II. J. Biol. Chem. **265:**2262–2268.
9. **Brandt, C. R., S. L. Morrison, B. K. Birshtein, and C. Milcarek.** 1984. Loss of a consensus splice signal in a mutant immunoglobulin gene eliminates the CH_1 domain exon from the mRNA. Mol. Cell. Biol. **4:**1270–1277.
10. **Breitbart, R. E., and B. Nadal-Ginard.** 1986. Complete nucleotide sequence of the fast skeletal troponin T gene. J. Mol. Biol. **188:**313–324.
11. **Breitbart, R. E., and B. Nadal-Ginard.** 1987. Developmentally induced, muscle-specific *trans* factors control the differential splicing of alternative and constitutive troponin T exons. Cell **49:**793–803.
12. **Brunak, S., J. Engelbrecht, and S. Knudsen.** 1991. Prediction of human mRNA donor and acceptor sites from the DNA sequence. J. Mol. Biol. **220:**49–65.
13. **Casanova, J.-L., C. Pannetier, C. Jaulin, and P. Kourilsky.** 1990. Optimal conditions for directly sequencing double-stranded PCR products with Sequenase. Nucleic Acids Res. **18:**4028.
14. **Chen, E. Y., Y. C. Liao, D. H. Smith, H. A. Barrera-Saldana, R. E. Gelinas, and P. H. Seeburg.** 1989. The human growth hormone locus: nucleotide sequence, biology, and evolution. Genomics **4:**479–497.

14a. **Cooper, T.** Personal communication.

15. **Cooper, T. A., and C. P. Ordahl.** 1985. A single cardiac troponin T gene generates embryonic and adult isoforms via developmentally regulated alternative splicing. J. Biol. Chem. **260:**11140–11148.
16. **Cooper, T. A., and C. P. Ordahl.** 1989. Nucleotide substitution within the cardiac troponin T alternative exon disrupt pre-mRNA alternative splicing. Nucleic Acids Res. **17:**7905–7921.
17. **Dominski, Z., and R. Kole.** 1991. Selection of splice sites in pre-mRNA with short internal exons. Mol. Cell. Biol. **11:**6075–6083.
18. **Dominski, Z., and R. Kole.** 1992. Cooperation of pre-mRNA sequence elements in splice site selection. Mol. Cell. Biol. **12:**2108–2114.
19. **Freyer, G. A., J. P. O'Brien, and J. Hurwitz.** 1989. Alterations in the polyY sequence affect the in vitro splicing reactions catalyzed by HeLa cell-free preparations. J. Biol. Chem. **264:**14631–14637.
20. **Fu, X.-D., R. A. Katz, A. M. Skalka, and T. Maniatis.** 1991. The role of branchpoint and 3′-exon sequences in the control of balanced splicing of avian retrovirus RNA. Genes Dev. **5:**211–220.
21. **Fu, X.-Y., H. Ge, and J. Manley.** 1988. The role of polypyrimidine stretch at the SV40 early pre-mRNA 3′ splice site in alternative splicing. EMBO J. **7:**809–817.
22. **Furdon, P. J., and R. Kole.** 1986. Inhibition of splicing but not cleavage at the 5′ splice site by truncating human beta-globin pre-mRNA. Proc. Natl. Acad. Sci. USA **83:**927–931.
23. **Furdon, P. J., and R. Kole.** 1988. The length of the downstream exon and the substitution of specific sequences affect pre-mRNA splicing in vitro. Mol. Cell. Biol. **8:**860–866.
24. **Gibbs, R. A., P. Nguyen, A. Edwards, A. B. Civitello, and C. T. Caskey.** 1990. Multiplex DNA deletion detection and exon sequencing of the hypoxanthine phosphoribosyltransferase gene in Lesch-Nyhan families. Genomics **7:**235–244.
25. **Grabowski, P. J., F. H. Nasim, H.-C. Kuo, and R. Burch.** 1991. Combinatorial splicing of exon pairs by two-site binding of U1 small nuclear ribonucleoprotein particle. Mol. Cell. Biol. **11:**5919–5928.
26. **Graham, G. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. DNA **52:**456–467.
27. **Green, M. R.** 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. **7:**559–99.
28. **Guo, W., G. J. Mulligan, S. Wormsley, and D. M. Helfman.** 1991. Alternative splicing of beta-tropomyosin pre-mRNA: cis-acting elements and cellular factors that block the use of a skeletal muscle exon in nonmuscle cells. Genes Dev. **5:**2096–2107.
29. **Hampson, R. K., L. La Follette, and F. M. Rottman.** 1989. Alternative processing of bovine growth hormone mRNA is influenced by downstream exon sequences. Mol. Cell. Biol. **9:**1604–1610.
30. **Hawkins, J. D.** 1988. A survey on intron and exon lengths. Nucleic Acids Res. **16:**9893–9908.
31. **Hayashi, S.-I., T. Kunisada, M. Ogawa, Y. Nishikawa, and S.-I. Nishikawa.** 1991. Exon skipping by mutation of an authentic splice site of c-kit gene in W/W mouse. Nucleic Acids Res. **19:**1267–1271.
32. **Hidada, Y., T. D. Palella, T. E. O'Toole, S. A. Tarle, and W. N. Kelly.** 1987. Human adenosine phosphorylribosyl transferase—identification of allelic mutations at the nucleotide level as a cause of a complete deficiency of the enzyme. J. Clin. Invest. **80:**1409–1415.
33. **Hodges, P. E., and L. E. Rosenberg.** 1989. The spf^{ash} mouse: a missense mutation in the ornithine transcarbamylase gene also causes aberrant mRNA splicing. Proc. Natl. Acad. Sci. USA **86:**4142–4146.
34. **Kuhne, T., B. Wieringa, J. Reiser, and C. Weissmann.** 1983. Evidence against a scanning model for RNA splicing. EMBO J. **2:**727–733.
35. **Kuo, H.-C., F. H. Nasim, and P. J. Grabowski.** 1991. Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. Science **251:**1045–1050.
36. **Libri, D., M. Goux-Pelletan, E. Brody, and M. Y. Fiszman.** 1990. Exon as well as intron sequences are *cis* elements for the mutually exclusive alternative splicing of the beta-tropomyosin gene. Mol. Cell. Biol. **10:**5036–5046.
37. **Ligtenberg, M. J. L., M. C. Gennissen, H. L. Vos, and J. Hilkens.** 1990. A single nucleotide polymorphism in an exon dictates allele dependent differential splicing of episialin mRNA. Nucleic Acids Res. **19:**297–301.
38. **Marvitt, J., A. G. DiLella, K. Brayton, F. D. Ledley, K. J. H. Robson, and S. L. C. Woo.** 1987. GT to AT transition at a splice donor site causes skipping of the preceding exon in phenylketonuria. Nucleic Acids Res. **15:**5613–5628.
39. **Mardon, H. J., G. Sebastio, and F. E. Baralle.** 1987. A role for exon sequences in alternative splicing of the human fibronectin gene. Nucleic Acids Res. **15:**7725–7733.
40. **Matsuo, M., T. Masumura, H. Nishio, T. Nakajima, Y. Kitoh, T. Takumi, J. Koga, and H. Nakamura.** 1991. Exon skipping during splicing of dystrophin mRNA precursor due to an intraexon deletion in the dystrophin gene of Duchenne muscular dystrophy Kobe. J. Clin. Invest. **87:**2172–2131.
41. **Mineo, I., P. R. H. Clarke, R. L. Sabina, and E. W. Holmes.** 1990. A novel pathway for alternative splicing: identification of an RNA intermediate that generates an alternative 5′ splice donor site not present in the primary transcript of *AMPD1*. Mol. Cell. Biol. **10:**5271–5278.
42. **Mitchell, P. A., G. Urlaub, and L. Chasin.** 1986. Spontaneous splicing mutations at the dihydrofolate reductase locus in Chinese hamster ovary cells. Mol. Cell. Biol. **6:**1926–1935.
43. **Miyazaki, H., A. Fukamizu, S. Hirose, T. Hayashi, H. Hori, H. Ohkubo, S. Nakanishi, and K. Murakami.** 1984. Structure of the human renin gene. Proc. Natl. Acad. Sci. USA **81:**5999–6003.
44. **Mullen, M. P., C. W. J. Smith, J. G. Patton, and B. Nadal-Ginard.** 1991. Alpha-tropomyosin mutually exclusive exon selection: competition between branchpoint/polypyrimidine tracts determines default exon choice. Genes Dev. **5:**642–655.
45. **Nikovits, W., Jr., G. Kuncio, and C. P. Ordahl.** 1986. The chicken fast skeletal troponin I gene: exon organization and sequence. Nucleic Acids Res. **14:**3377–3390.
46. **Ohno, K., and K. Suzuki.** 1988. Multiple abnormal beta-hex-

osaminidase alpha chain mRNAs in a compound heterozygous Ashkeazi Jewish patient with Tay-Sachs disease. J. Biol. Chem. **263**:18563–18567.

47. **Parker, B. A., and G. R. Stark.** 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. J. Virol. **31**:360–369.
48. **Patterson, B., and C. Guthrie.** 1991. A U-rich tract enhances usage of an alternative 3′ splice site in yeast. Cell **64**:181–187.
49. **Prediger, E. A., S. Hoffman, G. M. Edelman, and B. A. Cunningham.** 1988. Four exons encode a 93-base-pair insert in three neural cell adhesion molecule mRNAs specific for chicken heart and skeletal muscle. Proc. Natl. Acad. Sci. USA **85**:9616–9620.
50. **Reed, R.** 1989. The organization of 3′ splice-site sequences in mammalian introns. Genes Dev. **3**:2113–2123.
51. **Reed, R., and T. Maniatis.** 1986. A role for exon sequences and splice-site proximity in splice site selection. Cell **46**:681–690.
52. **Reed, R., and T. Maniatis.** 1988. The role of the mammalian branchpoint sequence in pre-mRNA splicing. Genes Dev. **2**:1268–1276.
53. **Robberson, B. L., G. J. Cote, and S. M. Berget.** 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell. Biol. **10**:84–94.
54. **Sabina, R. L., T. Morisaki, P. Clarke, R. Eddy, T. B. Shows, C. C. Morton, and E. W. Holmes.** 1990. Characterization of the human and rat myoadenylate deaminase genes. J. Biol. Chem. **265**:9423–9433.
55. **Santoni, M. J., D. Barthels, G. Vopper, A. Boned, C. Goridis, and W. Wille.** 1989. Differential exon usage involving an unusual splicing mechanism generates at least eight types of NCAM cDNA in mouse brain. EMBO J. **8**:385–392.
56. **Seidman, J. G., and P. Leder.** 1980. A mutant immunoglobulin light chain is formed by aberrant DNA and RNA splicing events. Nature (London) **286**:779–783.
57. **Sikder, S. K., E. A. Kabat, and S. L. Morrison.** 1985. Alternative splicing patterns in an aberrantly rearranged immunoglobulin kappa-light chain gene. Proc. Natl. Acad. Sci. USA **82**: 4045–4049.
58. **Smith, C. W. J., and B. Nadal-Ginard.** 1989. Mutually exclusive splicing of alpha tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. Cell **56**:749–758.
59. **Somasekhar, M. B., and J. E. Mertz.** 1985. Exon mutations that affect the choice of splice sites used in processing the SV40 late transcripts. Nucleic Acids Res. **13**:5591–5609.

59a. **Sterner, D. A., and S. M. Berget.** Unpublished data.

60. **Talerico, M., and S. M. Berget.** 1990. Effect of splice site mutations on splicing of the preceding intron. Mol. Cell. Biol. **10**:6299–6305.
61. **Tate, V. E., M. H. Finer, H. Boedtker, and P. Doty.** 1983. Chick pro alpha 2(1) collagen gene: exon location and coding potential for the prepropeptide. Nucleic Acids Res. **11**:91–104.
62. **Treisman, R., N. J. Proudfoot, M. Shander, and T. Maniatis.** 1982. A single base change at a splice in a beta^0-thalassemic gene causes abnormal RNA splicing. Cell **29**:903–911.
63. **Tsai, A. Y. M., M. Streuli, and H. Saito.** 1989. Integrity of the exon 6 sequence is essential for tissue-specific alternative splicing of human leukocyte common antigen pre-mRNA. Mol. Cell. Biol. **9**:4550–4555.
64. **Turnbull-Ross, A. D., A. J. Else, and I. C. Eperon.** 1988. The dependence of splicing efficiency on the length of the 3′ exon. Nucleic Acids Res. **16**:394–411.
65. **Wakamatsu, N., H. Kobayashi, T. Miyatake, and S. Tsuji.** 1992. A novel exon mutation in the human beta-hexosaminidase beta subunit gene affects 3′ splice site selection. J. Biol. Chem. **267**:2406–2413.
66. **Weil, D., M. Bernard, N. Combates, M. K. Wirtz, D. W. Hollister, B. Steinmann, and F. Ramirez.** 1988. Identification of a mutation that causes exon skipping during collagen pre-mRNA splicing in an Ehlors-Danlos syndrome variant. J. Biol. Chem. **263**:8561–8564.
67. **Weil, D., M. D'Alessio, F. Ramirez, W. deWet, W. G. Cole, D. Chan, and J. F. Bateman.** 1989. A base substitution in the exon of a collagen gene causes alternative splicing and generates a structurally abnormal polypeptide in a patient with Ehlors-Danlos syndrome type VII. EMBO J. **8**:1705–1710.
68. **Weil, D., M. D'Alessio, F. Ramirez, B. Steinmann, M. K. Wirtz, R. W. Glanville, and D. H. Hollister.** 1989. Temperature dependent expression of a collagen splicing defect in the fibroblasts of a patient with Ehlors-Danlos syndrome type VII. J. Biol. Chem. **264**:16804–16809.
69. **Wieringa, B., F. Meyer, J. Reiser, and C. Weissmann.** 1983. Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit beta-globin gene. Nature (London) **301**:38–43.
70. **Wilkinson, M. J., and R. J. A. Grand.** 1978. Comparison of amino acid sequences of troponin I from different striated muscles. Nature (London) **271**:30–35.
71. **Wu, J., and J. L. Manley.** 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. Genes Dev. **3**:1553–1561.
72. **Zhang, L.-H., H. Vrieling, A. A. van Zeeland, and D. Jenssen.** 1992. Spectrum of spontaneously occurring mutations in the *hprt* gene of V79 chinese hamster cells. J. Mol. Biol. **223**:627–635.
73. **Zhuang, Y., A. M. Goldstein, and A. M. Weiner.** 1989. UAC-UAAC is the preferred branch site for mammalian mRNA splicing. Proc. Natl. Acad. Sci. USA **86**:2752–2756.
74. **Zhuang, Y., and A. M. Weiner.** 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell **46**:827–835.
75. **Zhuang, Y., and A. M. Weiner.** 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. Genes Dev. **3**:1545–1552.
76. **Zhuang, Y., H. Leung, and A. M. Weiner.** 1987. The natural 5′ splice site of simian virus 40 large T antigen can be improved by increasing the base complementarity to U1 RNA. Mol. Cell. Biol. **7**:3018–3020.
77. **Zillmann, M., M. L. Zapp, and S. M. Berget.** 1988. Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. Mol. Cell. Biol. **8**:814–821.