Regulated Splicing of the *Drosophila* P Transposable Element Third Intron in Vitro: Somatic Repression

CHRISTIAN W. SIEBEL AND DONALD C. RIO

In eukaryotic cells alternative splicing of messenger RNA precursors (pre-mRNA's) is a means of regulating gene expression. Although a number of the components that participate in regulating some alternative splicing events have been identified by molecular genetic procedures, the elucidation of the biochemical mechanisms governing alternative splicing requires in vitro reaction systems. The tissue specificity of P element transposition in *Drosophila* depends on the germline restriction of pre-mRNA splicing of the P element third intron (IVS3). *Drosophila* P element IVS3 pre-mRNA substrates were spliced accurately in vitro in heterologous human cell extracts but not in *Drosophila* somatic cell splicing extracts. Components in *Drosophila* somatic cell extracts that specifically inhibited IVS3 splicing in vitro were detected by a complementation assay. Biochemical assays for *Drosophila* RNA binding proteins were then used to detect a 97-kilodalton protein that interacts specifically with 5′ exon sequences previously implicated in the control of IVS3 splicing in vivo. Inhibition of IVS3 splicing in vitro could be correlated with binding of the 97-kD protein to 5′ exon sequences, suggesting that one aspect of IVS3 tissue-specific splicing involves somatic repression by specific RNA-protein interactions.

IN EUKARYOTIC CELLS, THE FORMATION OF FUNCTIONAL protein-coding messenger RNA's (mRNA's) often requires the removal of intervening RNA sequences (introns) from precursor mRNA (pre-mRNA) and the rejoining of coding sequences (exons) by a process called RNA splicing. The development of soluble in vitro systems has led to a detailed understanding of the biochemical mechanism of this process (*1*). The first step in splicing involves cleavage at the conserved 5′ splice site sequence with the concomitant covalent joining of a conserved guanosine (G) residue at the 5′ end of the intron via a 2′,5′-phosphodiester bond to a site within the intron known as the branch point sequence, yielding the 5′ exon (E1) and intron–3′ exon (IVS-E2) intermediates. In the second step of splicing, cleavage at the conserved 3′ splice site occurs, and the two exons are joined by ligation, creating the mature mRNA and the excised intron (IVS). Splicing occurs in large ribonucleoprotein complexes known as spliceosomes (*1*), which are assembled from abundant small nuclear ribonucleoprotein particles (snRNP's) (*2*) in an ordered, adenosine triphosphate (ATP)–dependent pathway. In addition to snRNP's, other protein factors appear to facilitate spliceosome formation, presumably by stabilizing interactions between snRNP's and the pre-mRNA (*1*).

The splicing of messenger RNA precursors serves as an important mechanism for the posttranscriptional control of eukaryotic gene expression. Numerous examples of alternative splicing illustrate that regulation can occur in a temporal, tissue-specific, or developmental fashion (*3*). Despite the success of the biochemical approach toward elucidating the mechanism of pre-mRNA splicing, little detailed knowledge has emerged regarding the molecular events that direct alternative pre-mRNA splicing. However, recent molecular genetic experiments in *Drosophila melanogaster* have begun to suggest mechanisms responsible for the regulation of alternative pre-mRNA splicing (*4*). For example, the pathway of somatic sexual differentiation is controlled by a hierarchy of genes that are regulated by alternative pre-mRNA splicing. Several of these genes carry a conserved protein motif that is involved in direct protein-RNA interactions (*5*), suggesting that the translation product of one gene directly regulates splicing of its own pre-mRNA or those of downstream genes (or both). A second example of differential splicing in *Drosophila* is the *suppressor of white apricot, su(w^a)*, gene. The *su(w^a)* gene product apparently blocks splicing of its own pre-mRNA, thereby autoregulating *su(w^a)* protein expression (*6*).

While molecular genetic approaches in *Drosophila* have proved valuable for the identification of important cis- and trans-acting regulatory factors participating in alternative splicing events, the lack of in vitro systems has hindered progress toward clarifying how these factors function. The problem of alternative splicing in vitro is best appreciated by studying a simple alternative splicing event that has been well characterized in vivo, such as the tissue-specific splicing in *Drosophila* of the P transposable element third intron (IVS3). Differential splicing of P element pre-mRNA restricts P element transposition, and hence the syndrome of detrimental genetic traits known as hybrid dysgenesis, to germline tissues (*7*). In somatic cells, the first two introns of the P element pre-mRNA are removed, but IVS3 is retained. This somatic mRNA encodes a 66-kD protein that can function as a negative regulator of transposition in both the germline and soma (*8*). By contrast, in the germline IVS3 is removed, resulting in an mRNA species encoding the 87-kD transposase protein required to catalyze P element transposition

The authors are at the Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142 and the Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

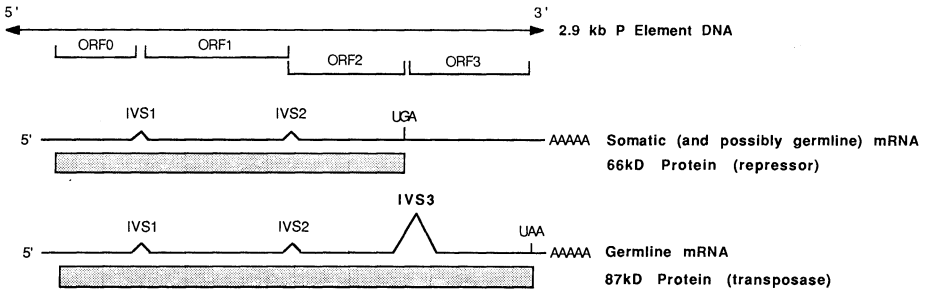


Fig. 1. Tissue-specific splicing of the third intron of P element pre-mRNA. Transcription of full-length 2.9-kb P element DNA yields a single pre-mRNA containing four long open reading frames (ORF's). The first three ORF's are joined in the mature mRNA after splicing of the first two introns (IVS1 and IVS2) in both the soma and germline. In contrast, IVS3 is removed only in the germline, thereby restricting production of the 87-kD transposase protein (indicated by the larger shaded box) to this tissue. In somatic cells, IVS3 is retained, and translation halts at a UGA stop codon within IVS3 to yield a 66-kD protein (smaller shaded box).

and excision (Fig. 1) (*9*). Genetic experiments have revealed that cis-acting regulatory sequences reside within exon sequences adjacent to the splice sites or within 26 nucleotides (nt) of intron sequence (*10*). The basis of this tissue-specific pre-mRNA splicing event is not known, but could involve negative control of IVS3 splicing in the soma, activation of IVS3 splicing in the germline, or both. We have used a biochemical complementation assay to study this tissue-specific splicing event in vitro. Our results suggest that IVS3 splicing is regulated by specific interactions between RNA binding proteins and the 5′ exon that inhibit splicing in somatic cells.

Accurate splicing of IVS3 in human HeLa cell nuclear extracts. Elucidating the mechanisms governing the tissue-specific splicing of the *Drosophila* P element third intron requires an in vitro reaction system suitable for dissecting this regulatory process biochemically. Although nuclear extracts derived from somatic *Drosophila* tissue culture cells or 0- to 12-hour embryos are active for splicing of the *fushi tarazu* (*ftz*) intron, the first two P element introns, and several other *Drosophila* and mammalian introns (*11, 12*), these extracts do not detectably process substrates containing the P element third intron (*12*) (see below). We therefore examined IVS3 splicing in heterologous human HeLa cell extracts that splice various mRNA precursors, including the *Drosophila ftz* intron (*13*). After incubation in a HeLa cell nuclear extract, an IVS3 pre-mRNA substrate was accurately spliced, yielding the expected products in an ATP-dependent reaction that followed the appropriate kinetics (Fig. 2, A and E). After a characteristic lag of approximately 30 minutes, the 5′ exon (E1) and the intron–3′ exon (I-E2) lariat appeared (Fig. 2A, lane 3). Later in the incubation (as expected), the amount of intron lariat (IVS) and mRNA surpassed the amount of 5′ exon and intron–3′ exon intermediates (Fig. 2A, lanes 6 to 8).

The designated bands (Fig. 2A) represent the products of an accurate IVS3 splicing reaction. The mRNA species was fully protected from ribonuclease (RNase) T1 digestion after hybridization with an M13 recombinant carrying a P element cDNA fragment lacking IVS3 (*12*). In addition, the putative IVS and IVS-E2 lariats were purified and incubated with a HeLa cell cytoplasmic S100 extract that contained debranching enzyme activity that specifically cleaved the 2′,5′-phosphodiester bonds present in lariat RNA's (*14*). In the absence of debranching enzyme, the putative IVS (Fig. 2B, lane 1) and IVS-E2 (Fig. 2B, lane 3) retain their lariat structures and thus migrate anomalously during gel electrophoresis. However, when treated with the debranching enzyme, the IVS (Fig. 2B, lane 2) and IVS-E2 (Fig. 2B, lane 4) migrate as linear RNA's of the expected sizes of 191 and 379 nt, respectively (Fig. 2E).

The site of branch point formation in the lariat RNA species generated during pre-mRNA splicing usually occurs at an adenosine (A) residue within a weak consensus sequence located 18 to 35 nt upstream of the 3′ splice site (*15, 16*). Because a number of alternatively spliced introns form a branch at a remote branch point (*17*), we mapped the position of the IVS3 branch point in the IVS-E2 and IVS lariat RNA's. First, the IVS and IVS-E2 lariats were treated with RNase T1 after incubation in the presence or absence of debranching enzyme. Both lariat RNA's, but not the pre-mRNA, gave rise to RNA oligonucleotides migrating at approximately 40 nt, the mobility of which was altered to 38 nt after debranching (Fig. 2C), an indication of the presence of an RNA branch. Examination of the IVS3 sequence (*18*) revealed that the only 38-nt RNase T1 oligonucleotide found in the IVS-E2 contained the consensus branch point sequence (*15*). The branch point was more precisely mapped by primer extension analysis (Fig. 2D) (*19*) with a ^{32}P-labeled DNA oligonucleotide complementary to sequences in the 3′ exon. A prominent band of approximately 87 nt was observed, consistent with an RNA branch blocking reverse transcription at the consensus branch point position at nucleotide 2114 (*18*) (Fig. 2D, lane 2). This 87-nt reverse transcription product was not observed when debranched IVS-E2 was used as the template (Fig. 2D, lane 3), confirming the conclusion that the 87-nt product resulted from the presence of an RNA branch. The branch point mapped here in mammalian extracts is likely to be the same as that used in *Drosophila* since the *Drosophila fushi tarazu* branch point is identical in both mammalian and *Drosophila* extracts (*11, 13*). Our data (Fig. 2) show that, unlike the splicing of other regulated introns, IVS3 splicing does not involve branch point formation at a distal position (*17*).

Specific inhibition of IVS3 splicing by somatic *Drosophila* extract. Because IVS3 was not spliced in somatic *Drosophila* splicing extracts but was accurately spliced in human cell extracts and microinjected *Xenopus* oocytes (*12*), regulation of IVS3 splicing could involve a somatic inhibitory activity. We therefore developed an in vitro biochemical complementation assay for factors that regulate IVS3 splicing in which we examined the effect of splicing extract from *Drosophila* somatic Kc cells (*20*) on IVS3 splicing in HeLa cell extract. The IVS3 pre-mRNA was first incubated with increasing amounts (0 to 60 μg of protein) of *Drosophila* Kc extract to allow potential regulatory factors to interact with the pre-mRNA. A constant amount (60 μg) of HeLa splicing extract was added, and incubation was continued. Small amounts of Kc extract (10 to 20 μg) had little effect on IVS3 splicing (Fig. 3, A and B, lanes 1 and 2); intermediate amounts (30 to 40 μg) inhibited IVS3 splicing by at least 50 percent (Fig. 3, A and B, lanes 3 and 4). When the amount of Kc extract (50 to 60 μg) was nearly equal to the amount of HeLa extract, splicing was completely inhibited (Fig. 3, A and B, lanes 5 and 6). The addition of increasing amounts of *Drosophila* Kc extract inhibited the production of each of the splicing intermediates and products (Fig. 2E), suggesting that the inhibition occurred at or before cleavage at the 5′ splice site. The observed inhibitory effect was not due to increased RNA degradation catalyzed by the added *Drosophila* Kc extract, as indicated by control experiments with other introns (see below) and the observation that an equal amount of RNA is present in each lane. The inhibitory effect was not due to changes in monovalent cation concentration, as the conductivity of both extracts was monitored, the salt concentration in each reaction

was kept constant, and IVS3 splicing in HeLa extracts was unaffected by a KCl concentration 3.5 times higher than that in the standard reaction (*12*). Heating the *Drosophila* Kc cell extract for 5 minutes at 100°C destroyed the inhibitory effect, suggesting that a protein component was involved (*12*).

Other introns were used in the inhibition assay to test the specificity of the inhibitory effect. The *Drosophila fushi tarazu* (*ftz*) intron was efficiently spliced in both *Drosophila* Kc and HeLa extracts (Fig. 4A). Addition of Kc extract, even in high amounts, to the HeLa splicing reaction had no effect on splicing of the *ftz* intron (Fig. 4A) or on the rabbit β-globin second intron (Fig. 4B). Splicing of a *Drosophila* actin 5C gene intron, which occurs inefficiently in HeLa cell extract, was also not inhibited by the addition of Kc extract (*12*). P element IVS3, the β-globin intron, and the *Drosophila* actin 5C intron were not spliced in Kc extract; however, of these introns, only splicing of IVS3 in HeLa cell extract was inhibited by Kc extract (Figs. 3 and 4B) (*12*). This observation rules out the possibility that inhibition of IVS3 splicing occurs because the *Drosophila* splicing machinery nonproductively interacts with the pre-mRNA during the initial incubation and prevents the HeLa splicing components from productively binding the pre-mRNA. Of all the control introns tested, only splicing of the adenovirus L1-2 intron was inhibited, but less so than IVS3 splicing (Fig. 4C). The adenovirus L1-2 intron, like the *Drosophila Sex-lethal* (*Sxl*) male-specific and *transformer* (*tra*) non–sex-specific 3′ splice sites, contains an unusual sequence of eight uridine residues and a cytosine residue (U_8C) in its polypyrimidine tract (*21*). Since binding of the *Sxl* protein is proposed to inhibit splicing by binding to these splice sites (*22*), *Sxl* expression in *Drosophila* Kc cells could account for the observed inhibition of splicing of the adenovirus L1-2 intron (*20, 23*). Alternatively, our RNA binding studies (Fig. 5) suggest that the adenovirus L1-2 intron binds the same protein that appears to

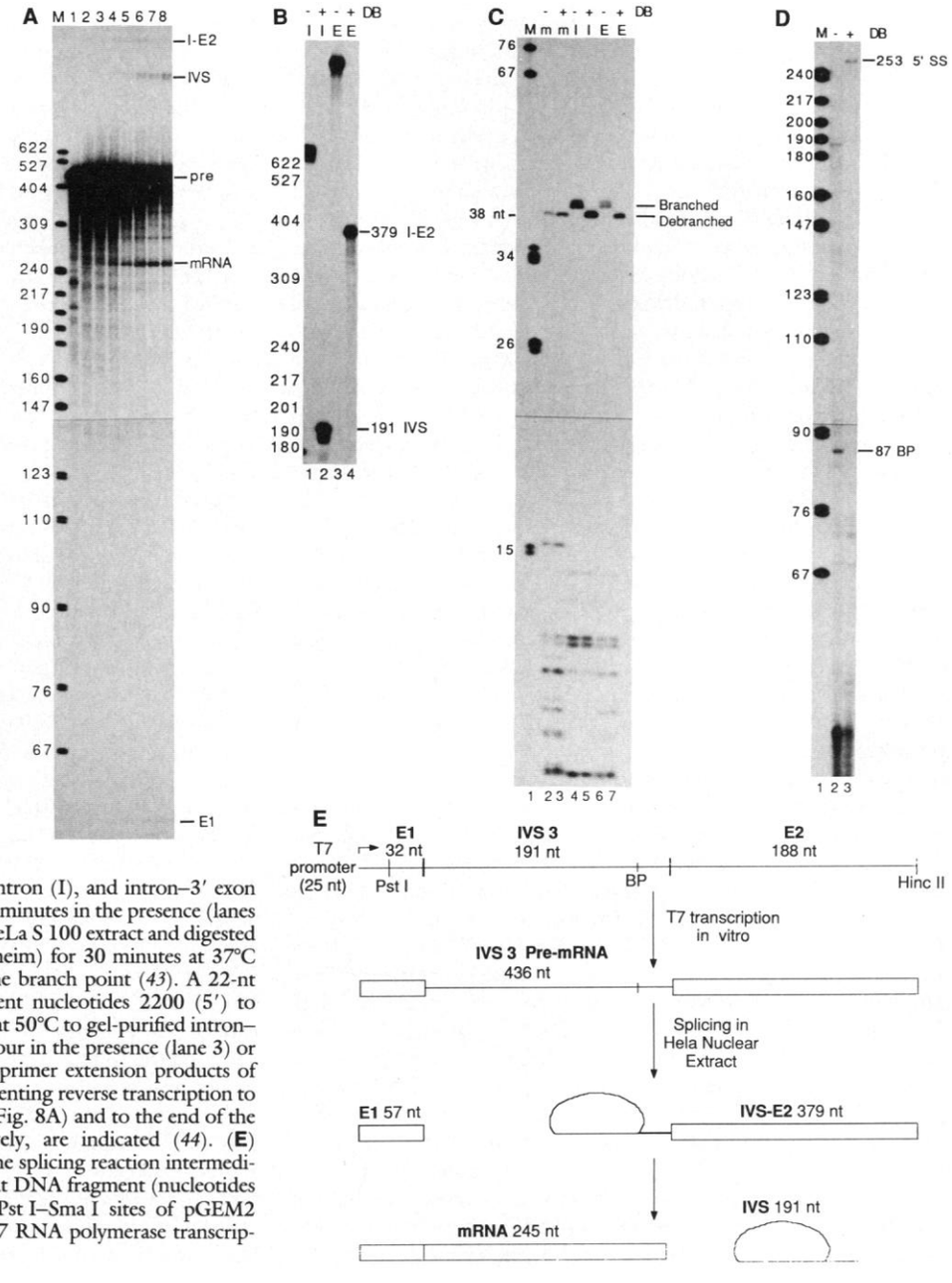


Fig. 2. The P element third intron is accurately spliced in HeLa cell nuclear extract. Reaction products were visualized by autoradiography after electrophoresis in denaturing polyacrylamide gels. (M, ^{32}P-labeled markers of Msp I-digested pBR322 DNA.) (**A**) Time course of the splicing reaction. IVS3 pre-mRNA (see part E) was incubated in a HeLa cell nuclear extract in the absence (lane 1) or presence (lanes 2 to 8) of ATP at 30°C: lane 1, 180 minutes; lane 2, 15 minutes; lane 3, 30 minutes; lane 4, 45 minutes; lane 5, 60 minutes; lane 6, 90 minutes; lane 7, 120 minutes; lane 8, 150 minutes. HeLa nuclear extract was made as described (*41*), except that 2.7 ml of buffer C containing 0.6 M NaCl was used. Reactions contained 80,000 cpm (2×10^6 cpm/pmol) of pre-mRNA, 20 percent (v/v, 100 μg of total protein) nuclear extract, 2.5 percent (w/v) polyvinyl alcohol, 5 mM $MgCl_2$, 10 units of RNasin (Promega), 3 mM ATP, 5 mM creatine phosphate, 27 mM Hepes (pH 7.6), 7.2 percent (v/v) glycerol, 62 mM KCl, 72 μM EDTA, and 0.18 mM dithiothreitol (DTT). Positions of the input pre-mRNA (pre), the intermediates of the reaction (5′ exon, E1; intron–3′ exon, I-E2) and the products of the reaction (mRNA and intron, IVS) are indicated. (**B**) The intron (I) and intron–3′ exon (E) species were purified from a gel similar to that shown in (A) and incubated for 1 hour in the absence (lanes 1 and 3) or presence (lanes 2 and 4) of HeLa S100 extract (*14, 41*). The numbers to the left of the gel show the positions of the size markers described above. (**C**) Debranching and T1 ribonuclease digestion analysis to map the branch point. The pre-mRNA (p), intron (I), and intron–3′ exon (E) species were gel-purified, incubated for 75 minutes in the presence (lanes 3, 5, and 7) or absence (lanes 2, 4, and 6) of HeLa S 100 extract and digested with 1 unit of RNase T1 (Boehringer Mannheim) for 30 minutes at 37°C (*42*). (**D**) Primer extension analysis to map the branch point (*43*). A 22-nt ^{32}P-labeled primer corresponding to P element nucleotides 2200 (5′) to 2179 (3′) (*18*) was hybridized for 90 minutes at 50°C to gel-purified intron–3′ exon RNA that had been incubated for 1 hour in the presence (lane 3) or absence (lane 2) of HeLa S100 extract. The primer extension products of about 87 nt (lane 2) and 253 nt (lane 3) representing reverse transcription to the branch point (BP) at position 2114 (*18*) (Fig. 8A) and to the end of the linearized intron–3′ exon (5′ SS), respectively, are indicated (*44*). (**E**) Structure of the input IVS3 pre-mRNA and the splicing reaction intermediates and products. The Pst I–Hinc II P element DNA fragment (nucleotides 1911 to 2316) (*18*) was inserted into the Pst I–Sma I sites of pGEM2 (Promega) and then used as a template for T7 RNA polymerase transcription with [α-^{32}P]GTP (*45*).

be involved in inhibition of IVS3 splicing. Therefore, splicing of the adenovirus L1-2 intron may be inhibited by the same mechanism as is splicing of IVS3. Indeed, the mechanism governing P element IVS3 tissue-specific splicing probably operates on other *Drosophila* genes because most laboratory strains lack P elements (*24*).

Preferential interaction of a 97-kD protein with IVS3. Putative inhibitors of splicing of IVS3 might bind to IVS3 RNA preferentially compared to other introns that are not subject to inhibition. Therefore, ultraviolet (UV) cross-linking was used to detect RNA binding proteins that interact specifically with IVS3 pre-mRNA (*25*). IVS3 and the other introns tested in the inhibition assay were incubated in Kc extract under splicing conditions, exposed to UV at 254-nm (a wavelength that covalently cross-links RNA-protein complexes) and then briefly treated with RNase A to remove the excess RNA that would alter the mobility of the cross-linked proteins in SDS-polyacrylamide gels. The IVS3 pre-mRNA cross-linked conspicuously to several polypeptides (Fig. 5, lane 3). The 40-kD species, which may correspond to the heterogeneous nuclear (hn) RNP C proteins (*26*), the 65-kD protein, which may recognize 5′ splice site consensus sequences (see below), and a number of other proteins cross-linked to all introns tested; therefore, they are not good candidates for a potential inhibitor. In contrast, only a 97-kD protein preferentially cross-linked to IVS3. In fact, cross-linking of the 97-kD protein to some of the other introns correlated with the degree of inhibition (Figs. 3 to 5). For example, the 97-kD protein cross-linked strongly to the IVS3 pre-mRNA, weakly to the adenovirus L1-2 pre-mRNA, and not at all to the *ftz* pre-mRNA (Fig. 5).

Competition experiments with unlabeled competitor RNA's confirmed that the differences in cross-linking signals reflected differences in binding and not label transfer (*12, 27*). However, one exception to the correlation between cross-linking the 97-kD protein and inhibition was observed with the β-globin intron. The β-globin pre-mRNA weakly cross-linked to the 97-kD protein (Fig. 5, lane 2), but splicing of this intron was not affected in the inhibition assay (Fig. 4B), implying that binding of the 97-kD protein in itself is not sufficient for inhibition, that the proper location of the 97-kD protein binding site within the pre-mRNA is critical for inhibition or that the 97-kD protein acts in concert with other factors not bound by the β-globin pre-mRNA.

In order to localize the 97-kD protein binding site on the IVS3 pre-mRNA, we used UV cross-linking to analyze a series of IVS3 RNA's with deletions from the 3′ end (*12*). A small RNA oligonucleotide from the IVS3 5′ exon termed Δ11 cross-linked to the 65-kD and 97-kD proteins (Fig. 5, lane 5). However, other regions within the intron may also have an affinity for these proteins (*12*). This Δ11 RNA fragment contains 35 nt of 5′ exon sequence ending 2 nt upstream from the IVS3 5′ splice site in addition to polylinker sequence (Fig. 7C). This exon region has been implicated genetically in the tissue specificity of IVS3 splicing (*10*), supporting the idea that a putative inhibitor might bind to the Δ11 RNA fragment. The pattern of bands observed when HeLa extract was used for cross-linking in place of Kc extract was similar in that polypeptides of 35 to 40 kD, 45 to 50 kD, 66 kD, and 116 kD cross-linked to Δ11 in both extracts; however, no band near 97 kD was observed with HeLa extract (Fig. 5, lane 6). Taken together, the UV cross-linking results suggest a correlation between inhibition of IVS3 splicing in vitro and interaction between a 97-kD RNA binding protein and IVS3 5′ exon sequences.

RNA that binds the 97-kD protein relieves inhibition of IVS3 splicing. The UV cross-linking data indicated that the 97-kD protein interacts preferentially with IVS3 pre-mRNA within the 5′ exon sequences in Δ11 RNA. If binding of the 97-kD or other *Drosophila* proteins to these sequences is responsible for inhibition of IVS3 splicing in HeLa extracts, then addition of excess Δ11 RNA might be expected to relieve the inhibition of IVS3 splicing in vitro

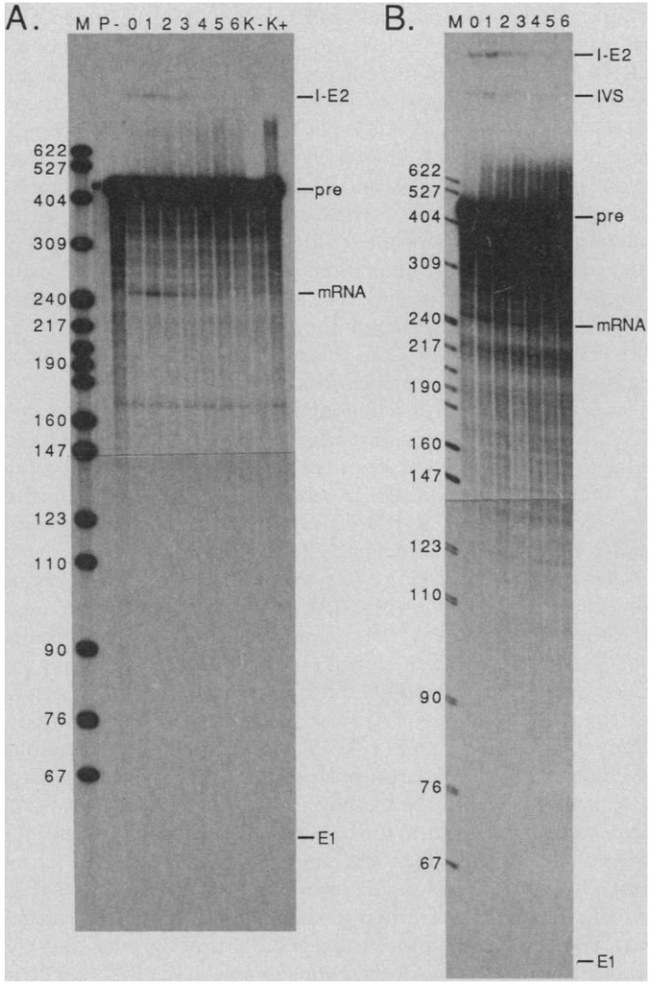


Fig. 3. *Drosophila* somatic cell nuclear extract inhibits splicing of the P element third intron in HeLa nuclear extract. IVS3 pre-mRNA (Fig. 2E) was first incubated in the absence of HeLa nuclear extract for 10 minutes at 20°C in a volume of 22 μl in the presence (lanes 0 to 6 and K+) or absence (lanes − and K−) of ATP under the splicing conditions described in Fig. 2A with the following amounts (v/v, total protein) of *Drosophila* somatic Kc nuclear extract (*11*): lanes − and 0, no Kc added; lane 1, 2.3 percent, 10 μg; lane 2, 4.5 percent, 20 μg; lane 3, 6.8 percent, 30 μg; lane 4, 9.1 percent, 40 μg; lane 5, 11.4 percent, 50 μg; lane 6, 13.6 percent, 60 μg; lanes K− and K+, 27 percent, 120 μg. HeLa nuclear extract (3 μl, 60 μg, 12 percent v/v) was then added (lanes − and 0 to 6), and incubation was continued for 105 minutes at 30°C. No HeLa extract was added to reactions K− and K+, which were allowed to incubate an additional 105 minutes at 20°C. Reaction conditions other than these have been tested, but no splicing of IVS3 in *Drosophila* Kc extract has been observed in this assay. P, input pre-mRNA. M, ^{32}P-labeled markers of Msp I–digested pBR322 DNA. The positions of the relevant species are indicated as in Fig. 2A. The heterogeneous population of RNA running above and adjacent to the pre-mRNA represents the products of an ATP-dependent polymerization reaction , catalyzed by the Kc extract. These products were not as prevalent when other pre-mRNA's were used (Fig. 4). (**A**) The products were detected by autoradiography after electrophoresis on a denaturing 7 percent polyacrylamide gel. The mRNA and pre-mRNA are resolved well, but the intron and pre-mRNA migrate together. (**B**) Same as (A), except that a denaturing 10 percent polyacrylamide gel was used to resolve the intron and the pre-mRNA.

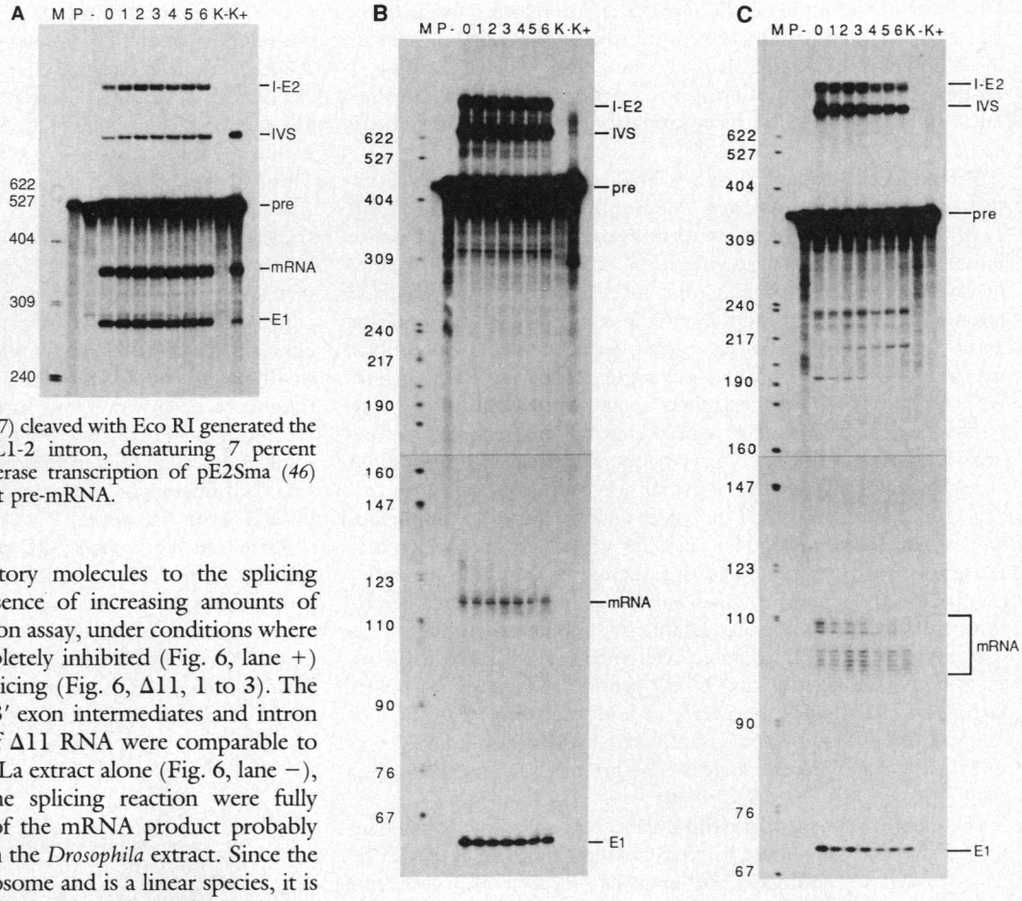


Fig. 4. Analysis of other *Drosophila* and mammalian introns in the somatic extract inhibition assay. Experiments were conducted as in Fig. 3, except that other pre-mRNA's were used and the $MgCl_2$ concentration was 3 mM for (B) and (C). The same extracts used for the experiment shown in Fig. 3 were used in these experiments. (**A**) *Drosophila fushi tarazu* (*ftz*), denaturing 12 percent polyacrylamide gel. T7 RNA polymerase transcription of pGEM2 V61 S/B (*11*) cleaved with Xho I generated the input RNA. (**B**) Rabbit β-globin second intron, denaturing 5 percent polyacrylamide gel. T3 RNA polymerase transcription of pBSAL4 (*37*) cleaved with Eco RI generated the input pre-mRNA. (**C**) Adenovirus L1-2 intron, denaturing 7 percent polyacrylamide gel. SP6 RNA-polymerase transcription of pE2Sma (*46*) cleaved with Sma I generated the input pre-mRNA.

by preventing binding of inhibitory molecules to the splicing substrate RNA. Indeed, the presence of increasing amounts of unlabeled Δ11 RNA in the inhibition assay, under conditions where splicing of IVS3 is normally completely inhibited (Fig. 6, lane +) resulted in restoration of IVS3 splicing (Fig. 6, Δ11, 1 to 3). The levels of the 5′ exon and intron–3′ exon intermediates and intron product formed in the presence of Δ11 RNA were comparable to those observed after splicing in HeLa extract alone (Fig. 6, lane −), indicating that both steps of the splicing reaction were fully restored. The partial restoration of the mRNA product probably reflects the presence of nucleases in the *Drosophila* extract. Since the mRNA is released from the spliceosome and is a linear species, it is the splicing product most susceptible to degradation by nucleases. No relief of inhibition of IVS3 splicing was observed when comparable molar amounts of the two control RNA's PL1 or PL2 were used (Fig. 6, lanes PL1, 1–3 and PL2, 1–3; and Fig. 7C), indicating the specificity of the effect. These results implicate binding of an inhibitory factor, perhaps the 97-kD protein, to the Δ11 IVS3 5′ exon sequences in the inhibition of IVS3 splicing by somatic *Drosophila* extracts. Also, since Δ11 contains only one 29-nt stretch of P element 5′ exon sequence not present in PL1 (Fig. 7C), and PL1 fails to relieve inhibition, this experiment maps the cis-acting regulatory region to these nucleotides.

Since only Δ11 RNA relieved inhibition of splicing of IVS3, this RNA must interact specifically, when compared to the RNA's that failed to relieve inhibition, with the factor (or factors) responsible for inhibition. Thus, in order to detect RNA binding proteins that bind specifically to Δ11 RNA and thereby identify the possible inhibitor protein (or proteins), Δ11 RNA and the two control RNA's PL1 and PL2 (Fig. 7C) were used in UV cross-linking competition experiments (Fig. 7A) and native gel electrophoresis RNA binding assays (Fig. 7B). In the UV-cross-linking experiment, Δ11 RNA cross-linked to several polypeptides; however, only the 97-kD protein and two 40-kD species did not cross-link to either of the control RNA's (Fig. 7A). The specificity of binding of these three proteins to labeled Δ11 RNA was confirmed by competition experiments in which a large molar excess of unlabeled competitor RNA was included in the binding reaction. Excess unlabeled Δ11 RNA competed efficiently for binding of both the 97-kD and the two 40-kD proteins (Fig. 7A, lanes 3 to 6), but neither excess unlabeled PL1 RNA (Fig. 7A, lanes 7 to 10) nor PL2 RNA (Fig. 7A, lanes 13 to 17) competed for binding of these three proteins, even when present at a molar excess 3200 times greater than the labeled substrate RNA. Of these three proteins, the 97-kD protein appears to be the best candidate for a negative regulatory molecule. The two 40-kD species bind to all introns tested (Fig. 5) and may correspond to the hnRNP C proteins based on their molecular mass (*26*) and fractionation properties on cesium chloride gradients (*12*). In addition, cross-linking to the different labeled RNA's and the competition experiments revealed that both Δ11 RNA and PL1 RNA, but not PL2 RNA, bound specifically to a 65-kD protein (Fig. 7A). Finally, the other polypeptides of 50-kD, 80-kD, and 130-kD that cross-link to Δ11 RNA appear to be general, nonspecific RNA binding proteins since they also cross-link to and are competed by PL1 RNA or PL2 RNA or both (Fig. 7A).

The interaction of proteins with these three RNA fragments was also assayed by a native gel electrophoresis assay; protein-RNA complexes were formed in solution and then separated from free RNA by polyacrylamide gel electrophoresis (PAGE) at low ionic strength. Both PL1 and Δ11 RNA's, but not PL2 RNA, formed protein-RNA complexes after incubation with *Drosophila* somatic cell nuclear extract (Fig. 7B). The specificity of these RNA-protein interactions was then tested by competition analysis. The three complexes (labeled A, B, and C in Fig. 7B) formed on labeled Δ11 RNA were specifically competed by increasing amounts of unlabeled Δ11 RNA but not by unlabeled PL1 RNA or PL2 RNA. These results indicated that specific RNA-protein complexes were formed between components in the *Drosophila* extract and Δ11 RNA. These interactions presumably reflect, at least in part, binding of the 97- and 65-kD proteins identified by UV cross-linking. Analysis of the protein-RNA complexes on a SDS-PAGE second dimension indicated that the 65-kD protein is present in the two prominent complexes observed with Δ11 RNA (Fig. 7B, complexes A and B) and that the 97-kD protein may be in one of the prominent complexes (Fig. 7B, complex B) as well as in the

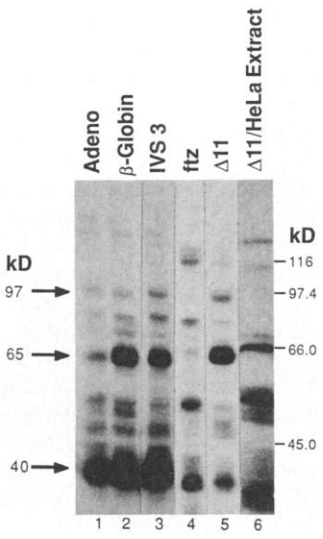


Fig. 5. Ultraviolet cross-linking assay of introns tested in the inhibition assay. [^{32}P]GTP-labeled substrate RNA's from the appropriate transcription templates were made as described (Figs. 2E and 4) (*45, 47*). For UV cross-linking (*25*) the indicated substrate RNA (about 150,000 cpm; 2×10^6 cpm/pmol) was incubated for 1 hour at 20°C in a 25-μl reaction mixture containing *Drosophila* somatic Kc extract (20 percent v/v, 100 μg of total protein) as in Fig. 2A, except that 3 mM $MgCl_2$ was used. The samples were then irradiated for 15 minutes [254 nm; 3 mW/cm^2 on the lamp surface, UVP model UVG-54] at 3.5 cm from the surface. Ribonuclease A (Boehringer Mannheim) was added to 1 mg/ml, and the samples were incubated for 30 minutes at 30°C. SDS-PAGE sample buffer was then added, and the samples were boiled for 5 minutes and placed on a discontinuous 7.5 percent SDS-polyacrylamide gel (*48*). The gel was fixed and stained with Coomassie blue. ^{32}P-labeled proteins were detected by autoradiography. Lanes 1 to 3 are from the same gel, and lanes 4 and 5 are from another. The mobility of the cross-linked proteins varied slightly among experiments, probably because of small differences in RNase digestion (*25*).

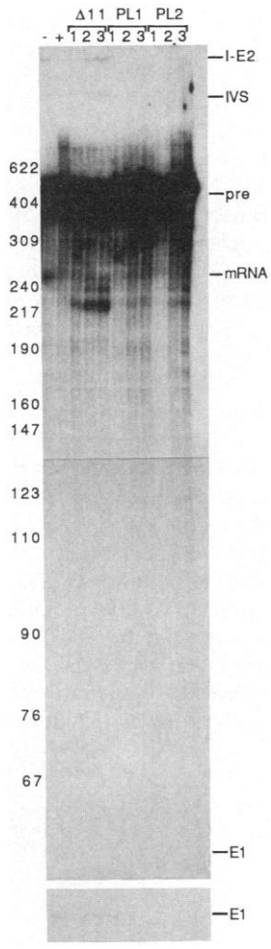


Fig. 6. Relief of inhibition of splicing of P element IVS3 in vitro by an RNA oligonucleotide that binds the 97-kD protein. Inhibition was assayed in the absence (lane −) or presence (all other lanes) of *Drosophila* Kc extract (10 percent, v/v, 50 μg) as in Fig. 3. Numbered lanes also contained the following molar excess, relative to the ^{32}P-labeled pre-mRNA, of the indicated unlabeled competitor RNA oligonucleotide: lanes 1, 3000; lane 2, 6000; lane 3, 9000. The requirement for such a large excess of the competitor RNA's to achieve an effect may reflect the presence of abundant, nonspecific RNA binding proteins or nucleases in the extract (*25*) or may indicate that the labeled substrate RNA contains multiple binding sites for the inhibitory factor. Also, since the reaction to synthesize the [^{32}P]GTP-labeled substrate RNA was conducted in the presence of the cap analog diguanosine triphosphate, which may have been incorporated into the nascent RNA and thus reduced the specific activity, the calculated amount of substrate RNA could have been an underestimate (*45*). The sequences of the competitor RNA's are shown in Fig. 7C. Synthesis of Δ11 RNA is described in Fig. 5 (*47*). T7 transcription of pGEM2 (Promega) cleaved with Eco RI generated PL1. PL1 and Δ11 share extensive sequence homology (Fig. 7C). SP6 transcription of pGEM2 cleaved with Hind III generated PL2 with a sequence unrelated to both Δ11 and PL1. ^{32}P-Labeled RNA's were detected by autoradiography after electrophoresis on a denaturing 10 percent polyacrylamide gel. The positions of the relevant RNA species and the size markers to the left of the gel are as in Fig. 2A. The lower panel is a darker exposure of the lower portion of the gel and has been included to better reveal levels of the 5′ exon.

slowest migrating complex (Fig. 7B, complex C) (*12*). Our data indicate that the 97-kD protein interacts specifically with Δ11 RNA. The 65-kD protein, which interacts with both Δ11 and PL1 RNA's, may recognize the 5′ splice site–like sequences in these RNA's (Fig. 7C). Chemical modification of the bases comprising the 5′ splice site–like sequence in Δ11 RNA interferes with formation of the complex that contains the 65-kD protein (*12*). Furthermore, this protein cofractionates with the major snRNP's on cesium chloride gradients, suggesting that it may be associated with or an integral component of an RNP complex (*12*).

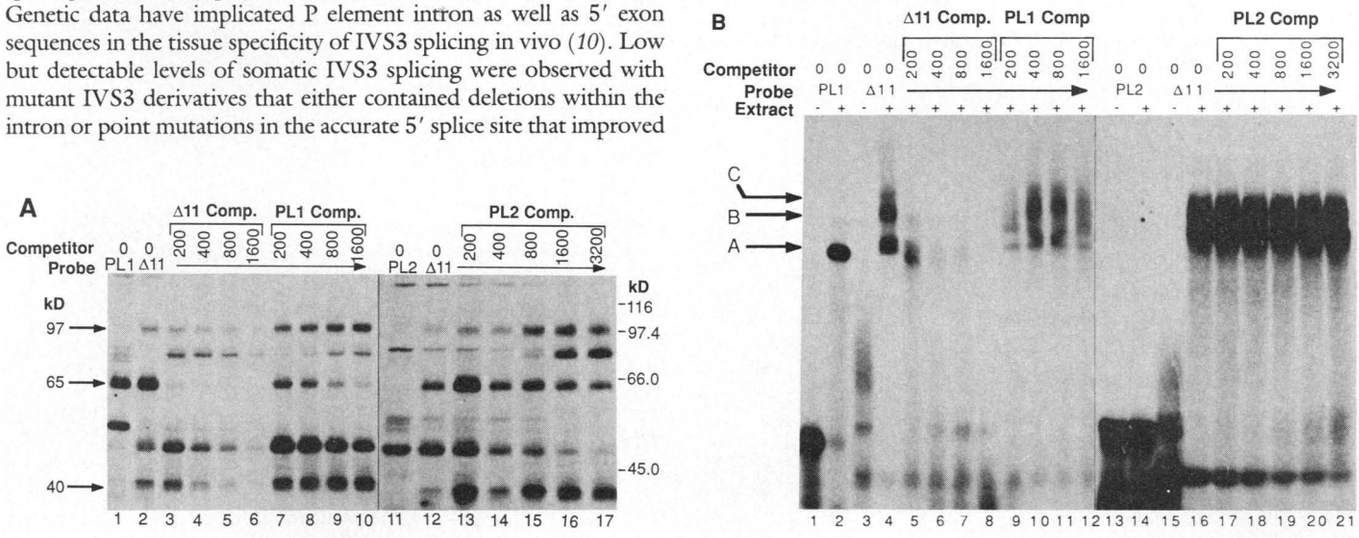
Exon sequences and control of P element IVS3 splicing. Although no other case of regulated splicing in *Drosophila* has been analyzed biochemically, it is useful to compare the somatic inhibition of P element IVS3 splicing to other regulated splicing events in *Drosophila*. Regulated splicing in the sexual differentiation pathway involves differential use of 3′ splice sites (*4*). However, negative control of splicing of the *suppressor of white apricot* first and second introns, like the rat troponin T and α-tropomyosin genes (*17*), apparently involves use of distal branch point sequences (*6*). The regulatory factors in these systems apparently control splicing by recognizing sequences within the intron such as the 3′ splice site or branch point-polypyrimidine tract (*4, 6, 22*) or in the neighboring 3′ exon (*28*). By contrast, negative control of P element IVS3 splicing in vitro involves 5′ exon sequences that interact with specific RNA binding proteins. Mutagenesis studies have also implicated these same exon sequences in the in vivo control of IVS3 splicing (*10*).

The fact that exon sequences could play a role in alternative splicing had been established by in vitro experiments where RNA substrates with duplicated 5′ splice sites, 3′ splice sites (or both), including variable lengths of flanking exon sequence, were spliced in HeLa cell extracts. These experiments showed that varying exon length and sequence could dramatically alter splice site selection (*29*). Other examples have also implicated exon sequences in alternative splicing (*3, 30*), but the biochemical mechanisms of these effects are not clear. Inhibition of IVS3 splicing in vitro can be relieved by competition with excess 5′ exon RNA sequences, suggesting that those sequences interact with RNA binding proteins that negatively regulate IVS3 splicing in the soma. Among the proteins that cross-linked to the IVS3 5′ exon, only a 97-kD protein cross-linked specifically to this RNA fragment and bound preferentially to the IVS3 pre-mRNA splicing substrate. In fact, specific binding of the 97-kD protein to the IVS3 5′ exon correlated with inhibition of splicing. Definition of the 97-kD protein binding site should provide insights into how the 97-kD protein might act to inhibit IVS3 splicing and whether additional 97-kD protein binding sites lie elsewhere in the IVS3 pre-mRNA. In the simplest model, specific binding of the 97-kD protein to IVS3 pre-mRNA might inhibit splicing by sterically occluding binding of positive splicing factors. Since the 97-kD protein binds in an exon fragment immediately adjacent to the 5′ splice site, the most likely candidates for such positive splicing factors are those implicated in the recognition of 5′ splice sites, including U1 snRNP (*31, 32*). In this respect, it is potentially important for elucidating the mechanism of inhibition that we have also detected binding of a 65-kD protein that seemed to recognize 5′ splice sites. The 97-kD protein may act in conjunction with other factors, perhaps as part of a specific inhibitory complex, or by modifying the activity of general splicing components.

Inhibition of IVS3 splicing in vitro occurs before the first step of splicing, cleavage at the 5′ splice site and lariat formation. Therefore, the 97-kD protein may act by blocking one of the early steps in spliceosome assembly. These steps include recognition of the 5′

splice site and branch point by U1 and U2 snRNP's, respectively (*1, 33, 34*). Both of these interactions are facilitated by other protein factors (*1, 32, 34–36*). Formation of pre-splicing complexes appears to require interaction between the snRNP complexes formed at the 5′ splice site and branch point (*36–38*). Studies of the 97-kD binding and spliceosome assembly experiments should clarify at which step in the splicing pathway the IVS3 block occurs. It is also possible that a block at the earliest stages of spliceosome assembly might result in increased transport of unspliced IVS3-containing RNA to the cytoplasm, an idea that is supported by experiments in yeast and with the HIV *rev* protein (*39*).

5′ splice site competition model. The IVS3 5′ exon sequences implicated in the regulation of IVS3 splicing in vivo and in vitro contain a 5′ splice site–like sequence adjacent to the accurate 5′ splice site. Nearby 5′ exon and intron sequences contain additional 5′ splice site sequences (*18*). The presence of these six 5′ splice site–like sequences (referred to as pseudo–5′ splice sites, two in the 5′ exon and four in the intron; see Fig. 8) within 196 nt surrounding the accurate IVS3 5′ splice site raises the possibility that competition between these sites, either for binding of U1 snRNP or other splicing factors, may play a role in the inhibition of IVS3 splicing. Genetic data have implicated P element intron as well as 5′ exon sequences in the tissue specificity of IVS3 splicing in vivo (*10*). Low but detectable levels of somatic IVS3 splicing were observed with mutant IVS3 derivatives that either contained deletions within the intron or point mutations in the accurate 5′ splice site that improved its match to the 5′ splice site consensus sequence (*10*). Indeed, the low level of somatic IVS3 splicing observed in several IVS3 deletion mutants in vivo might have been due to loss of some of the internal pseudo–5′ splice sites. The possibility that the remaining pseudo–5′ splice sites retained some inhibitory influence may explain why such a low level of somatic splicing activity was observed. The notion of redundant regulatory signals to efficiently inhibit somatic IVS3 splicing is not surprising given the potentially lethal effects of somatic transposition. Such reiterated negative regulatory signals may preclude the simple transfer of this splicing control to heterologous introns. The idea of nearby 5′ splice site sequences influencing splice site utilization is supported by in vitro experiments in the mammalian globin system, where the juxtaposition of an introduced 5′ splice site sequence near an authentic site inhibited splicing from both the accurate and introduced sites, even though both sites appeared to bind U1 snRNP (*40*). The efficiency with which the introduced site was used correlated with its match to the 5′ splice site consensus sequence, presumably because an improved match allowed for increased binding of U1 snRNP by increasing comple-



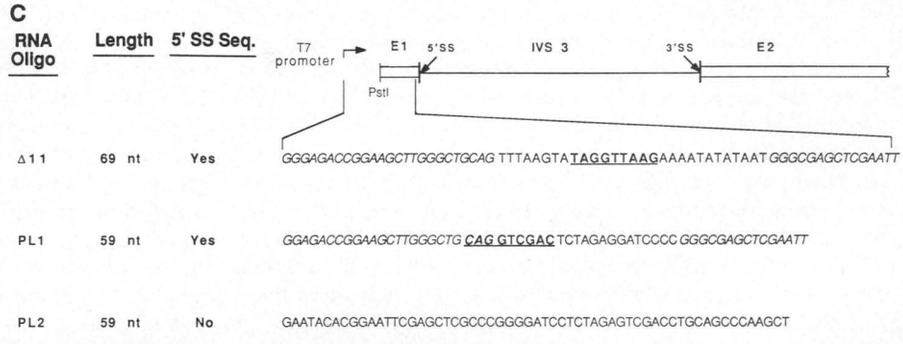


Fig. 7. Biochemical assays for RNA binding proteins. [^{32}P]GTP-labeled substrate RNA's and unlabeled competitor RNA's were made as described (Fig. 6) (*45*). The ^{32}P-labeled substrate RNA (probe) in each lane is Δ11 (C) except where indicated otherwise. Numbers above the lanes denote molar excess, relative to the probe, of the unlabeled competitor (Comp.) RNA's (Fig. 6). (**A**) Ultraviolet photochemical cross-linking. The indicated ^{32}P-labeled RNA's (about 80,000 cpm; 5×10^5 cpm/pmol), together with or without (0) the indicated molar excess of unlabeled competitor RNA, were subjected to incubation and UV cross-linking as in Fig. 5. The labeled proteins seen when [^{32}P]UTP-labeled substrate was used in place of [^{32}P]GTP-labeled substrate were the same, suggesting that all RNA binding proteins detectable by UV cross-linking were visualized after cross-linking to the GTP-labeled RNA's as shown. (**B**) Native gel electrophoresis. The indicated ^{32}P-labeled RNA (probe, about 40,000 cpm; 5×10^5 cpm/pmol) together with or without (0) the indicated molar excess of unlabeled competitor RNA was incubated for 80 minutes under the splicing conditions described in Fig. 5 in the absence (−) or presence (+) of *Drosophila* somatic Kc extract (20 percent v/v, 100 μg of protein). Xylene cyanol and bromophenol blue (BPB) dyes were each added to 0.1 percent (w/v), and the samples were subjected to electrophoresis on a native 4.2 percent polyacrylamide (acrylamide:bis = 60:1) gel (17 by 14.7 by 0.15 cm) at 8W until the BPB was approximately 2 cm from the bottom of the gel (*49*). Three RNA-protein complexes, labeled A, B, and C, as well as the free RNA at the bottom of the gel, were detected by autoradiography. (**C**) Sequence of the RNA oligonucleotides. The diagram at the top shows the position of Δ11 relative to IVS3. Δ11 contains 35 nt of 5′ exon sequence ending 2 nt upstream from the 5′ splice site (SS) in addition to 19 nt and 15 nt from the pGEM2 polylinker at its 5′ and 3′ ends, respectively. The lengths of the RNA's and the positions of the 5′ SS sequences (bold, underlined) are also indicated. The 5′ SS sequences in Δ11 and PL1 contain the absolutely conserved GU dinucleotide at positions 4 and 5 of the consensus sequence and an overall match of 5 nt out of the 9-nt consensus sequence (Fig. 8B) (*50*). Other regions of sequence homology between Δ11 and PL1 are shown in italics. PL2 contains a GU dinucleotide at only one position, and at this position only one other nucleotide matches the consensus sequence; thus, PL2 is described as lacking a 5′ SS sequence.

mentarity to U1 RNA. This observation may explain why some somatic IVS3 splicing was observed in vivo when point mutations in the accurate 5′ splice site improved its match with the consensus sequence, thereby allowing better U1 recognition of the accurate IVS3 5′ splice site relative to the pseudo–5′ splice sites (*10*). In the presence of a wild-type 5′ splice site, the pseudo–5′ splice sites, although not utilized to make functional mRNA's in vivo (*7*), may compete efficiently with the accurate 5′ splice site since they match quite favorably to the consensus sequence; in fact, some of the pseudo–5′ splice sites match the 5′ splice site consensus better than the accurate site (Fig. 8B). The 97-kD protein may act to negatively regulate IVS3 splicing by either sterically excluding U1 snRNP or other factors from recognizing the accurate IVS3 5′ splice site or by actively channeling these factors to the pseudo–5′ splice sites. Taken together, these observations suggest that 5′ splice site recognition factors as well as intron and 5′ exon sequences may be involved in the somatic control of IVS3 splicing.

The 97-kD protein, since it was detected in extracts of *Drosophila* somatic cells derived from a strain lacking P elements, likely modulates splicing of other pre-mRNA's during development. Our in vitro results, which suggest that regulation of P element IVS3 splicing involves inhibition of splicing in the soma, predict that the 97-kD protein (and possibly other factors) may be found in an active form only in somatic cells (*7*). Thus, the mechanism of negative control governing IVS3 splicing may normally operate as part of the germline-soma dichotomy in *Drosophila*. This type of control may serve to regulate the splicing of germline-specific genes in the soma, perhaps genes normally involved in germline differentiation. Furthermore, it is possible that in addition to somatic repression of IVS3 splicing, splicing of IVS3 could be activated specifically in the germline.

A

TTGTGGAGCATTTTTTTGGCAGCATGCGATCGAGAGGTGGACA

ATTCGACCATCCCACTCCACTGCAGTTTAAGTATAGGTTAAG

5′SS ➡

AAAATATATAATAGGTATGACAAATTTAAAAGAATGCGTAA

ACAAAAATGTAATTCCATGATTTATAATTGTTTAATGTTTAG

CTATATGTTTCAGGAAAGTTTCAGTTGAGAATGTAGGTAGTT

ATGTGCTGTCTATTGTGTTTTGTCTTTTATCTGTTTCTTTTCA

BP ⬇ 3′SS ⬅

TTTTATTATTTAATCATTATCCTTTTGCTTATCCAGCCAGGA

ATACAGAAATGTTAAGAAATTCGGGAAATATCGAAGAGGAC

B SEQUENCE OF 5′ SPLICE SITES

Splice Site		Sequence	Match	Location
		C A		
consensus		AAG/GUGAGU		
P IVS 3		UAG/GUAUGA	(6/9)	nt 1945 -1953
fushi tarazu		CAG/GUAGGC	(7/9)	
P IVS3 pseudo 5′ splice sites				
upstream	1	UAG/GUUAAG	(5/9)	nt 1925 - 1933
	2	GAG/GUGGAC	(5/9)	nt 1882 - 1890
downstream	1	AAU/GUAAUU	(7/9)	nt 1980 - 1988
	2	AAU/GUAGGU	(7/9)	nt 2046 - 2054
	3	UAG/GUAGUU	(6/9)	nt 2050 - 2058
	4	AUU/GUGUUU	(6/9)	nt 2070 - 2078

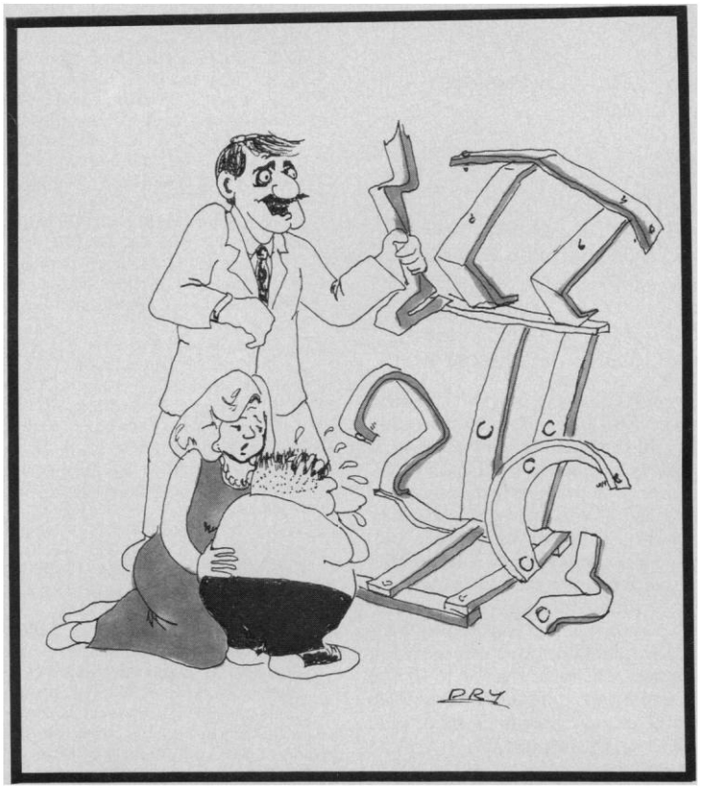
Fig. 8. Location of the pseudo–5′ splice sites within and near the P element third intron. (**A**) Sequence of the P element DNA between nucleotides 1849 and 2184 (*18*). The positions of the splice sites and branch point are denoted by arrows. The accurate 5′ splice site and the six pseudo–5′ splice sites are underlined. Pseudo–5′ splice sites were defined as containing the absolutely conserved GT dinucleotide at positions 4 and 5 of the consensus sequence and an overall match of at least 5 nt with the 9-nt consensus sequence (*50*). The IVS3 5′ splice sites (six pseudo sites, one accurate site) are found within 196 nt of each other, and the shortest distance between the accurate site and a pseudo site is 20 nt. Computer analysis of 10,000 random sequences of 20 nt (taking into account the average 40 percent GC content of *Drosophila* DNA) revealed the presence of approximately 200 (2 percent) sequences that contained two or more pseudo–5′ splice sites. The same analysis of 10,000 random sequences of 196 nt revealed the presence of approximately 1400 (14 percent) sequences that contained seven or more pseudo–5′ splice sites. (**B**) Sequences of the consensus, IVS3, *fushi tarazu* (*ftz*) (*51*) and IVS3 pseudo–5′ splice sites. The overall match with the consensus sequence (*50*) as well as the positions of the IVS3 5′ splice sites within the P element sequence (*18*) are also listed.

REFERENCES AND NOTES

1. P. Sharp, *Science* **235**, 766 (1987); R. A. Padgett, P. J. Grabowski, M. M. Konarska, S. Seiler, P. Sharp, *Annu. Rev. Biochem.* **55**, 1119 (1986); M. R. Green, *Annu. Rev. Genet.* **20**, 671 (1986); A. R. Krainer and T. Maniatis, in *Frontiers in Molecular Biology: Transcription and Splicing*, B. D. Hames and D. M. Glover, Eds. (IRL Press, Oxford, 1988), p. 131; G. Dreyfuss, L. Philipson, I. W. Mattaj, *J. Cell Biol.* **106**, 1419 (1988).
2. J. A. Steitz *et al.*, in *The Structure and Function of Small Nuclear Ribonucleoprotein Particles*, M. L. Birnstiel, Ed. (Springer-Verlag, Berlin, 1988), p. 115; T. Maniatis and R. Reed, *Nature* **325**, 673 (1987); C. Guthrie and B. Patterson, *Annu. Rev. Genet.* **22**, 387 (1988).
3. S. E. Leff, M. G. Rosenfeld, R. M. Evans, *Annu. Rev. Biochem.* **55**, 1091 (1986); R. E. Breitbart, A. Andreadis, B. Nadal-Ginard, *ibid.* **56**, 467 (1987); A. Andreadis, M. E. Gallego, B. Nadal-Ginard, *Annu. Rev. Cell Biol.* **3**, 207 (1987).
4. B. S. Baker, *Nature* **340**, 521 (1989); J. Hodgkin, *Cell* **56**, 905 (1989); P. M. Bingham, T.-B. Chou, I. Mims, Z. Zachar, *Trends Genet.* **4**, 134 (1988).
5. I. W. Mattaj, *Cell* **57**, 1 (1989); R. J. Bandziulis, M. S. Swanson, G. Dreyfuss, *Genes Dev.* **3**, 431 (1989).
6. T.-B. Chou, Z. Zachar, P. M. Bingham, *EMBO J.* **6**, 4095 (1987); Z. Zachar, T.-B. Chou, P. M. Bingham, *ibid.*, p. 4105.
7. F. A. Laski, D. C. Rio, G. M. Rubin, *Cell* **44**, 7 (1986).
8. H. M. Robertson and W. R. Engels, *Genetics* **123**, 815 (1989); S. Misra and D. C. Rio, *Cell*, in press.
9. D. C. Rio, F. A. Laski, G. M. Rubin, *Cell* **44**, 21 (1986).
10. F. A. Laski and G. M. Rubin, *Genes Dev.* **3**, 720 (1989).
11. D. C. Rio, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2904 (1988).
12. C. W. Siebel and D. C. Rio, unpublished results.
13. R. Reed and T. Maniatis, *Cell* **41**, 95 (1985).
14. B. Ruskin and M. R. Green, *Science* **229**, 135 (1985).
15. E. B. Keller and W. A. Noon, *Nucleic Acids Res.* **13**, 497 (1985).
16. ———, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7417 (1984); R. Reed and T. Maniatis, *Genes Dev.* **2**, 1268 (1988); B. Ruskin, J. M. Greene, M. R. Green, *Cell* **41**, 833 (1985).
17. C. W. J. Smith and B. Nadal-Ginard, *Cell* **56**, 749 (1989); D. Helfman and W. M. Ricci, *Nucleic Acids Res.* **17**, 5633 (1989).
18. K. O'Hare and G. M. Rubin, *Cell* **34**, 25 (1983).
19. B. Ruskin, A. R. Krainer, T. Maniatis, M. R. Green, *ibid.* **38**, 317 (1984); H. Domdey *et al.*, *ibid.* **39**, 611 (1984); J. R. Rodriguez, C. W. Pikielny, M. Rosbash, *ibid.*, p. 603.
20. A. Debec, *Roux's Arch. Dev. Biol.* **174**, 1 (1974); G. Echalier and A. Ohanessian, *C. R. Acad. Sci. (Paris)* **D268**, 1171 (1969).
21. P. J. Grabowski, R. A. Padgett, P. A. Sharp, *Cell* **37**, 415 (1984).
22. B. A. Sasnowski, J. M. Belote, M. McKeown, *ibid.* **58**, 449 (1989).
23. B. S. Baker, personal communication.
24. W. R. Engels, in *Mobile DNA*, D. E. Berg and M. M. Howe, Eds. (American Society for Microbiology, Washington, DC, 1989), p. 437; S. B. Daniels, K. R. Peterson, L. D. Strausbaugh, M. G. Kidwell, A. Chovnick, *Genetics* **124**, 339 (1990).
25. H. J. P. Shoemaker and P. R. Schimmel, *J. Mol. Biol.* **84**, 503 (1974); J. T. Wilusz and T. Shenk, *Cell* **52**, 221 (1988); M. A. Garcia-Blanco, S. F. Jamison, P. A. Sharp, *Genes Dev.* **3**, 1874 (1989).
26. G. Dreyfuss, *Annu. Rev. Cell. Biol.* **2**, 459 (1986); J. Wilusz, D. I. Feig, T. Shenk, *Mol. Cell. Biol.* **8**, 4477 (1988).
27. Competition experiments were performed by adding excess unlabeled P element IVS3, adenovirus, β-globin, or *ftz* pre-mRNA to a binding reaction containing *Drosophila* Kc extract and labeled IVS3 pre-mRNA. These experiments indicated that a molar excess of 50 to 100 times of unlabeled IVS3 pre-mRNA competed for binding of the 97-kD protein. In contrast, a molar excess of 200 to 400 times of the adenovirus or β-globin pre-mRNA's was needed to achieve the same effect, and a molar excess of 800 times of *ftz* pre-mRNA failed to compete for 97-kD protein binding. Thus, the 97-kD protein bound to IVS3 pre-mRNA with the highest affinity, approximately two to four times higher than the affinity with which this protein bound to the adenovirus or β-globin pre-mRNA's and at least eight times

higher than the affinity with which this protein bound to the *ftz* pre-mRNA. These data led to conclusions that were consistent with those reached from the cross-linking studies (Fig. 5).
28. R. N. Nagoshi and B. S. Baker, *Genes Dev.* **4**, 89 (1990).
29. R. Reed and T. Maniatis, *Cell* **46**, 681 (1986).
30. H. J. Mardon, G. Sebastio, F. E. Baralle, *Nucleic Acids Res.* **15**, 7725 (1987); R. K. Hampson, L. LaFollette, F. M. Rottman, *Mol. Cell. Biol.* **9**, 1604 (1989); M. Streuli and H. Saito, *EMBO J.* **8**, 787 (1989).
31. S. M. Mount, I. Petterson, M. Hinterberger, A. Karmas, J. A. Steitz, *Cell* **33**, 509 (1983); D. L. Black, B. Chabot, J. A. Steitz, *ibid.* **42**, 737 (1985); B. Chabot and J. A. Steitz, *Mol. Cell. Biol.* **7**, 698 (1987); A. Bindereif and M. R. Green, *EMBO J.* **6**, 2415 (1987); V. Heinrichs *et al.*, *Science* **247**, 69 (1990).
32. A. Mayeda, K. Tatei, H. Kitayama, K. Takemura, Y. Oshima, *Nucleic Acids Res.* **14**, 3045 (1986); M. L. Zapp and S. M. Berget, *ibid.* **17**, 2655 (1989); A. Krämer and W. Keller, *EMBO J.* **4**, 3571 (1985); A. Krainer and T. Maniatis, *Cell* **42**, 725 (1985); A. Krainer, personal communication.
33. B. Seraphin and M. Rosbash, *Cell* **59**, 349 (1989); M. Zillman, S. D. Rose, S. M. Berget, *Mol. Cell Biol.* **7**, 2877 (1987); M. Zillman, M. L. Zapp, S. M. Berget, *ibid.* **8**, 814 (1989); A. Bindereif and M. R. Green, *EMBO J.* **6**, 2415 (1987).
34. A. Krämer, *Genes Dev.* **2**, 1155 (1988).
35. B. Ruskin, P. D. Zamore, M. R. Green, *Cell* **52**, 207 (1988).
36. A. Krämer, *J. Mol. Biol.* **196**, 559 (1987).
37. A. I. Lamond, M. M. Konarska, P. A. Sharp, *Genes Dev.* **1**, 532 (1987).
38. S. W. Ruby and J. Abelson, *Science* **242**, 1028 (1988); B. Chabot and J. A. Steitz, *Mol. Cell. Biol.* **7**, 281 (1987); I. W. Mattaj, W. J. Habets, W. J. vanVenrooj, *EMBO J.* **5**, 997 (1986).
39. P. Legrain and M. Rosbash, *Cell* **57**, 573 (1989); D. A. Chang and P. A. Sharp, *ibid.* **59**, 789 (1989).
40. K. K. Nelson and M. R. Green, *Genes Dev.* **2**, 319 (1988).
41. J. D. Dignam, R. M. Lebovitz, R. G. Roeder, *Nucleic Acids Res.* **11**, 1475 (1983).
42. The branched species migrates as a doublet due to either incomplete RNase T1 digestion at the 5′ end of the intron or because there may be two different adenosine residues to which the 5′ guanosine of the intron is joined. The consensus sequence for a branch point in *Drosophila* introns is $\frac{C}{T}$-T-$\frac{A}{G}$-**A**-$\frac{C}{T}$ with the branch occurring at a conserved adenosine (A) residue (boldface) 18 to 35 nt upstream from the 3′ splice site (*15*). The A residue at position 2114 within the IVS3 sequence contained within this 38-nt RNase T1 fragment meets these criteria (Fig. 8A) (*18*).
43. K. A. Jones, K. R. Yamamoto, R. Tijan, *Cell* **42**, 559 (1985).
44. DNA sequencing reactions used as markers indicated that the 87-nt primer extension product mapped the branch point to a uridine residue, one nucleotide downstream from the consensus branch point adenosine residue at position 2114 (Fig. 8A) (*18*). This observation most likely reflects the fact that reverse transcriptase is sterically blocked by the RNA branch and is, therefore, unable to add nucleotides immediately adjacent to the branch point. Given that such observations have been made (*19*), uridine residues are rarely used as branch points (*1, 16*) and that this uridine residue does not fall within the *Drosophila* branch point consensus sequence (*15*), we infer that the branch point is the adenosine residue at position 2114 (*18*).
45. P. A. Krieg and D. A. Melton, *Methods Enzymol.* **155**, 396 (1987); J. K. Yisraeli and D. A. Melton, *ibid.* **180**, 42 (1989).
46. D. Solnick, *Cell* **42**, 157 (1985); pE2Sma is a modified version of pSPAd in which a Sma I linker has been inserted into the Rsa I site in the adenovirus L1-2 second exon.
47. The Δ11 template was made with Bal 31 exonuclease to delete 15 nt, including the IVS3 5′ splice site, from the 3′ end of the Pst I–Dra I fragment (nucleotides 1911 to 1958) of P element DNA (*18*). The resulting 5′ exon fragment, consisting of P element nucleotides 1916 to 1945 (*18*) (5′ splice site at 1947), was cloned into the Pst I–Sma I sites of pGEM2 (Promega). Cleavage of this recombinant plasmid with Eco RI provided the Δ11 transcription template.
48. U. K. Laemmli, *Nature* **270**, 680 (1970).
49. M. M. Konarska, *Methods Enzymol.* **180**, 442 (1989).
50. S. M. Mount, *Nucleic Acids Res.* **10**, 459 (1982).
51. A. Laughon and M. Scott, *Nature* **310**, 25 (1984).
52. We thank D. Black, M. Garcia-Blanco, M. Konarska, A. Sachs, and P. Sharp for helpful suggestions and stimulating discussions; S. Mount, D. Black, and the Sharp laboratory for plasmid DNA's; D. Black, P. Kaufman, P. Kim, S. Misra, B. Ruskin, and A. Sachs for critically reading the manuscript; F. Laski for communicating results prior to publication; T.-M. Yi for calculating 5′ splice site frequencies; A. Hannaford for help with tissue culture cells; and A. Kron for preparation of the manuscript. Supported by the NIH (HD 22587), NSF (DMB8857176), the Lucille P. Markey Charitable Trust, the DuPont Center for Molecular Genetics at the Whitehead Institute, an NSF predoctoral fellowship (C.W.S.), and a Lucille P. Markey Scholar Award (D.C.R.).

19 December 1989; accepted 25 April 1990



" Never mind that. Ask him just* what *'got away '."