The mechanism of somatic inhibition of *Drosophila* P-element pre-mRNA splicing: multiprotein complexes at an exon pseudo-5' splice site control U1 snRNP binding

Christian W. Siebel,1,2 Lucille D. Fresco, and Donald C. Rio1,2

Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142 USA; 1Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

Somatic inhibition restricts splicing of the *Drosophila* P-element third intron (IVS3) to the germ line. We have exploited this simple system to provide a model for a mechanism of alternative pre-mRNA splicing. Biochemical complementation experiments revealed that *Drosophila* somatic extracts inhibited U1 snRNP binding to the 5' splice site. Using sensitive RNase protection and modification-interference assays, we found that U1 snRNP bound to a pseudo-5' splice site in the 5' exon and that multiprotein complexes bound to an adjacent site. Binding of these factors appeared to mediate the inhibitory effect, because mutations in the pseudo-5' splice sites blocked binding and activated splicing in vitro. Likewise, wild-type, but not mutant, 5' exon RNA titrated inhibitory factors away from the pre-mRNA and activated splicing. Thus, we have defined the pseudo-5' splice sites as crucial components of the regulatory element, correlated the inhibitory activity with specific RNA binding factors from *Drosophila* somatic cells, and provided a mechanistic description of somatic inhibition. Because the inhibitory activity involves general splicing functions such as protein recognition of 5' splice site sequences and changes in the distribution of bound U1 snRNP, our data may also provide insights into how splice sites are selected.

[Key Words: Alternative splicing; RNA-binding proteins; U1 snRNP; P elements; 5' splice sites; RNA processing]

Received April 16, 1992; revised version accepted May 11, 1992.

Regulation of RNA splicing, the process that removes intervening sequences (introns) from mRNA precursors (pre-mRNAs) and rejoins coding sequences (exons), controls the expression of many eukaryotic genes, often according to tissue-specific or temporal cues (for reviews, see Smith et al. 1989; Green 1991; Maniatis 1991). Variations in the activities of general splicing factors may control some alternative splicing events. The general splicing factor ASF/SF2 alters 5' splice site selection in vitro and may control alternative splicing of the SV40 T/t antigen pre-mRNA (Ge and Manley 1990; Krainer et al. 1990a,b). The binding of regulatory factors to specific pre-mRNA sequences can also control alternative splicing. For example, the *Sex-lethal* (*Sxl*), *transformer* (*tra*), and *tra-2* proteins appear to bind directly to specific RNA target sequences to regulate a cascade of alternative splicing events in the *Drosophila* pathway of somatic sexual differentiation (for review, see Baker 1989). Although these and other studies provide tools for understanding mechanisms of alternative splicing, these mechanisms have remained largely speculative.

The *Drosophila* P transposable element, which displays a simple alternative splicing pattern and is accessible to genetic and biochemical techniques, provides a valuable system to approach such mechanistic questions. The P-element third intron (IVS3) is spliced only in germ-line cells, restricting the 87-kD transposase protein and, hence, transposition, to the germ line (Laski et al. 1986). In contrast, somatic cell mRNA retains IVS3 and encodes a 66-kD protein that represses transposition (Robertson and Engels 1989, Misra and Rio 1990). Somatic inhibition of IVS3 splicing plays a role in this regulation. *Drosophila* somatic cell extracts, which splice a variety of introns, fail to splice IVS3 and specifically in-
hibit IVS3 splicing in mammalian splicing extracts (Siebel and Rio 1990). Also, mutations in the IVS3 5′ exon, defined as a cis-acting regulatory region in vivo (Laski and Rubin 1989) and in vitro (Siebel and Rio 1990), activate IVS3 splicing in somatic cells of transformed flies (Chain et al. 1991). RNA-binding proteins from Drosophila somatic cells specifically bind to the IVS3 5′ exon, and addition of 5′ exon RNA relieves somatic inhibition of IVS3 splicing in vitro, suggesting that these RNA–protein interactions block IVS3 splicing (Siebel and Rio 1990). Pseudo-5′ splice sites (5′ splice site-like sequences) in the 5′ exon and third intron could play a role in somatic inhibition by binding regulatory proteins and/or competing with the accurate 5′ splice site (Siebel and Rio 1990). However, testing this model and determining how RNA–protein interactions block IVS3 splicing has required further experimentation.

We now describe biochemical experiments that support a mechanism for somatic inhibition in which the binding of U1 small nuclear ribonucleoprotein particle (snRNP) and multiprotein complexes to two exon pseudo-5′ splice sites prevents U1 snRNP binding to the accurate 5′ splice site. Our experiments provide direct support for the 5′ splice site competition model and may have identified proteins involved in the general process of 5′ splice site selection.

Results

Drosophila somatic cell extracts inhibit binding of U1 snRNP to the accurate 5′ splice site

Drosophila somatic extracts specifically inhibit formation of both the products and intermediates of the IVS3 splicing reaction in mammalian extracts, demonstrating that the block to splicing occurs before the first step of the reaction (Siebel and Rio 1990). Before this step, a large RNP complex that catalyzes splicing, called the spliceosome, assembles on pre-mRNA in an ordered pathway. Spliceosomes are comprised of snRNPs and other protein factors that appear to facilitate assembly and splice site recognition (for reviews, see Maniatis and Reed 1987; Guthrie and Patterson 1988; Steitz et al. 1988; Green 1991). Inhibition of IVS3 splicing must occur before or during spliceosome assembly. One of the earliest steps in spliceosome assembly is the binding of U1 snRNP to 5′ splice sites (Ruby and Abelson 1988; Scrarphin and Rosbash 1989; Michaud and Reed 1991). Following our observation that the cis-acting regulatory region in the IVS3 5′ exon contains pseudo-5′ splice sites (Siebel and Rio 1990), we hypothesized that these sites are involved in inhibiting IVS3 splicing and that inhibition occurs at the level of U1 snRNP binding. To test this hypothesis, we analyzed U1 snRNP binding with a ribonuclease (RNase) protection–immunoprecipitation assay that had been used to characterize snRNP/pre-mRNA interactions (Parker and Steitz 1989). Transcripts containing the IVS3 5′ splice site and flanking exon and intron sequences were incubated in Drosophila somatic and mammalian extracts at 0°C without ATP, conditions that allow binding of U1 snRNP but not other snRNPs. The RNA was then digested with RNase T1, and RNA fragments that were protected from digestion by bound U1 snRNP were precipitated with antiserum that recognizes U1 snRNP (see Fig. 1A, legend).

Two prominent fragments of 22 and 14 nucleotides, as well as three other fragments of 11, 12, and 15 nucleotides, were detected in mammalian extracts (Fig. 1A, lane 3). Secondary RNase T1 digestions (data not shown) and use of mutant or truncated RNAs (Fig. 1A, panels II–IV) confirmed that the 22-nucleotide fragment represented the accurate 5′ splice site, whereas the other fragments represented the F1 pseudo-5′ splice site in the 5′ exon (Fig. 1B). Therefore, approximately equal amounts of U1 snRNP bound to the accurate and F1 pseudo-5′ splice sites in mammalian extracts. Interestingly, in Drosophila somatic extracts, little U1 snRNP bound to the accurate 5′ splice site in contrast to the large amount bound to the F1 pseudo-5′ splice site (Fig. 1A, lane 5). Surprisingly, U1 snRNP did not bind to the F2 pseudo-5′ splice site in either extract, although this site is used in cryptic splicing reactions and better matches the consensus 5′ splice site sequence than does the F1 site (Fig. 1A, lanes 2, 4) nor in the absence of nuclear extract (Fig. 1A, lane 6). Also, disruption of the F1 and F2 pseudo-5′ splice sites (Fig. 1B, panels II–IV) abolished U1 snRNP binding to the F1 pseudo-5′ splice site and increased U1 snRNP binding to the accurate 5′ splice site, especially in Drosophila extracts (Fig. 1A, cf. lanes 11 and 5 and lanes 9 and 3). These results suggest that binding of U1 snRNP and/or protein factors to these pseudo-5′ splice sites might prevent U1 snRNP from binding to the accurate 5′ splice site, a hypothesis supported by the protein binding and splicing data presented below.

The ratio of U1 snRNP bound at the accurate 5′ splice site to that bound at the F1 pseudo-5′ splice site appeared very low in Drosophila extracts relative to mammalian extracts. Therefore, either a mammalian factor that is absent in an active form in Drosophila extracts promotes U1 snRNP binding to the accurate 5′ splice site, or a Drosophila factor prevents U1 snRNP binding to this site. To distinguish between these possibilities, we performed extract mixing experiments. The same wild-type IVS3 RNA fragment (Fig. 1B, panel I) was preincubated with various amounts of Drosophila somatic extract followed by the addition of a constant amount of mammalian extract. Increasing amounts of Drosophila extract progressively and dramatically decreased U1 snRNP binding to the accurate 5′ splice site (Fig. 2, lanes 2–5). This effect was specific because U1 snRNP binding to the F1 pseudo-5′ splice site decreased only slightly (Fig. 2, lanes 2–5). Thus, components in the Drosophila somatic extract specifically inhibit or alter binding of U1 snRNP to the accurate 5′ splice site. Although the effect was less pronounced than when Drosophila extract was included in the first incubation, Drosophila extract
Siebel et al.

Figure 1. U1 snRNP binding to the IVS3 5' exon in mammalian and Drosophila somatic extracts. (A) U1 snRNP binding to IVS3 5' exon RNA fragments was assayed by using RNase T1 protection and immunoprecipitation. As determined by secondary RNase digestions (data not shown), the predominant fragment of 14 nucleotides (nt), as well as the 11-, 12-, and 15-nucleotide fragments, represent the F1 pseudo-5' splice site, and the 22-nucleotide fragment represents the accurate 5' splice site. Binding to the F2 pseudo-5' splice site, which would be expected to yield an 11- or 16-nucleotide fragment, was not observed. The four panels (I-IV) denote parallel experiments performed with different 5' exon fragments, as indicated below each panel. The + and − signs indicate wild-type and mutant 5' splice sites, respectively. The antisera and extracts used are listed above each panel. (HeLa) Mammalian extract; (Kc) Drosophila somatic extract; (buffer) no extract (control); (NHS) normal human serum (control); [c,-(U 1) RNP] antiserum recognizing U1 and U2 snRNP (see Materials and methods). Recognition of U2 snRNP by this antiserum does not affect our results because [1] binding was performed at 0°C without ATP, conditions that allow U1 snRNP binding but not U2 snRNP binding, and [2] similar results were obtained with antisera that specifically recognizes U1 snRNP (data not shown). We routinely used the α-U1/U2 serum because it precipitated a greater amount of the RNA fragments than did the U1 snRNP-specific antisera. Binding at 20°C or 30°C and the addition of ATP also did not change our results (data not shown). Tenfold more of the Drosophila extract and buffer D reactions were examined relative to the mammalian reactions to detect the low level of accurate 5' splice site binding in the Drosophila extract. The differences in the mammalian and Drosophila extracts were reproducible and independent of the batch of extract. Lanes 1–12 and 13–24 are from different portions of the same gel representing results from a single experiment. (B) Diagrams of the 5' exon RNAs. The immunoselected RNA fragments protected from digestion owing to U1 snRNP binding are marked with arrows below each map. The sequences of the accurate and pseudo-5' splice sites, and the restriction enzyme sites used to generate transcription templates, are listed above each map.

added after preincubating the RNA in mammalian extract also reduced U1 snRNP binding to the accurate 5' splice site [Fig. 2, lanes 7–10]. Therefore, a Drosophila factor may destabilize U1 snRNP bound to the accurate 5' splice site or prevent repeated binding of U1 snRNP to this site.
Likewise, wild-type competitor RNA, but labeled wild-type RNA for formation of complexes b–d (lanes 2, 11). No complexes were formed when the exon pseudo-5' splice sites were disrupted (Fig. 3A, lane 9; Fig. 3B, -5' SS) formed only an extremely low amount of complexes comigrating with complexes a–d (Fig. 3A, lanes 8, 6). Formation of these complexes depended on the 5' splice site sequence because disruption of this sequence (Fig. 3B, −5' SS) allowed only a low level of complex a to form (Fig. 3A, lanes 7, 5). Thus, the complexes that form on the IVS3 5' exon relative to the 5' splice site consensus oligonucleotide argues that complex formation involves recognition of the specific sequences of the F1 and F2 pseudo-5' splice sites in general, however, the large amount of complexes that formed on the IVS3 5' exon relative to the 5' splice site consensus oligonucleotide in conditions that inhibited splicing of a control intron, and 3) inclusion of anti-(U1) RNP antiserum (data not shown).

These RNA–protein interactions could reflect binding to the specific F1 and F2 pseudo-5' splice site sequences or general recognition of any 5' splice site-like sequence. An RNA oligonucleotide containing a consensus 5' splice site (Fig. 3B, +5' SS) formed only an extremely low amount of complexes comigrating with complexes a–d (Fig. 3A, lanes 8, 6). Formation of these complexes depended on the 5' splice site sequence because disruption of this sequence (Fig. 3B, −5' SS) allowed only a low level of complex a to form (Fig. 3A, lanes 7, 5). Thus, the complexes that form on the IVS3 5' exon may include proteins that recognize 5' splice sites in general; however, the large amount of complexes that formed on the IVS3 5' exon relative to the 5' splice site consensus oligonucleotide argues that complex formation involves recognition of the specific sequences of the F1 and F2 pseudo-5' splice sites or their context within the IVS3 5' exon, not just their homology to 5' splice sites.

To precisely define the nucleotides required for RNA–protein complex formation, we performed chemical modification/interference binding experiments using UA11 + RNA (Fig. 4B, nucleotide sequence listed at left). This RNA is identical to the UA11 5' exon RNA described above, except that it contains an additional 9 nucleotides of 5' exon sequence to improve resolution of the terminal RNA sequences involved in protein binding. Following chemical modification and incubation in Drosophila extract, UA11 + RNA formed four RNA–protein complexes, which we renamed 1–4 to denote the use of the longer RNA (Fig. 4A, lane 2 and legend). Free RNA

**Figure 2.** Drosophila somatic extract inhibits U1 snRNP binding to the accurate 5' splice site: extract mixing experiments. U1 snRNP binding to an IVS3 5' exon RNA (Fig. 1B, panel I) was assayed as in Fig. 1A, except that the mRNA was first incubated in various amounts (µl, indicated above each lane) of Drosophila somatic extract and then challenged with one amount (20 µl) of mammalian extract [A, lanes 2–5]. In B (lanes 7–10), the order of extract addition was reversed. Lanes 6 and 11 show U1 snRNP binding in Drosophila extract alone. Lanes 1 and 12 display a ladder of size markers. Equivalent amounts of the Drosophila and mammalian reactions were loaded (unlike the experiment in Fig. 1A, where 10-fold more of the mammalian extract was loaded). The reaction volume was doubled relative to that used in Fig. 1A to accommodate the increased amount of extract in this experiment. For this reason and because a darker exposure of the autoradiograph is shown for this experiment, the overall signal appears stronger relative to that in Fig. 1A. However, the ratio of fragments representing the F1 pseudo-5' splice site to those representing the accurate 5' splice site was similar in both (and other) experiments and independent of the batch of extract (data not shown).

**Multiprotein complexes recognize an exon pseudo-5' splice site**

The factors preventing U1 snRNP binding to the 5' splice site might include the proteins that we previously detected binding specifically to the IVS3 5' exon (Siebel and Rio 1990). To address this question, we used a variety of methods to further characterize the specific 5' exon–protein interactions. First, we used native gel electrophoresis, under conditions that preclude U1 snRNP binding (see below), to study protein complexes assembled on the IVS3 5' exon. In Drosophila somatic extract, a 35-nucleotide 5' exon RNA fragment [Fig. 3B, WT] was completely converted from free RNA [Fig. 3A, lanes 1, 10] to the RNA–protein complexes labeled a–d [Fig. 3A, lanes 2, 11]. No complexes were formed when the exon pseudo-5' splice sites were disrupted [Fig. 3A, lane 9; Fig. 3B, −F1,2]. Likewise, wild-type competitor RNA, but not mutant RNA, competed efficiently with the radiolabeled wild-type RNA for formation of complexes b–d (Fig. 3A, lanes 11–18). Therefore, these complexes represent specific interactions between Drosophila proteins and the F1 and/or F2 pseudo-5' splice sites. The mutant RNA competed weakly for formation of complexes c and d, and competition correlated with an increase in complex b (Fig. 3A, lane 18); thus, given that complexes c and d may represent a more complicated form of complex b containing additional proteins (see below), the mutant RNA may compete weakly for binding of one of these proteins. The wild-type and mutant RNAs both competed for formation of complex a (Fig. 3A, lanes 11–18), indicating that the mutations may not significantly affect the binding of proteins in this complex. However, complex a does not represent nonspecific binding because this complex failed to form on the labeled mutant RNA (Fig. 3A, lane 9). In mammalian extract, protein complexes that migrated similarly and depended on the pseudo-5' splice sites also formed (Fig. 3A, lanes 3, 4), although we have not characterized these complexes further. The Drosophila protein–RNA interactions were independent of U1 snRNP because all of the binding reactions included heparin, a polyanion thought to disrupt U1 snRNP binding (Bindereif and Green 1987; Konarska 1989). Furthermore, complex formation was not affected by [1] micrococcal nuclease digestion of nucleic acids in the extract, [2] cleavage of U1 snRNA using an oligonucleotide complementary to the 5' end of U1 snRNA and RNase H under conditions that inhibited splicing of a control intron, and [3] inclusion of anti-(U1) RNP antiserum (data not shown).

These RNA–protein interactions could reflect binding to the specific F1 and F2 pseudo-5' splice site sequences or general recognition of any 5' splice site-like sequence. An RNA oligonucleotide containing a consensus 5' splice site (Fig. 3B, +5' SS) formed only an extremely low amount of complexes comigrating with complexes a–d (Fig. 3A, lanes 8, 6). Formation of these complexes depended on the 5' splice site sequence because disruption of this sequence (Fig. 3B, −5' SS) allowed only a low level of complex a to form (Fig. 3A, lanes 7, 5). Thus, the complexes that form on the IVS3 5' exon may include proteins that recognize 5' splice sites in general; however, the large amount of complexes that formed on the IVS3 5' exon relative to the 5' splice site consensus oligonucleotide argues that complex formation involves recognition of the specific sequences of the F1 and F2 pseudo-5' splice sites or their context within the IVS3 5' exon, not just their homology to 5' splice sites.

To precisely define the nucleotides required for RNA–protein complex formation, we performed chemical modification/interference binding experiments using UA11 + RNA (Fig. 4B, nucleotide sequence listed at left). This RNA is identical to the UA11 5' exon RNA described above, except that it contains an additional 9 nucleotides of 5' exon sequence to improve resolution of the terminal RNA sequences involved in protein binding. Following chemical modification and incubation in Drosophila extract, UA11 + RNA formed four RNA–protein complexes, which we renamed 1–4 to denote the use of the longer RNA (Fig. 4A, lane 2 and legend). Free RNA
Siebel et al.

Figure 3. Protein complexes specifically bound to the IVS3 5’ exon. (A) Native gel electrophoresis. Radiolabeled RNA oligonucleotides (probe) were incubated in the absence (−) [lanes 1, 10] or presence (+) of Drosophila somatic extract [lanes 2, 5–9, 11–18] or mammalian extract [H] [lanes 3, 4]. Native gel electrophoresis resolved free RNA from the RNA–protein complexes [Drosophila complexes labeled a, b, c, and d]. The indicated molar excess of unlabeled competitor RNA [WT comp., UA11; mutant comp., UA11–F1,F2] was included in lanes 12–14 and 16–18. Lanes 1–4 and 5–18 represent results from two experiments. Lanes 5 and 6 represent a darker exposure of lanes 7 and 8. (B) Sequence of the RNA oligonucleotides. [UA11WT] A 29-nucleotide IVS3 5’ exon RNA fragment that ends 2 nucleotides upstream from the accurate 5’ splice site and contains an additional 6 nucleotides of T7 promoter sequence at its 5’ end. [UA11–F1,2] A mutant version of UA11WT carrying the mutations indicated with arrows. (+ 5’ SS) A 25-nucleotide RNA containing a 5’ splice site consensus sequence; (− 5’ SS) a mutant version of the + 5’ SS RNA lacking any homology to the 5’ splice site consensus sequence except for the GT dinucleotide.

Analysis of complexes c and d that formed on UA11 RNA yielded similar results (data not shown). Formation of complexes 1 and 2 required only a subset of the nucleotides crucial for formation of complexes 3 and 4; and even at these positions, chemical modification often only weakly affected complexes 1 and 2 [Fig. 4B, lanes 7, 8, 13, 14]. We used 5’-end-labeled RNA with nearly identical results, except that modification of 2 nucleotides upstream of the pseudo-5’ splice sites also appeared to disrupt complex formation [Fig. 4C]. In summary, the modification–interference experiments demonstrate that Drosophila somatic factors (excluding U1 snRNP) recognize the F2 pseudo-5’ splice site and thus form specific RNA–protein complexes on the IVS3 5’ exon.

To characterize the protein components of these complexes, we cross-linked the RNA–protein complexes in the native gel with ultraviolet light and detected the polypeptides following electrophoresis in an SDS–poly-
acrylamide gel. We obtained similar results using a number of 5' exon RNAs, including UΔ11 and UΔ11 +. Complexes 3 and 4 [UΔ11 + RNA] and complexes c and d [UΔ11 RNA], which each required the F2 pseudo-5' splice site for binding [as defined by the modification-interference experiments; Fig. 4B,C], contained 65- and 50-kD cross-linked proteins and perhaps a 40- to 45-kD protein (Fig. 4D, lanes 2,3,6). Because the 50-kD protein was not detected in any other complexes, its presence correlated with the formation of the F2-specific complexes [3,4,c,d; Fig. 4D, lanes 2,3,6], suggesting that this protein binds to the F2 pseudo-5' splice site. In contrast,
the 65- and 45-kD proteins appeared in other complexes [a, b, and 2] that failed to generate a discernible modification-interference pattern [Fig. 4D, lanes 4–10 and data not shown]. Therefore, the 65- and 45-kD proteins may bind less specifically than the 50-kD protein or in a manner that remains unaffected by chemical modification at a single nucleotide. A 97- to 116-kD protein appeared in complex 4; this polypeptide likely represents the 97-kD protein that has been correlated with the somatic inhibitory activity (Siebel and Rio 1990; Chain et al. 1991). In fact, preliminary fractionation of the extract suggests that formation of complex 4 requires the 97-kD protein [C.W. Siebel and D.C. Rio, unpubl.]. The low level of 97-kD protein binding detected in this assay probably reflects the fact that this protein, like some nucleic acid-binding proteins, fails to bind stably to RNA in the presence of heparin or is stripped off of the RNA during native gel electrophoresis (see Fig. 4D, legend).

Mutations in exon pseudo-5’ splice sites activate IVS3 splicing

The RNA-binding data suggest that binding of U1 snRNP and multiple proteins to two exon pseudo-5’ splice sites inhibits IVS3 splicing. Because mutations in these sites disrupted binding, we predicted that these mutations would also activate IVS3 splicing. To test this prediction and the 5’ splice site competition model, we examined the splicing efficiency of IVS3 substrates carrying mutations in three regions: (1) upstream of the accurate 5’ splice site, (2) downstream of the accurate 5’ splice site, and (3) at the accurate 5’ splice site [Fig. 5A].

In Drosophila somatic extracts, mutations in the upstream pseudo-5’ splice sites activated both steps of IVS3 splicing. For example, splicing of mutant IVS3 premRNA [Fig. 5A, F1, F2-GT] generated a significantly greater amount of 5’ exon intermediate than did splicing of wild-type premRNA under three different splicing conditions [Fig. 5B, cf. E1 in lanes 2–4 and 6–8]. Fewer nucleotide changes or changes in either pseudo-5’ splice site alone also activated splicing to a similar level [C.W. Siebel and D.C. Rio, unpubl.]. Although Drosophila extracts generated the accurate 5’ exon intermediate, they also used a number of novel cryptic splice sites and thus produced mRNA and lariats different from those observed following accurate splicing in mammalian extracts [Fig. 5B]. These cryptic sites include (1) a 3’ splice site [position 2017] upstream from the accurate 3’ splice site [position 2138] and (2) the D2 pseudo-5’ splice site. This splicing pattern in vitro seemed to reflect events observed in vivo, because mutations in the IVS3 5’ exon also activate cryptic splicing at both of these sites in Drosophila somatic cells [Chain et al. 1991]. Furthermore, conclusions from our studies in vitro appear relevant to regulated IVS3 splicing in vivo because mutations in the exon pseudo-5’ splice sites, but not the downstream sites, activate accurate IVS3 splicing in transfected Drosophila somatic tissue culture cells [P. Seshaiha, C.W. Siebel, and D.C. Rio, unpubl.].

The same mutations in the upstream pseudo-5’ splice

Figure 4. Characterization of the RNA–protein complexes. [A] The four RNA–protein complexes analyzed by modification-interference. Native gel electrophoresis resolved four RNA–protein complexes that formed on an IVS3 5’ exon RNA (UAI1 +) in the presence [lane 2] but not absence [lane 1] of Drosophila somatic extract. UAI1 + RNA is identical to UAI1 RNA [Fig. 3B], except that it contains an additional 9 nucleotides of 5’ exon sequence to improve resolution of terminal RNA sequences following modification-interference. In the modification-interference and two-dimensional gel assays, complexes 3 and 4 [UAI1 + RNA] and complexes a and d [UAI1 RNA] yielded similar results as did complex 2 [UAI1 + RNA] and complexes a and b [UAI1 RNA] (see below and data not shown). Complex 1 appeared only at high ratios of RNA to protein, a condition required to generate enough RNA–protein complex for modification-interference analysis. [B] Modification-interference assay with 3’-end-labeled 5’ exon RNA. 3’-End-labeled UAI1 + RNA that had been treated with DEPC to modify purines, lanes 3–8) or hydrazine (to modify pyrimidines, lanes 9–14) was incubated in Drosophila somatic extract, and the RNA–protein complexes were resolved as in A. RNA was either eluted from the indicated RNA–protein complexes [lanes 5–8, 11–14] or from the free remaining after incubation in the absence [lanes 4, 10] or presence [lanes 3, 9] of nuclear extract. Denaturing gel electrophoresis resolved the products generated by aniline cleavage at the modified positions. Unmodified RNA eluted from a complex [lane 2] served as a control for nonspecific degradation of the RNA. RNA ladders prepared by alkaline hydrolysis of 3’ [lane 1] or 5’ [lane 15] end-labeled RNA served as size markers. The DNA sequence of the 5’ exon, with the positions of the pseudo-5’ splice sites marked, is listed at left. Arrows and asterisks mark the nucleotides where chemical modification interfered with formation of the RNA–protein complexes strongly or weakly, respectively. The nucleotides comprising the F2 pseudo-5’ splice site and 3–5 nucleotides upstream were critical for assembling the protein complexes. A potential consensus-binding sequence (YAGNTTAAG) that is repeated in tandem coincides with these nucleotides. C is the same as B, except that 5’ end-labeled RNA was used. [D] Two-dimensional gel assay to detect RNA-binding proteins in the RNA–protein complexes. RNA-binding proteins from the indicated complexes were radiolabeled in the native gel by UV cross-linking and transfer of label from the cross-linked 97-kD protein appears to be spread between 97 and 116 kD. Mod-Int Pattern? denotes whether chemical modification of the bases comprising the F2 pseudo-5’ splice site interfered with formation of the indicated complexes [see B and C]. Y Yes, N no, na not applicable, nd not determined. Lanes 1–4 and 5–10 represent the results from two gels. The positions of molecular mass markers [in kD] are indicated between the gels. Preliminary fractionation experiments support the results from this assay, except that the 65-kD protein does not appear to be required for assembly of complexes 3 or 4. Thus, the 65-kD protein detected in complexes 3 and 4 may result from nonspecific binding or from contamination of these complexes with complex 2, a complex containing a strong 65-kD cross-linking signal.
Mechanism regulating P-element splicing

Figure 5. [See following page for legend.]
sites also dramatically activated both steps of IVS3 splicing in mammalian extracts, which splice wild-type IVS3 pre-mRNAs inefficiently (Fig. 5C, cf. lanes 4 and 8). Therefore, some components of the inhibitory apparatus may be evolutionarily conserved, consistent with our observation that inhibition, in part, involves conserved general splicing factors such as U1 snRNP and that specific RNA–protein complexes assembled on the exon pseudo-5' splice sites in mammalian and Drosophila extracts. Mutations in the downstream pseudo-5' splice sites combined with those in the upstream sites (Fig. 5A, -all) did not activate splicing to a greater extent than did mutations in the upstream sites alone, suggesting that the downstream sites play little role in inhibiting IVS3 splicing (Fig. 5C, cf. lanes 4 and 6); disruption of only the downstream sites failed to activate splicing (C.W. Siebel and D.C. Rio, unpubl.). Thus, the upstream F1 and F2 pseudo-5' splice sites, but not the downstream sites, act as regulatory sequences, perhaps by competing with the accurate 5' splice site. This competition requires juxtaposition of the competing sites, because separating the sites by an additional 40 nucleotides, but not 15 nucleotides, activated splicing (C.W. Siebel and D.C. Rio, unpubl.). In further support of the competition model and previous genetic results (Laski and Rubin 1989), improvement of the accurate 5' splice site [from a 6 of 9 to an 8 of 9 match to the consensus sequence] increased U1 snRNP binding ~20-fold and activated splicing ~5- to 10-fold [L.D. Fresco and D.C. Rio, unpubl.]. Disruption of the accurate 5' splice site [Fig. 5A, -5'SS] also activated splicing, except that splicing occurred from the F2 pseudo-5' splice site (Fig. 5C, cf. lanes 2, 4, and 8). All of these mutations yielded similar results in Drosophila and mammalian extracts, except that (1) Drosophila extracts yielded the cryptic splicing products described above [Fig. 5B], and (2) disruption of the accurate 5' splice site had no effect in Drosophila extracts other than to inhibit splicing from this site (data not shown).

IVS3 5' exon RNA activates IVS3 splicing

If factor binding to the exon pseudo-5' splice sites inhibits IVS3 splicing, then IVS3 5' exon RNA might activate IVS3 splicing by titrating these factors away from the splicing substrate. Consistent with this prediction, IVS3 5' exon RNA relieves the inhibition of IVS3 splicing in mammalian extracts caused by Drosophila extracts (Siebel and Rio 1990). To examine this effect further, we analyzed the ability of a number of RNAs to activate

![Figure 5. Activated splicing of mutant IVS3 pre-mRNAs.](image-url)
IIVS3 splicing in mammalian and Drosophila somatic extracts alone. A 500-fold molar excess of 5' exon RNA competitor, called Δ11, dramatically increased splicing efficiency in mammalian extracts [Fig. 6A, cf. lanes 1 and 2], consistent with published results [Tseng et al. 1991]. Larger amounts of Δ11 RNA failed to activate splicing further [Fig. 6A, lane 3], and at a 25,000-fold molar excess, splicing was inhibited, probably because the competitor titrated necessary splicing factors away from the pre-mRNA [data not shown]. Addition of Δ11 RNA had little, if any, effect on the splicing efficiency of other Drosophila and mammalian introns (a total of five other introns was tested), demonstrating the specificity of the activating effect for IVS3 pre-mRNA [data not shown]. This activating effect was also specific for IVS3 5' exon competitor RNAs. The PL1 RNA fragment, which contains a 5' splice site-like sequence, failed to activate splicing [Fig. 6A, lanes 6,7]. Likewise, RNA fragments containing the adenovirus L1-L2 5' splice site or the 5' splice site consensus sequence (Fig. 3B) failed to activate splicing [data not shown]. Consistent with the protein-binding experiments, these results suggest that binding of regulatory factors to the IVS3 5' exon reflects recognition of the specific F1 and F2 pseudo-5' splice site sequences, not the general recognition of any 5' splice site sequence. Thus, Δ11 RNA may activate IVS3 splicing by

Figure 6. Activated splicing of wild-type IVS3 pre-mRNA in the presence of IVS3 5' exon RNA. (A) Wild-type IVS3 pre-mRNA was spliced in mammalian extract in the absence [0, lane 1] or presence [lanes 2-7] of the indicated molar excess of competitor RNAs. [Δ11 WT] 5' Exon RNA, [Δ11 Mut] 5' exon RNA carrying the -F1,F2-GT mutation [Fig. 5A], [PL1] a control RNA containing a 5' splice site-like sequence. Asterisks mark the positions of lariats generated by cryptic splicing at the F2 pseudo-5' splice site. Competition with IVS3 5' exon RNA activated the accurate and F2 cryptic splicing reactions that are normally observed in mammalian extracts, but it did not change the ratio of accurate to cryptic splicing. (B) Competition experiments were performed as in A with Drosophila somatic extracts. A mammalian splicing reaction was included [H, lane 1] to mark the positions of accurately spliced products, as listed at left. The Drosophila reactions generated the products listed at right; asterisks mark the positions of novel lariats with a new branchpoint, as seen in lanes 3 and 4. To reveal these lariats, we used a 10% acrylamide gel. Thus, the normal branched products also migrated at positions different from those in Fig. 5B, where an 8% acrylamide gel was used. Δ11 RNA also weakly affected branchpoint selection in mammalian extracts. For example, debranching experiments revealed that the two novel branched species that appeared slightly above the accurate intron lariat and pre-mRNA [see A, lane 3] represented faster migrating forms of the accurate intron-exon and intron lariats [C.W. Siebel and D.C. Rio, unpubl.]; thus, high concentrations of Δ11 RNA appeared to weakly shift the branchpoint from position 2114 usually employed in mammalian extracts to a branchpoint farther upstream. We have not determined whether this effect represents the direct titration of factors involved in branch formation or the disruption of the usual ratio of various splicing activities. To better resolve the different splicing products generated by mammalian and Drosophila extracts, pre-mRNAs with 3' exons of different lengths were used in A and B [see Materials and methods]. The abbreviations of the products and the size markers are as in Fig. 5.
Siebel et al.

titrating regulatory factors, as opposed to general 5’ splice site recognition factors such as U1 snRNP. The same mutations in the exon pseudo-5’ splice sites that prevented protein binding also strongly inhibited the ability of the 5’ exon RNA to activate splicing (Fig. 6A, lanes 4,5), further supporting the functional importance of RNA–protein interactions at the pseudo-5’ splice sites. The residual activating effect of the mutant Δ11 competitor RNA and the fact that activation seems to involve titration of more than just U1 snRNP suggest that sequences other than the pseudo-5’ splice sites are also involved in regulating IVS3 splicing; these sequences might include the nucleotides upstream from the pseudo-5’ splice sites that were identified in the modification–interference experiments (Fig. 4B,C).

The 5’ exon RNA also activated IVS3 splicing in Drosophila extracts (Fig. 6B, lanes 2–4). A 500-fold molar excess of Δ11 RNA strongly increased the amount of 5’ exon that was produced (Fig. 6B, cf. E1 in lanes 2 and 3). Splicing in the Drosophila extracts yielded a number of cryptic products, as noted above (Fig. 5B); therefore, the lariat RNAs typical of the mammalian reaction (Fig. 6B, lane 1) were not observed. The lariats that appeared in the Drosophila extracts at a 1500-fold molar excess of Δ11 RNA (Fig. 6B, lane 4) were analyzed by gel purification, debranching, and ribonuclease T1 digestion. They apparently represent the products of the usual cryptic splicing events observed in Drosophila extracts (Fig. 5B), except that Δ11 RNA caused a shift in the branchpoint (see Fig. 6B, legend). As in mammalian extracts, very high concentrations of 5’ exon RNA inhibited splicing (Fig. 6B, lanes 5–7).

Discussion

Biochemical complementation experiments provide two fundamental conclusions about the mechanism regulating alternative splicing of the Drosophila P-element third intron (IVS3). As we reported previously, Drosophila somatic extracts specifically inhibit IVS3 splicing in mammalian extracts, suggesting that regulation is the result of inhibition of IVS3 splicing in somatic cells (Siebel and Rio 1990). We have now demonstrated that Drosophila somatic extracts significantly reduced U1 snRNP binding to the accurate IVS3 5’ splice site, suggesting that somatic inhibition occurs at this step. RNase protection assays revealed that rather than binding to the accurate 5’ splice site, U1 snRNP bound to an inactive pseudo-5’ splice site (F1) in the 5’ exon. Multiprotein complexes, detected by native gel electrophoresis, appeared to restrict U1 snRNP binding to this site. As revealed by modification–interference experiments, these complexes specifically recognized another pseudo-5’ splice site (F2) adjacent to the F1 site. Mutations in the F1 and F2 sites activated IVS3 splicing, demonstrating that these sites function as negative regulatory sequences that compete with the accurate 5’ splice site. Similar mutations also prevented binding of U1 snRNP and the multiprotein complexes, providing a strong correlation between the binding of these factors and the somatic inhibitory activity. Taken together, these experiments lead to a model for the mechanism inhibiting IVS3 splicing in which a complex of proteins prevents recognition of the accurate 5’ splice site by misdirecting U1 snRNP to a nonfunctional pseudo-5’ splice site.

Aspects of this mechanism may be evolutionarily conserved. Mammalian extracts splice IVS3 inefficiently (Siebel and Rio 1990), and 5’ exon RNA competitor or mutations in the pseudo-5’ splice sites that activated IVS3 splicing efficiency in Drosophila extracts also did so in mammalian extracts. The juxtaposition of 5’ splice sites also inhibits splicing of a mammalian β-globin intron (Nelson and Green 1988, I.C. Eperon, pers. comm.) and may limit IVS3 splicing efficiency in both species. Of the factors that inhibit IVS3 splicing, one (U1 snRNP), and perhaps others, are conserved general splicing factors. For example, similar complexes assembled specifically on the IVS3 5’ exon in mammalian and Drosophila extracts, and competition with IVS3 5’ exon RNA activated IVS3 splicing in both extracts. Because other RNAs containing 5’ splice sites (and likely bound by U1 snRNP) failed to activate splicing in this assay, activation probably reflects titration of conserved factors other than U1 snRNP.

On the other hand, our U1 snRNP-binding studies suggest that the mechanisms inhibiting IVS3 splicing are not identical in mammals and Drosophila. In mammalian extracts, U1 snRNP bound to the F1 pseudo-5’ splice site and the accurate 5’ splice site. These results resemble previous experiments where U1 snRNP binding to two juxtaposed 5’ splice sites in a β-globin intron inhibits its splicing (Nelson and Green 1988, I.C. Eperon, pers. comm.). In contrast, Drosophila extracts promoted efficient U1 snRNP binding only to the F1 pseudo-5’ splice site. Furthermore, the addition of Drosophila extract specifically prevented U1 snRNP binding to the accurate 5’ splice site in mammalian extracts. Mammalian extracts could lack Drosophila-specific inhibitory factor[s], or species-specific differences in the concentrations of general splicing factors that influence U1 snRNP binding could inhibit IVS3 splicing. These explanations are not mutually exclusive. For example, Drosophila-specific factor[s] could act with general splicing factors, including U1 snRNP, perhaps by stabilizing their interactions with the IVS3 5’ exon and enhancing the effects of general 5’ splice site competition. The binding of Drosophila-specific factor[s] to the F2 pseudo-5’ splice site may explain why splicing from this site never occurs in Drosophila extracts but can occur in mammalian extracts.

In Drosophila extracts, U1 snRNP bound efficiently to only one of the upstream pseudo-5’ splice sites (F1), even though this site weakly matches the 5’ splice site consensus sequence (Fig. 5) and is not used as a 5’ splice site. Although the F2 pseudo-5’ splice site and the accurate 5’ splice site both better match the consensus sequence and are within 20 nucleotides of the F1 site, these sites were bound poorly, if at all, by U1 snRNP. Given previous demonstrations of the promiscuity of U1 snRNP binding to various 5’ splice site sequences (Chabot and Steitz 1987; Nelson and Green 1988), this specificity is strik-
The proteins that we detected binding to the 5' exon may direct or stabilize U1 snRNP binding to the F1 site or block U1 snRNP binding to the F2 and accurate sites. Modification–interference experiments defined the F2 site as crucial for protein binding, suggesting that the multiprotein complexes sterically occlude recognition of at least the F2 site. Others have noted a lack of a correlation between U1 snRNP binding and splicing [Chabot and Steitz 1987; Nelson and Green 1988], perhaps because functional U1 snRNP binding requires additional factors [Mayeda et al. 1986; Zapp and Berget 1989; Newman and Norman 1991]. RNA context effects, attributed to surrounding sequence and location, influence 5' splice site selection [Reed and Maniatis 1986; Nelson and Green 1988, 1990]. Perhaps the proteins bound at the F2 pseudo-5' splice site play a role in defining this context. Thus, the characterization and purification of these proteins should not only help to clarify the mechanism of IVS3 splicing inhibition but may also further our understanding of how functional 5' splice sites are recognized.

Using a two-dimensional gel assay, we detected a number of RNA-binding proteins in the 5' exon RNA-protein complexes. A 50-kD RNA-binding protein correlated with the appearance of the specific modification–interference pattern at the F2 pseudo-5' splice site. Thus, the 50-kD protein may directly contact the nucleotides that comprise this site or it may affect the binding affinity or specificity of other proteins, directing them to the F2 site [Fig. 7]. Other components of the RNA–protein complexes included a 40- to 45-kD protein and a 65-kD protein, previously detected by UV cross-linking [Siebel and Rio 1990]. We detected these proteins in complexes on the IVS3 5' exon that failed to generate a modification–interference pattern and in complexes on other RNAs containing 5' splice site sequences, suggesting that these proteins alone bind with less specificity than the 50-kD protein.

Previously, we implicated a 97-kD RNA-binding protein, based on its specific UV cross-linking to the IVS3 5' exon and pre-mRNA, as a candidate component of a somatic-specific inhibitory activity [Siebel and Rio 1990]. Probably because binding of the 97-kD protein is unstable during treatment of the RNA–protein complexes with heparin or native gel electrophoresis, we now detect only a small amount of the 97-kD protein in the slowest migrating RNA–protein complex. How the 97-kD protein interacts with the IVS3 5' exon is not yet clear, but preliminary fractionation experiments [C.W. Siebel and D.C. Rio, unpubl.] suggest that formation of the slowest migrating complex requires the 97-kD protein. These experiments support our two-dimensional gel analysis and should allow us to elucidate how the 97-kD protein interacts with other components of the complexes to inhibit IVS3 splicing.

The Drosophila proteins that we have characterized could include previously identified RNA-binding proteins. The best candidates are splicing factors involved early in the splicing reaction or proteins influencing 5' splice site selection such as SF2/ASF [Ge and Manley 1990; Kramer et al. 1990a, b] and SC-35 [Fu and Maniatis 1990]. The 52- to 55-kD Drosophila SRp55/B52 protein [Champlin et al. 1991; Roth et al. 1991] is structurally related to and can functionally substitute for SF2/ASF [Mayeda et al. 1992]; however, preliminary experiments using available antibodies and purified protein fractions [Roth et al. 1991] as probes suggest that this protein is not the Drosophila 50-kD RNA-binding protein that we have identified. The heterogenous nuclear RNP [hnRNP] protein family appears to include splicing factors and proteins that influence 5' splice site selection [Choi et al. 1986; Gil et al. 1991; Paterson et al. 1991; Mayeda and Kainer 1992], and some of the proteins that we have detected may belong to this family [E.L. Matunis et al. 1992; M.J. Matunis et al. 1992].

The absence of somatic-specific inhibitory factors could activate IVS3 splicing in the germ line. Alternatively, germ-line-specific activators may be required. Efficient splicing of mutant IVS3 pre-mRNAs in vitro [this paper, Siebel and Rio 1990; Tseng et al. 1991] and in vivo [Chain et al. 1991] does not require such activators. These findings, however, do not exclude the possible requirement for such factors to activate splicing of wild-type IVS3 pre-mRNA. The Drosophila doublesex (dsx) gene has provided a clear example of positive regulation of alternative splicing. The tra and tra-2 proteins, perhaps in a complex with a conserved RNA-binding protein, may stabilize assembly of the U2 snRNP complex at the weak [nonconsensus] dsx polypyrimidine tract [Tian and Maniatis 1992]. Like the P-element regulatory
mechanism that we have proposed, these mechanisms involve interactions between regulatory RNA-binding factors, specific exon RNA sequences, and basic splicing factors. Thus, the specific recognition of pre-mRNA regulatory sequences by multiprotein complexes is a general mechanism that can positively and negatively control alternative splicing.

We can summarize our results in a model for the mechanism of somatic inhibition (Fig. 7). The specific binding of U1 snRNP to the F1 pseudo-5' splice site and multiprotein complexes to the F2 pseudo-5' splice site inhibits IVS3 splicing by preventing stable binding of U1 snRNP to the accurate 5' splice site. The bound factors may sterically block or destabilize U1 snRNP binding to the accurate 5' splice site. The multiprotein complexes likely influence U1 snRNP binding, and their assembly at the F2 site may explain the unusual specificity of U1 snRNP binding only to the weakest [F1] of three adjacent 5' splice site sequences. For example, the protein complexes may stabilize U1 snRNP binding at the F1 site and/or direct U1 snRNP away from the F2 and accurate sites. The complexes contain 97-, 65-, and 40- to 45-kD proteins as well as a 50-kD protein that likely contacts the F2 site directly. Nucleotide contacts upstream from the pseudo-5' splice sites are also required for complex assembly, suggesting that recognition of sequences in addition to 5' splice site sequences is important for somatic inhibition of IVS3 splicing.

This model points to several remaining questions. For example, how do the multiprotein–RNA complexes affect U1 snRNP binding to 5' splice sites? Does a Drosophila somatic-specific factor play a role, or is regulation mediated entirely by controlling the relative levels of general splicing factors? What other, non-P-element splicing events, perhaps involved in maintaining the germ-line/soma dichotomy, exploit the same regulatory mechanism? The further characterization and purification of the components of the RNA–protein complexes, as well as the cloning of the genes that encode them, should help in answering these questions and clarifying the fundamental problem of how splice sites are selected.

Materials and methods

RNase protection/immunoprecipitation assay

We performed this assay essentially as described (Parker and Steitz 1989). Capped IVS3 RNA [500,000 cpm] was added to 250 μg of mammalian [7.3 μg/μl] or Drosophila extract [7.8 μg/μl] or to buffer D (Dignam et al. 1983) alone. The mixture was incubated for 15 min on ice under splicing conditions without ATP and creatine phosphate, conditions that allow binding of U1 snRNP, but not U2 snRNP [Mount et al. 1983; Black et al. 1985; Bindereif and Green 1987]. RNase T1 [750 units, Boehringer Mannheim], which cleaves 3' to unprotected G residues, was then added, and RNA fragments bound by U1 snRNP were immunoprecipitated by use of protein A–Sepharose CL-4B beads [Sigma] and an autoimmune serum against U1 snRNP [α-U1(U2)] RNP patient autoimmune serum [EW, Fresco et al. 1987], a gift of Jack Keene [Duke University Medical Center, Durham, NC]]. Normal human serum served as a control. The immunoprecipitates were washed six times in NET2 buffer [150 mM NaCl, 50 mM Tris-Cl at pH 7.4, 0.05% NP-40] and resuspended in NET2, 182 mM NaOAc, 9 mM MgCl2, and 33 μg of yeast RNA. The RNAs were phenol extracted, ethanol precipitated, separated on a denaturing 16% polyacrylamide gel, and sized by comparison to RNA ladders produced by alkaline hydrolysis of 5'-end-labeled RNA [Query et al. 1989]. To monitor the digestion, RNA from one one-hundredth of the reaction prior to washing of the immunoprecipitates was analyzed. (Fig. 1, E) Secondary RNase T1 digestion of gel-purified RNA fragments was performed as described (Rio 1988 and data not shown).

To help distinguish RNA fragments containing each of the U1 snRNP-binding sites in the 5' exon, site-directed mutagenesis was used to introduce a guanosine residue between the F2 pseudo-5' splice site and the accurate 5' splice site. A cytosine residue was also introduced to create an EcoRV restriction site that could be used to generate 5' exon RNAs lacking the accurate 5' splice site (Fig. 1B). These two nucleotide changes did not affect IVS3 splicing in Drosophila or mammalian extracts [data not shown]. Mutagenesis was performed as described below using the following oligonucleotides: 5'-CATACCTAAGATA-TCTTTTCTAACCC-3' and 5'-CATACCTAGATATCTTT- TGAATAAG-3'. The PstI–HincII IVS3 DNA fragment was cloned into pGEMZf+ [Promega], and transcription templates were prepared with EcoRV or Drel. The RNAs contained 22 nucleotides of polynucleoty sequence followed by 30 [EcoRV] or 50 [DraI] nucleotides of IVS3 5' exon or 5' exon and intron sequence. RNAs were gel purified following transcription in the absence of unlabeled GTP by use of 75 μCi of [α-32P]GTP [>800 Ci/m mole, Amersham or DuPont/NEN], GpppG cap analog [Pharmacia], and SP6 RNA polymerase [New England Biolabs or Ambion] in a standard 20-μl reaction.

Native gel electrophoresis

For analysis of RNA–protein complex formation on the 5' exon RNA oligonucleotides, the RNAs [29,000–54,000 cpm, 5 × 10^6 cpm/pmol] were incubated for 1 hr at 20°C under the following conditions: 100 μg of total protein from the extract, 25 mM KCl, 10 mg/ml of heparin, 24 mM HEPEs [pH 7.6], 4% [vol/vol] glycerol, 40 μM EDTA, 100 μM DTT, 114 μM PMSE, 4 mM MgCl2, 3 mM ATP, and 5 mM creatine phosphate. Four microliters of each 25-μl reaction was then electrophoresed for ~2 hr and 30 min (until the bromphenol blue dye was 3 cm from the bottom of the gel) at 140 V and 2 W maximum on a 4.2% acrylamide native gel in 50 mM Tris base and 50 mM glycine [Konarska 1989].

Modification–interference assays

Approximately 25 pmoles of the 5' exon fragment (UA11 + ) was labeled with 32P at the 5' end essentially as described (Conway and Wickens 1987) by using calf intestinal phosphatase [Boehringer Mannheim], T4 polynucleotide kinase [U.S. Biochemicals], and [γ-32P]ATP [7000 Ci/m mole] [ICN]. 3'-End-labeling reactions with ~75 pmoles of RNA were performed by use of T4 RNA ligase [U.S. Biochemical] and [5'-32P]CP synthesized essentially as described [England et al. 1980].

One-third of the end-labeled RNA was then chemically modified (approximately zero to one modification per molecule) at purine or pyrimidine residues with DEPC or hydrazine, respectively, essentially as described [Conway and Wickens 1987]. For the DEPC reactions, the RNA pellet in a 1.5-ml microcentrifuge tube was resuspended in 200 μl of 50 mM sodium acetate [pH 4.5], 1 mM EDTA, and chilled for at least 10 min on ice. DEPC [2.3 μl, Sigma] was added, and the tube was vortexed for 5 sec.
and incubated for 135 sec at 90°C. For the hydrazine reactions, the RNA pellet was resuspended in 10 μl of H₂O prior to the addition of 10 μl of hydrazine (Sigma) and an incubation for 6.5 min on ice. The reactions were stopped as described. The final pellet was resuspended in 50 μl of loading buffer (Conway and Wickens 1987) before gel-purifying the RNA.

Binding reactions in Drosophila somatic extract and native gel electrophoresis were performed as described above. Free RNA and the RNA-protein complexes were cut out of the gel. The gel slices were soaked in 30 μl of loading buffer in 1.5-ml microcentrifuge tubes and placed in large wells (1 cm wide) of a denaturing 10% acrylamide gel (16 x 19.7 x 0.15 cm) to purify full-length RNA. The acrylamide slices containing the RNA were chopped into small (1 x 1 mm) pieces, placed into 12 x 75-mm tubes (Falcon), containing 10 μg of tRNA and 800 μl of elution buffer (0.5 mM ammonium acetate, 10 mM magnesium acetate, 1 mM EDTA, 0.1% SDS), and eluted for 20 hr at 4°C shaking at 200 rpm. The RNA was extracted twice with an equal volume of butanol, ethanol precipitated, washed with 95% ethanol, and dried.

The RNA was then cleaved at the modified nucleotides by dissolving the pellet in 20 μl of 1 M aniline (Fluka, redistilled and stored at ~80°C) diluted in 0.3 M sodium acetate (pH 3.8). The sample was vortexed for 1–2 min before a 21-min incubation at 60°C in the dark. The reaction was stopped and treated as described before electrophoresis on a denaturing 16% acrylamide gel.

**Two-dimensional gel assay for RNA-binding proteins**

Following native gel electrophoresis, RNA–protein complexes were cut out of the gel and placed on a glass plate on ice. Ultraviolet (UV) cross-linking reactions were performed for 17 min by placing a hand-held UV monitor [254 nm, UVC-54 (Ultraglo Products, Inc., San Gabriel, CA); 3 mW/cm² on the lamp surface] directly above the gel slice (lamp to gel distance equalled 2.5 mm). For the UAll + RNA–protein complexes, the RNA was digested by soaking the gel slice in 50 μl of a 1:1 (vol/vol) mix of buffer D (Dignam et al. 1983) and 10 mg/ml of RNase A [Pharmacia]. The samples were incubated for 30 min at 37°C with shaking at 300 rpm, the excess buffer D and RNase A were removed, and the incubation was continued for 14 hr at room temperature. For the UA11 + RNA–protein complexes, the RNA was digested in the same manner with 100–150 μl of 2.5 mg/ml of RNase A [Boehringer Mannheim] diluted in native gel buffer. This reaction was incubated for 30 min at 37°C, the excess solution was removed, and the incubation was continued for 20 min at room temperature. The gel slices were soaked in 50 μl of 2x SDS-PAGE sample buffer until they were well permeated; for example, the samples were either left at room temperature for 1–2 hr or agitated at 300 rpm for 1 hr at 37°C. Finally, the samples were boiled for 5 min and loaded onto an 8.5% [UA11 + ] or 7.5% [UA11] acrylamide SDS–polyacrylamide gel [Laemmli 1970].

For the control UV cross-linking reactions in solution, the RNA was incubated under the binding conditions described above for the native gel assay. The reactions were placed as drops on paraffilm on ice and were cross-linked for 10.5 min as described above. Fifteen microliters of 2x SDS-PAGE sample buffer was added, and the samples were boiled for 5 min and loaded onto SDS–polyacrylamide gels.

**Synthesis of RNA substrates and competitors for splicing assays in vitro**

Wild-type substrates were described previously [Siebel and Rio 1990]. Substrates with shortened 3’ exons (Figs. 5B and 6B) were prepared by use of the polymerase chain reaction [PCR] (Scharf 1990). First, synthetic DNA oligonucleotides (made with an Applied Biosystems PCR Mate), corresponding to the top and bottom strands of the IVS3 branchpoint region containing nucleotides 2103–2149 (O’Hare and Rubin 1983), were hybridized together to generate double-stranded DNA containing 5’ single-stranded overhangs corresponding to Xhol and BamHI restriction sites. Second, PCR with primers 1 and 2 was used to generate a copy of IVS3 from the PsiI site at 1911 to a newly created Xhol site at position 2103. This DNA was then digested with Xhol and BamHI [New England Biolabs] and ligated, together with the synthetic DNA described above, into the PsiI and BamHI sites of pGEM2 [Promega]. The final product was sequenced (Sequeneq, U.S. Biochemical). Beginning with the C at position 2095 and ending at the newly introduced BamHI site in the 3’ exon, the sequence read (5’—3’): CTCTAGTGAT-TTTATTATTTAATCATTATCCTTTTGCTTATCCAGCAGGAATACAGAG; bottom, GATTCTCTGTATTTCCTGGCTGGATAGCAAAAGGATATTGGAATTAAATATAACT; primer 1, CTCCTCGACTTTAAAG-TATAG; primer 2, CACTAGAGAAGACAGATAAAAG.

For disruption of the pseudo-5’ splice sites, the EcoRI–HinII fragment containing P-element sequences 1711–2316 [O’Hare and Rubin 1983] was cloned into the EcoRI–BamHI sites of M13mp18 [Messing 1983]. For disruption of the pseudo-5’ splice sites in the 5’ exon fragment [Δ11], the HindIII–EcoRI fragment from pGEM2 P/D Δ11 [Siebel and Rio 1990] was subcloned into the HindIII–EcoRI sites of M13mp19 [Messing 1983]. Mutant DNAs from site-directed mutagenesis reactions [Kunkel 1989] were sequenced (Sequeneq, U.S. Biochemical), and the correct mutants were subcloned into the PsiI–BamHI sites (for splicing substrates) or HindIII–EcoRI sites (for Δ11 mutants) of pGEM2 [Promega]. Mutant sequences are listed in Figure 5A. Pre-mRNAs were prepared as described [Siebel and Rio 1990], by use of [α-32P]UTP (800 Ci/m mole) and T7 RNA polymerase [New England Biolabs or prepared as described previously [Grodberg and Dunn 1988]].

The competitor RNAs PL1, Δ11, and Δ11 Mut were prepared as described [Siebel and Rio 1990]. The amounts of competitor RNAs were measured by comparing the intensities of ethidium bromide staining to a standard amount of tRNA following denaturing gel electrophoresis of serial dilutions of each RNA sample. The other RNA oligonucleotides, UA11, UA11-F1,F2, +5’S, −5’S and UA11+, were all prepared according to Miliqan and Uhlenbeck [1989] by use of [α-32P]UTP (800 Ci/m mole).

**In vitro splicing reactions**

Splicing or debranching reactions in mammalian (HeLa) cell nuclei or cytoplasmic (S-100) extracts [Dignam et al. 1983], respectively, were performed essentially as described [Siebel and Rio 1990] in a total volume of 25 μl. The splicing reactions included (except where noted otherwise) ~60 μg of total protein from the extract, 25 mM HEPES [pH 7.6], 4.8% (vol/vol) glycercol, 80 mM KCl, 48 μM EDTA, 120 μM dithiothreitol (DTT), 137 μM phenylmethylsulfonyl fluoride (PMSF), 2.5% [wt/vol] polyvinyl alcohol, 3 mM ATP, 5 mM creatine phosphate, 4 or 5
mm MgCl₂, 10 units of RNasin (Promega), and ~125,000 cpmpre-mRNA (3 × 10⁶ to 4 × 10⁶ cpmp/mole).

Splicing reactions in Drosophila Kc cell nuclear extracts (Rio 1988) were performed under the same conditions except that 120 µg of total protein was used. Mammalian and Drosophila reactions were incubated for ~95 min at 30°C or 20°C, respectively. The optimal concentration of MgCl₂ with either extract was determined to be ~5 mM (data not shown), although splicing in the Drosophila extracts was often detected up to 10 mm. Products from splicing the large and small [3’ exon truncated] pre-mRNAs were resolved on denaturing 10% acrylamide gels, except in Figure 5B, where an 8% gel was used.

Acknowledgments

We thank S. Misra, D. Black, T.-M. Yi, and P. Kaufman for comments on the manuscript, S. Misra and D. Black for encouragement and helpful criticisms throughout the course of the work, K. Lang for insightful advice on the modification–interference experiments, and A. Hannaford for maintaining tissue culture cells and synthesizing oligonucleotides. This work was supported by National Science Foundation and Ayukawa predoctoral fellowships to C.W.S., by the American Cancer Society postdoctoral fellowship (PF-3505) to L.D.F., and by grants from the Lucille P. Markey Charitable Trust, the National Institutes of Health (R01-HD28063), and the National Science Foundation (DMB-8857176) to D.C.R. D.C.R. is a recipient of the Lucille P. Markey Scholar Award. The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby

References

Baker, B.S. 1989. Sex in flies: The splice of life. Nature 340: 521–524.

Bindereif, A. and M.R. Green. 1987. An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. EMBO J. 6: 2415–2424.

Black, D.L., B. Chabot, and J.A. Steitz. 1985. U2 as well as U1

England, T.E., A.G. Bruce, and O.C. Uhlenbeck. 1980. Specific labeling of 3’ termini of RNA with T4 RNA ligase. Methods Enzymol. 65: 65–74.

Fresco, I.D., M.G. Kurilla, and J.D. Keene. 1987. Rapid inhibition of processing and assembly of small nuclear ribonucleoproteins after infection with vesicular stomatitis virus. Mol. Cell. Biol. 7: 1148–1155.

Fu, X.D. and T. Maniatis. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. Nature 343: 437–441.

Ge, H. and J.L. Manley. 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. Cell 62: 25–34.

Gil, A., P.A. Sharp, S.F. Jamison, and M.A. Garcia-Blanco. 1991. Characterization of cDNAs encoding the polypyrimidine tract-binding protein. Genes & Dev. 5: 1224–1236.

Green, M.R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. 7: 559–599.

Großberg, J. and J.J. Dunn. 1988. ompT encodes the Escherichia coli outer membrane protease that cleaves T7 RNA polymerase during purification. J. Bacteriol. 170: 1245–1253.

Guthrie, C. and B. Patterson. 1988. Spliceosomal snRNAs. Annu. Rev. Genet. 22: 387–419.

Konoarska, M.M. 1989. Analysis of splicing complexes and small nuclear ribonucleoprotein particles by native gel electrophoresis. Methods Enzymol. 180: 442–453.

Kraemer, A.R., G.C. Conway, and D. Kozak. 1990a. The essential pre-mRNA splicing factor SF2 influences 5’ splice site selection by activating proximal sites. Cell 62: 35–42.

Kunkel, T.A. 1989. Oligonucleotide-directed mutagenesis without phenotypic selection. In Current protocols in molecular biology (ed. F.M. Ausebel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl), pp. 8.1.1–8.1.6. Greene/John Wiley, New York.

Lacmii, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685.

Laski, F.A., D.C. Rio, and G.M. Rubin. 1986. Tissue specificity of Drosophila P element transposition is regulated at the level of mRNA splicing. Cell 44: 7–19.

Laski, F.A. and G.M. Rubin. 1989. Analysis of the cis-acting requirements for germ-line-specific splicing of the P-element ORF2-ORF3 intron. Genes & Dev. 3: 720–728.

Maniatis, T. 1991. Mechanisms of alternative pre-mRNA splicing. Science 251: 33–34.

Maniatis, T. and R. Reed. 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. Nature 325: 673–678.

Matunis, E.L., M.J. Matunis, and G. Dreyfuss. 1992. Characterization of the major hnRNP proteins from Drosophila melanogaster. J. Cell Biol. 116: 257–269.

Matunis, M.J., E.L. Matunis, and G. Dreyfuss. 1992. Isolation of hnRNP complexes from Drosophila melanogaster. J. Cell Biol. 116: 245–255.

Mayeda, A. and A.R. Krainer. 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell 68: 365–375.

Mayeda, A., K. Tatei, H. Kitayama, K. Takekura, and Y. Ohshima. 1986. Three distinct activities possibly involved in mRNA splicing in a nuclear fraction lacking U1 and U2 RNA. Nucleic Acids Res. 14: 3045–3057.

Mayeda, A., A.M. Zahler, A.R. Krainer, and M.B. Roth. 1992. Two members of a conserved family of nuclear phosphopro
Mechanism regulating P-element splicing

Drosophila P transposable element third intron in vitro: Somatic repression. Science 248: 1200–1208.

Smith, C.W.J., J.G. Patton, and B. Nadal-Ginard. 1989. Alternative splicing in the control of gene expression. Annu. Rev. Genet. 23: 527–577.

Steitz, J.A., D.L. Black, V. Gerke, K.A. Parker, A. Krämer, D. Frendeway, and W. Keller. 1988. Functions of the abundant U-snRNPs. In The structure and function of small nuclear ribonucleoprotein particles [ed. M.L. Birnstiel], pp. 115–154. Springer-Verlag, Berlin.

Tian, M. and T. Maniatis. 1992. Sequence-specific activation of splicing by tra and tra-2 proteins in vitro. Science 256: 237–240.

Tseng, J.C., S. Zollman, A.C. Chain, and F.A. Laski. 1991. Splicing of the Drosophila P element ORF2-ORF3 intron is inhibited in a human cell extract. Mech. Dev. 35: 65–72.

Zapp, M.L. and S.M. Berget. 1989. Evidence for nuclear factors involved in recognition of 5′ splice sites. Nucleic Acids Res. 17: 2655–2674.
The mechanism of somatic inhibition of Drosophila P-element pre-mRNA splicing: multiprotein complexes at an exon pseudo-5' splice site control U1 snRNP binding.

C W Siebel, L D Fresco and D C Rio

*Genes Dev.* 1992, 6: Access the most recent version at doi:10.1101/gad.6.8.1386