MOLECULAR AND CELLULAR BIOLOGY, Dec. 1993, p. 7689–7697
0270-7306/93/127689-09$02.00/0

Vol. 13, No. 12

AU-Rich Intronic Elements Affect Pre-mRNA 5′ Splice Site Selection in *Drosophila melanogaster*

ANDREW J. McCULLOUGH AND MARY A. SCHULER*

Department of Plant Biology, 190 Plant and Animal Biotechnology Laboratory, 1201 West Gregory Drive, University of Illinois, Urbana, Illinois 61801-3838

Received 10 May 1993/Returned for modification 9 July 1993/Accepted 8 September 1993

***cis*-spliced nuclear pre-mRNA introns found in a variety of organisms, including *Tetrahymena thermophila*, *Drosophila melanogaster*, *Caenorhabditis elegans*, and plants, are significantly richer in adenosine and uridine residues than their flanking exons are. The functional significance of this intronic AU richness, however, has been demonstrated only in plant nuclei. In these nuclei, 5′ and 3′ splice sites are selected in part by their positions relative to AU-rich elements spread throughout the length of an intron. Because of this position-dependent selection scheme, a 5′ splice site at the normal (+1) exon-intron boundary having only three contiguous consensus nucleotides can compete effectively with an enhanced exonic site (−57E) having nine consensus nucleotides and outcompete an enhanced site (+106E) embedded within the AU-rich intron. To determine whether transitions from AU-poor exonic sequences to AU-rich intronic sequences influence 5′ splice site selection in other organisms, alleles of the pea *rbcS3A1* intron were expressed in *Drosophila* Schneider 2 cells, and their splicing patterns were compared with those in tobacco nuclei. We demonstrate that this heterologous transcript can be accurately spliced in transfected *Drosophila* nuclei and that a +1 G-to-A knockout mutation at the normal splice site activates the same three cryptic 5′ splice sites as in tobacco. Enhancement of the exonic (−57) and intronic (+106) sites to consensus splice sites indicates that potential splice sites located in the upstream exon or at the 5′ exon-intron boundary are preferred in *Drosophila* cells over those embedded within AU-rich intronic sequences. In contrast to tobacco, in which the activities of two competing 5′ splice sites upstream of the AU-rich intron are modulated by their proximity to the AU transition point, *D. melanogaster* utilizes the upstream site which has a higher proportion of consensus nucleotides. The enhanced version of the cryptic intronic site is efficiently selected in *D. melanogaster* when the normal +1 site is weakened or discrete AU-rich elements upstream of the +106E site are disrupted. Selection of this internal site in tobacco requires more drastic disruption of these motifs. We conclude that 5′ splice site selection in *Drosophila* nuclei is influenced by the intrinsic strengths of competing sites and by the presence of AU-rich intronic elements but to a different extent than in tobacco.**

Most eukaryotic protein-coding genes are interrupted by introns which must be precisely removed from precursor transcripts in order to produce functional mRNAs. The process of intron removal, termed pre-mRNA splicing, proceeds by two sequential transesterification reactions (12, 22). In the first reaction, cleavage at the 5′ exon-intron border occurs with the concomitant formation of a 2′-5′ phosphodiester bond between the first nucleotide of the intron (+1 G) and a branchpoint nucleotide approximately 20 to 50 nucleotides upstream of the 3′ end of the intron. In the second reaction, cleavage at the 3′ intron-exon border releases the intron as a lariat structure and the exons are ligated together.

In mammalian and yeast systems, three *cis*-acting elements are known to be required for intron recognition. One of these elements, the 5′ splice site, is a moderately conserved set of nine nucleotides spanning the 5′ exon-intron border. Within this element, intron position +1 is an invariant guanosine and position +2 is a highly conserved uridine residue. The second critical element is a loosely conserved internal branchpoint sequence that is utilized for formation of the lariat intermediate. The third element, located at the 3′ intron-exon border, consists of a conserved AG dinucleotide which in mammalian introns is preceded by an extended polypyrimidine tract.

Pre-mRNA splicing takes place within a large ribonucleoprotein complex termed a spliceosome. This complex contains five small nuclear RNAs (snRNAs), designated U1, U2, U4, U5, and U6, within four unique small nuclear ribonucleoprotein particles (14, 19, 30). Biochemical and genetic evidence has demonstrated that the 5′ end of U1 snRNA forms complementary base pairings at the 5′ exon-intron border (2, 28, 29, 35) and at the 3′ terminus of the intron (27). U5 snRNA has been implicated in base-pairing recognition of exonic sites adjacent to the 5′ and 3′ splice sites (24, 25, 32), and U2 snRNA has been shown to base pair with sequences surrounding the branchpoint (26, 34, 36).

While the chemistry of pre-RNA splicing is well understood, the mechanism(s) of intron recognition remains unclear. It is increasingly apparent that to ensure proper splice site selection, the splicing machinery must recognize features of introns other than these simple border and branchpoint sequences. Similar splice site motifs, which are not utilized for splicing, exist at multiple locations in most transcription units, and different organisms utilize the consensus motifs to different degrees. *Saccharomyces cerevisiae*, which represents one extreme in a continuum of intron recognition motifs, requires consensus 5′ splice site and branchpoint sequences (33). Mammals represent another point in the continuum in that they require pyrimidine tracts upstream from their 3′ splice sites but have more flexible splice site and branchpoint requirements (12, 22). A large group of other organisms, including plants and invertebrates,

* Corresponding author.

contain no strongly conserved splice site motifs and often no prominent pyrimidine tracts (6).

Considerable experimental evidence supports a role for intronic AU-rich sequence elements in plant intron recognition. Plant introns are 15 to 20% higher in A+U residues than their flanking exons are (11), and this AU richness has been shown to be essential for efficient splicing in dicot nuclei (8, 10). On the basis of our analysis of 5′ and 3′ splice site selection schemes in *Nicotiana benthamiana* (tobacco), we have proposed a model for intron recognition in dicot nuclei suggesting that 5′ and 3′ splice sites are selected in position-dependent manners relative to AU-rich elements spread throughout intronic sequences (17, 18, 21). The experimental support for this model is derived from our demonstration that AU-rich elements upstream from the 3′ end of maize *Adh1* intron 3 are essential for defining the 3′ splice site (17, 18) and that AU-rich sequences strongly modulate 5′ splice site selection in the pea *rbcS3A1* intron (21). Because of these AU-dependent selection schemes, 5′ splice sites at the normal (+1) exon-intron boundary with only limited agreement to the 5′ consensus sequence compete effectively with enhanced exonic sites having absolute identity with the 5′ splice site consensus. Intronic sites which are buried within AU-rich intronic sequences are not utilized to any extent until upstream AU-rich sequences are removed (21).

As in plant systems, the *cis*-spliced nuclear pre-mRNA introns present in a multitude of organisms, including *Tetrahymena thermophila*, *Drosophila melanogaster*, and *Caenorhabditis elegans*, are richer in adenosine and uridine residues than their flanking exons are (6). This, and the absence of strong 5′ and 3′ splice site consensus motifs, suggests that some aspects of the model that we have described for intron recognition may extend to other organisms. Significantly, a recent study has established that the human and *Drosophila* splicing machineries are not necessarily interchangeable and led to the suggestion that small *Drosophila* introns are recognized by a distinct mechanism (13). To determine whether transitions from AU-poor exonic sequences to AU-rich intronic sequences are capable of influencing 5′ splice site selection in *D. melanogaster*, we have expressed heterologous pea *rbcS3A1* constructs in *Drosophila* Schneider 2 (S2) cells and directly compared the spliced products with those obtained in *N. benthamiana*. We demonstrate that in both *Nicotiana* and *Drosophila* nuclei, the activity of a cryptic 5′ splice site located at position +106 within the *rbcS3A1* intron is determined, in part, by the sequence composition upstream from this site. This finding provides the first experimental evidence that AU-rich sequences can affect 5′ splice site selection in *Drosophila* nuclei.

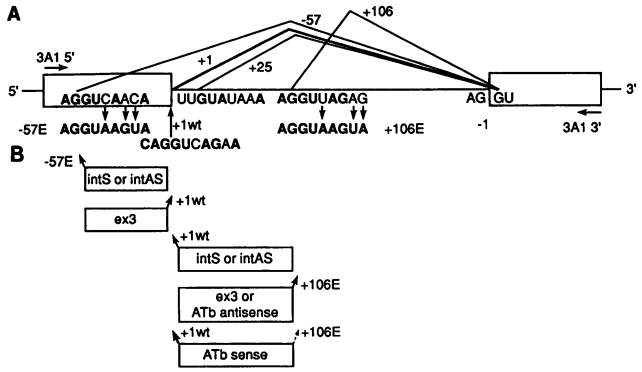

FIG. 1. Summary of pea *rbcS3A1* pre-mRNA mutant constructs and splicing patterns in tobacco nuclei. (A) The *rbcS3A1* pre-mRNA transcript generated in vivo by transcription from the coat protein promoter in tobacco nuclei is drawn to scale. Open boxes represent exons, and a solid line represents the intron. Sequences of the wild-type 5′ splice site at +1 and the cryptic 5′ splice sites at −57, +25, and +106 are indicated, with bold letters representing nucleotides complementary to the 5′ end of U1 snRNA. Sequences of the enhanced 5′ splice sites at −57 and +106 used in this study are shown below each wild-type sequence. Lines connecting the 5′ and 3′ splice sites summarize the splicing patterns obtained in tobacco nuclei. (B) Replacement substitutions between the −57E and +1wt splice sites or between the +1wt and +106E splice sites are designated as intS (sense orientation of β-conglycinin intron 4; Fig. 7), intAS (antisense orientation of β-conglycinin intron 4; Fig. 5), ex3 (antisense orientation of *rbcS3A* exon 3), ATb sense (ATbS; Fig. 7), and ATb antisense (ATbAS; Fig. 5). Solid arrows designate the predominant 5′ splice sites used in tobacco nuclei; the dashed arrow designates the +106E site which is used at low efficiency in the +1 to +106E ATb sense replacement mutant.

MATERIALS AND METHODS

Plant expression constructs. The wild-type and mutant *rbcS3A1* constructs used in this study contain intron 1 and the entire first and second exons of the pea (*Pisum sativum* L.) *rbcS3A* gene (7) inserted into the unique *Bgl*II site of pMON458 (15) as previously described (20). Construction of the splice site mutant and replacement alleles has been described elsewhere (21). The ATb sense mutant was generated from the wild-type +1 (+1wt) to +106E intS replacement construct, using polymerase chain reaction (PCR)-mediated site-directed mutagenesis (31). The ATb antisense construct was generated by inverting the ATb mutant fragment between the +1wt and +106E sites. The nucleotide compositions of the replacement fragments between +1 and +106E are as follows: *rbcS3A1* wild-type +8 to +103 (24% A, 45% U, 12% C, 19% G); intS (29% A, 45% U, 16% C, 10% G); ATbS (26% A, 38% U, 22% C, 14% G); intAS (45% A, 29% U, 10% C, 16% G); ATbAS (38% A, 26% U, 14% C, 22% G); ex3 (33% A, 21% U, 22% C, 24% G); and *rbcS3A1* wild-type −49 to −4 (34% A, 24% U, 18% C, 24% G). Structures of the constructs are shown in Fig. 1.

Recombinant pMON458 constructs were introduced into *N. benthamiana* leaf discs, samples were harvested 4 days after transfection, and total RNA samples were prepared as follows. Tissue (10 to 12 leaf discs) was ground for 1 min with a Mini-Beadbeater (Biospec Products) in an RNA isolation buffer (1 ml) containing 50% phenol-chloroform (1:1), 50 mM LiCl, 50 mM Tris-HCl (pH 8.0), 5 mM EDTA, and 0.5% sodium dodecyl sulfate (SDS). Samples were centrifuged for 5 min at 12,000 × *g* to separate the aqueous and organic phases, after which each aqueous phase was reextracted with an equal volume of phenol-chloroform (1:1). The aqueous phases were made 2 M LiCl by addition of 8 M LiCl, and high-molecular-weight RNA was precipitated on wet ice for 8 to 12 h. RNA was collected by centrifugation at 12,000 × *g* for 15 min. Residual contaminating DNA was removed from RNA preparations by digestion with 5 U of RNase-free DNase (Promega) at 37°C for 60 min.

***Drosophila* expression constructs.** For expression in *Drosophila* S2 cells, each pea *rbcS3A1* allele was excised from pBluescript (Stratagene) as a *Bam*HI-*Sal*I fragment and ligated into a nonreplicating pUC-based vector between the *Drosophila* actin 5C promoter and *Adh* terminator (kindly

provided by L. Cherbas). S2 cells were maintained at 28°C in M3 medium (Sigma) supplemented with 10% fetal calf serum (Sigma). Twelve to sixteen hours prior to transfection, cells were subcultured into 10 ml of fresh medium at a density of approximately 10^8 cells per ml. Cells were transfected by the calcium phosphate method as described previously (3), using 20 μg of plasmid DNA. After 24 h, the medium was replaced with 10 ml of fresh medium, and the cells were collected for RNA isolation 48 h after transfection. RNA was isolated and DNase treated as described for *N. benthamiana* except that the cells were lysed simply by vortexing in 1 ml of RNA isolation buffer. Total RNA from approximately 5×10^9 cells was resuspended in 100 μl of sterile water, and 1 μl of this solution was used for reverse transcription-PCR (RT-PCR) analysis.

Analysis of transcripts. First-strand cDNA synthesis and PCR amplifications were done in a single reaction mixture containing 1 μg of total RNA, 50 mM KCl, 10 mM Tris-Cl (pH 8.4), 2.5 mM $MgCl_2$, 200 μg of gelatin per ml, 200 μM each deoxynucleoside triphosphate, 5 U of avian myeloblastosis virus reverse transcriptase (Promega), 2.5 U of *Taq* DNA polymerase (Bethesda Research Laboratories), 20 U of RNasin (Promega), and 50 pmol of the 3A1 5′ and 3A1 3′ primers (20) complementary to the 5′ and 3′ ends of *rbcS3A* exons 1 and 2, respectively. First-strand cDNA was synthesized for 30 min at 50°C and subsequently amplified by 15 cycles of PCR for plant RNA and 20 cycles of PCR for S2 cell RNA. Each PCR cycle consists of 94°C denaturation for 1 min, 55°C annealing for 2 min, and 72°C extension for 2 min. PCR products were fractionated on 2% agarose gels containing 1× Tris-borate-EDTA buffer, transferred to GeneScreen (Du Pont), and probed with a random-hexamer ^{32}P-labeled *rbcS3A* exon 1-specific DNA probe complementary to the region upstream of −57. Blots were hybridized in 50% formamide–5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–25 mM sodium phosphate buffer (pH 6.5)–0.5% (wt/vol) SDS–5× Denhardt's solution for 16 h at 42°C, and membranes were washed twice in 0.2× SSC–0.1% SDS for 60 min at 68°C. Hybridization signals were quantified with a PhosphorImager (Molecular Dynamics). Coamplification of spliced and unspliced transcripts indicates that this assay quantitatively amplifies precursor and spliced products between 16 and 25 cycles over a wide range of RNA concentrations (see Fig. 3). Each reported splicing efficiency represents the average of at least three independent transfection experiments with the corresponding standard error of the mean (SEM). Splice site selection patterns were defined by cloning PCR products into Bluescript II SK+ (Stratagene), using restriction sites in the PCR primers, and sequencing with T7 DNA polymerase (U.S. Biochemicals) and the 3A1 exon 2 oligonucleotide primer (20).

RESULTS

Splicing of wild-type *rbcS3A1* transcripts in *Drosophila* nuclei. We have previously demonstrated that the pea *rbcS3A1* intron (Fig. 1), which is 469 nucleotides long and 73% AU rich, is accurately and efficiently spliced from precursor transcripts in transfected *N. benthamiana* leaf disc nuclei (Fig. 2A, lane 1) (20, 21). To determine whether this intron could be spliced by the *Drosophila* splicing machinery, the wild-type *rbcS3A1* construct was expressed in S2 cells by using a nonreplicating pUC-based vector which contains a *Drosophila* actin 5C promoter and *Adh* terminator. RNA isolated from transfected S2 cells was reverse transcribed and PCR amplified for 20 cycles, using

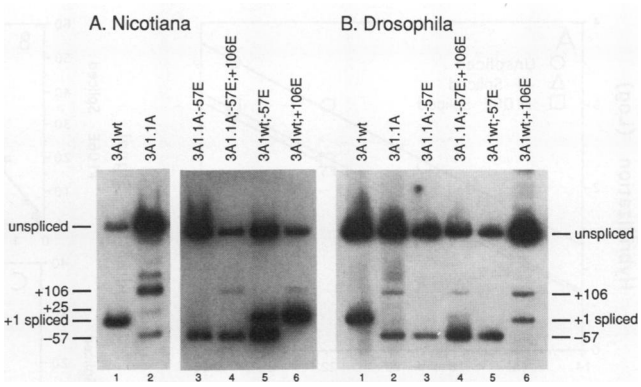

FIG. 2. Splicing of *rbcS3A1* constructs in *Nicotiana* and *Drosophila* nuclei. Constructs containing a wild-type or mutant splice site at +1 and a wild-type or enhanced splice site at −57 or +106 were transfected into *N. benthamiana* leaf discs (A) or *Drosophila* S2 cells (B). RNA from the each transfection was analyzed by 15-cycle (A) or 20-cycle (B) RT-PCR analysis using the 3A1 5′ and 3A1 3′ oligonucleotide primers as described in Materials and Methods. The percentage of accumulated transcript that was spliced was determined by using a ^{32}P-labeled PCR product complementary to *rbcS3A* exon 1 upstream of −57. The construct analyzed is designated above each lane. The positions of the PCR products corresponding to unspliced transcripts, the correctly spliced transcript (+1), and the transcripts spliced between the −57, +25, and +106 cryptic 5′ splice sites and the normal 3′ splice site are shown at the right and left of each panel. In lane 2, the fourth band above the +106 spliced product corresponds to a hybrid formed between +1 and +106 spliced products during the last PCR cycle.

3A1 5′ and 3′ oligonucleotide primers complementary to the 5′ and 3′ ends of *rbcS3A1* exons 1 and 2, respectively (Fig. 1). Two products, 809 and 340 nucleotides, corresponding to precursor and accurately spliced transcripts are generated from the wild-type transcript (Fig. 2B, lane 1). The fidelity of the cleavage and ligation reactions for this and all other constructs described in this report has been monitored by sequencing the cloned PCR product(s) corresponding to each spliced transcript. The efficiencies of splicing at this normal and all cryptic splice sites have been defined by hybridization with a randomly primed exon 1-specific probe. PhosphorImager analysis indicates that this RT-PCR Southern assay quantitatively reflects the levels of precursor and spliced transcripts between 16 and 25 cycles (Fig. 3). The splice site selection patterns in *N. benthamiana* and *D. melanogaster* are summarized in Fig. 4.

Competition between the normal and cryptic 5′ splice sites. Mutation of the +1 G to A in the *rbcS3A1* 5′ splice site abolishes usage of the +1 site and activates three cryptic 5′ splice sites (−57, +25, and +106) in tobacco nuclei (Fig. 2A, lane 2) (21). In *Drosophila* S2 cell nuclei, these cryptic sites at −57, +25, and +106 are activated in the 3A1.1A mutant but at efficiencies different from those observed in tobacco (compare lanes 2 in Fig. 2A and B). The cryptic −57 wild-type splice site is used in preference to the +106 site, which predominates in *N. benthamiana*. The +25 site is used at an efficiency which is barely detectable in *D. melanogaster* compared with *N. benthamiana*.

In *Nicotiana* nuclei, the cryptic 5′ splice site at −57 can be used efficiently and exclusively in the 3A1.1A background when it is enhanced to a consensus sequence which complements U1 snRNA at nine contiguous positions (−57E) (Fig. 2A, lane 3) (21). As shown in Fig. 2B, the 3A1.1A;−57E transcript is also spliced exclusively to the −57E site in

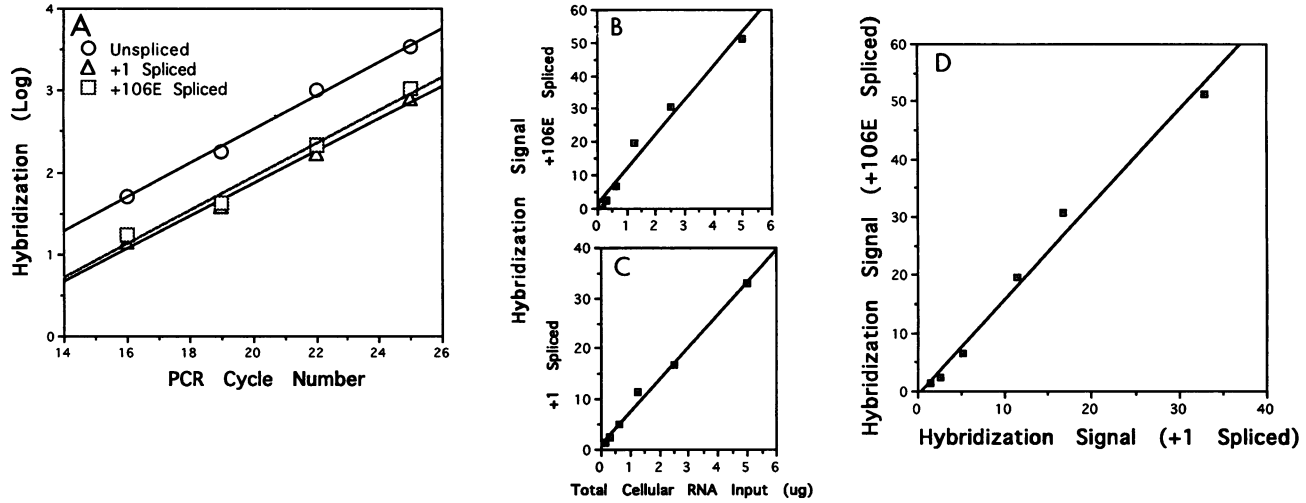

FIG. 3. Calibration of the RT-PCR assay. Total cellular RNA prepared from S2 cells transfected with the *rbcS3A1* wt;+106E construct was analyzed by RT-PCR Southern analysis, using the 3A1 5′ and 3A1 3′ oligonucleotide primers as described in Materials and Methods. The levels of PCR product corresponding to unspliced transcripts and transcripts spliced to the +1 and +106E 5′ splice sites were determined with a PhosphorImager (Molecular Dynamics). (A) Identical RT-PCR mixtures containing 1 μg of S2 cell RNA were collected after 16, 19, 22, and 25 cycles of PCR, quantified, and plotted in actual PhosphorImager units versus PCR cycle number. Least-squares linear regression analysis indicated that the PCR products corresponding to unspliced, +1 spliced, and +106E spliced transcripts amplified exponentially between 16 and 25 cycles of PCR. The overall splicing efficiency for *rbcS3A1* wt;+106E in this experiment is 30%. (B and C) RT-PCR mixtures containing 5.00, 2.50, 1.25, 0.625, 0.312, and 0.157 μg of RNA were analyzed after 20 cycles of PCR amplification. Regression analysis demonstrated linear amplification of the +106E (B) and +1 (C) spliced transcripts. The relative amplification efficiencies of the +1 and +106E PCR products at different RNA concentrations obtained from panels B and C were compared by correlation analysis (D). This analysis demonstrated that these PCR products amplify with the same relative efficiency across the tested range of RNA concentrations. About 1 μg of total RNA was used for the S2 cell splicing analysis reported in this paper.

Drosophila nuclei. In contrast, when the +106 site is enhanced to match the consensus sequence (+106E) and introduced into the 3A1.1A;−57E background, it competes with the −57E site for only 22% of the spliced transcript in *Nicotiana* nuclei and 15% (SEM = 0.88%) in *Drosophila* nuclei (Fig. 2A and B, lanes 4). Thus, *Nicotiana* and *Drosophila* nuclei effectively utilize an enhanced 5′ splice site in the sequence context surrounding the −57 site and not in the sequence context surrounding the +106 site.

In transcripts containing a competing exonic −57E site and the +1wt site, which has seven consensus nucleotides, splicing at the −57E site accounts for about 57% of the splicing in *Nicotiana*, compared with 43% at the +1wt site (Fig. 2A, lane 5) (21). In *Drosophila* nuclei, this 3A1wt;−57E construct is spliced almost exclusively to the −57E site (Fig. 2B, lane 5). In transcripts containing a competing intronic +106E site and a +1wt site, there was 10% usage of the +106E site, compared with 90% usage of the +1wt site in *N. benthamiana* (Fig. 2A, lane 6) (21). In *D. melanogaster*, this intronic +106E site competes more effectively (61%, SEM = 4.9%) with the +1 site. This finding suggests that the *Drosophila* splicing machinery is more dependent than the tobacco splicing machinery on the degree of complementarity to U1 snRNA at each 5′ splice site but that this dependence is not exclusive. The sequence identity of the −57E and +106E splice sites in the region expected to interact with the 5′ end of U1 snRNA compared with the preferential activity of the −57E site indicates that the sequence context and location of these sites play a critical role in 5′ splice site selection.

Effects of sequence context on 5′ splice site selection. While the +106E 5′ splice site is inefficiently utilized in transfected *Nicotiana* nuclei (Fig. 5A, lane 1) (21), it is used almost exclusively when the AU-rich intronic sequences (69% AU) between +1wt and +106E are replaced with AU-poor exonic sequences (ex3, 54% AU) derived from exon 3 of the *rbcS3A* gene (Fig. 5A, lane 2). In contrast, when these sequences are replaced with AU-rich intronic sequences (intS, 74% AU) derived from intron 4 of the soybean β-conglycinin gene, the +1 site is effectively reactivated (Fig. 5A, lane 3) (21). With a +1wt site, this holds true even when the intron sequences are inserted in the antisense orientation (intAS; Fig. 5A, lane 4). In constructs containing a weakened +1 5′ splice site (3A1.−2T), sense and antisense intronic replacements have different effects (Fig. 5A, lanes 5 and 6). Most notably, the antisense intron replacement activates four different 5′ splice sites between −57 and +106E, compared with almost exclusive use of the +1 (−2T) site in the sense replacement.

To determine whether the sequences upstream of the +106E splice site affect its ability to compete with the +1wt site, the +1 to +106E replacement constructs were introduced into *Drosophila* S2 cells. In such a competition, the +106E site predominated when preceded by exonic sequences (ex3; Fig. 5B, lane 2). The +1 site predominated when the +106E site was preceded by sense intron sequences (intS; Fig. 5, lanes 3). This strong activation of the +106E site by the introduction of upstream exonic sequences is identical to the splice site selection patterns obtained in plant nuclei and suggests that, like the plant splicing machinery, the *Drosophila* machinery is capable of recognizing the nucleotide composition of adjacent intron and exon sequences.

Insertion of the intronic sequences in the antisense orientation resulted in a complicated *Drosophila*-specific splicing pattern (Fig. 5B, lane 4). In addition to the splice site choices which occur in tobacco nuclei (intAS; Fig. 5A, lane 4), a

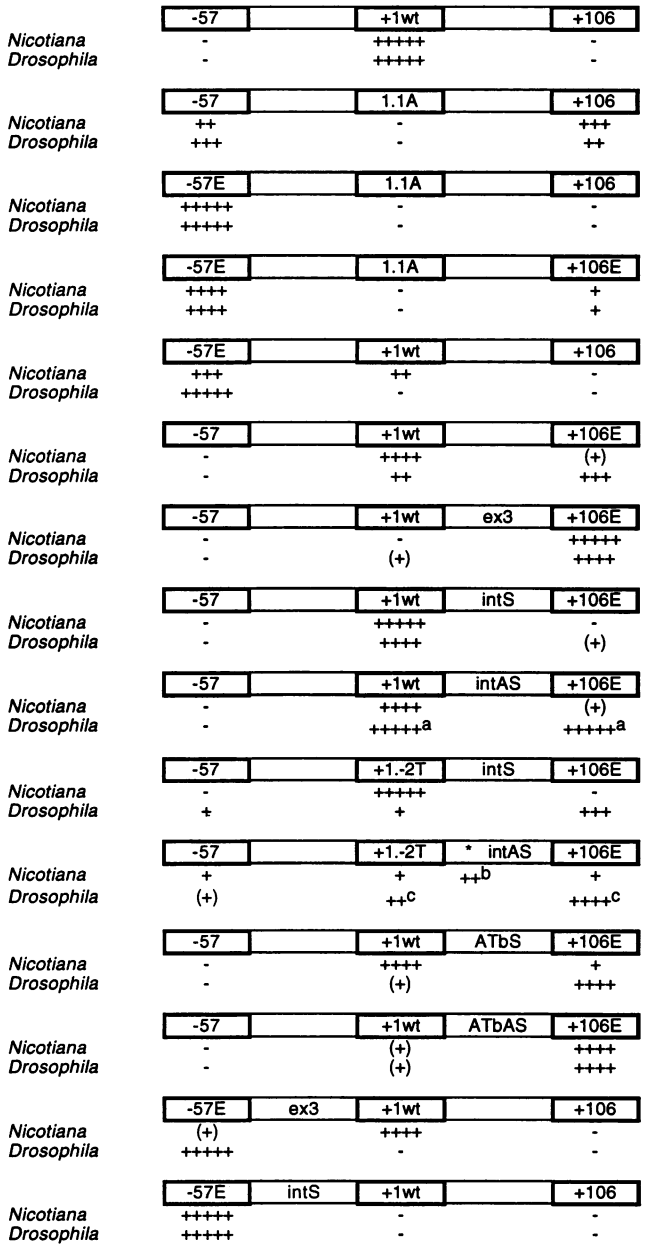

FIG. 4. Efficiencies of splicing at the −57, +1, and +106 5′ splice sites in *Nicotiana* and *Drosophila* cells. (+), low but detectable usage of a site; +, approximately 20% usage; +++++, essentially 100% usage. [a]The most abundant spliced transcripts in *Drosophila* cells are spliced from the +1 5′ splice site to a fortuitous cryptic 3′ splice site (** in Fig. 5) in the intAS sequence and from the +106E 5′ splice site to the normal 3′ splice site at −1 (see Fig. 5). Thus, the +1 and +106E 5′ splice sites in this construct are used with similar efficiencies. [b]The most abundant spliced transcripts in *Nicotiana* cells are spliced between a cryptic 5′ splice site (*) located in the intAS sequence and the normal 3′ splice site at −1 (see Fig. 5). [c]While the most abundant splice in this construct is between the +106E 5′ splice site and the normal 3′ splice site at −1, low levels of the doubly spliced transcripts described in note a do accumulate.

short cryptic intron is spliced from +1 to +106E intAS transcripts produced in *Drosophila* nuclei. The most abundant spliced product derived from this construct has undergone two separate splicing events (Fig. 5C): one in which the +1 5′ splice site has been ligated to a fortuitous 3′ splice site located in the intAS replacement sequence and another in which the +106E 5′ splice site has been ligated to the normal 3′ splice site at −1. This results in the excision of 51- and 363-nucleotide introns and the inclusion of a 52-nucleotide exon derived from the antisense intron sequences. The cryptic 3′ splice site (double asterisk in Fig. 5) occurs directly downstream of a local concentration of AU-rich elements and is preceded by sequences which are similar to the *Drosophila* 3′ splice site consensus (23). These features and the ability of *D. melanogaster* to splice very small introns (13, 23) probably contribute to the functionality of this 3′ splice site in *Drosophila* cells. The 51-nucleotide cryptic intron is substantially shorter than the minimum length spliced in plant nuclei (9).

To estimate the relative contributions of various intronic elements to 5′ splice site selection, constructs containing a weakened +1 site (3A1.−2T) and either the sense (intS) or antisense (intAS) intronic replacement sequence preceding +106E was introduced into *Drosophila* S2 cells. In both cases, the +106E site represented the most favored site. In the sense replacement (intS), the weakened +1 (−2T) and wild-type −57 splice sites were also used to appreciable extents (Fig. 5B, lane 5); in the antisense replacement (intAS), the +106E site was used almost exclusively and very low levels of doubly spliced (see above), +1 spliced, and −57 spliced transcripts were detected (lane 6). This result again indicates that the intrinsic strength of a 5′ splice site and its surrounding context are important determinants in the 5′ splice site selection patterns of *Drosophila* and *Nicotiana* cells but that the dependence on each of these determinants differs in these two organisms. In some contexts (intS), a single nucleotide change in the +1 splice site which has little effect in *Nicotiana* cells dramatically decreases recognition of the +1 site in *Drosophila* cells (Fig. 5, lanes 3 versus 5). In other contexts (intAS), single base changes decrease recognition of the +1 site in both *Drosophila* and *Nicotiana* cells and activate alternate sites (Fig. 5, lanes 4 versus 6). In *Drosophila* but not in *Nicotiana* cells, the fractional usage of these alternate sites is clearly correlated with the number of consensus nucleotides at each site (compare lane 5 or lane 6, Fig. 5A versus Fig. 5B).

To determine whether replacements downstream from an exonic 5′ splice site have similar effects, the ex3 and intS sequences were substituted downstream of the −57E site. While the activities of the +1wt and −57E 5′ splice sites in *Nicotiana* nuclei are strongly affected by these sequence replacements (Fig. 6A, lanes 2 and 3), the −57E site is utilized exclusively in transfected *Drosophila* S2 nuclei, regardless of its downstream sequences (Fig. 6B).

Mutation of an intronic AU-rich element affects splice site selection. Our previous results (21) and those described above demonstrate that *Nicotiana* and *Drosophila* nuclei are capable of distinguishing between exonic and intronic sequence contexts. To further define critical intronic elements, we generated a mutant of the +1wt to +106E intS replacement allele in which U residues present in three central AU-rich islands in the intS sequence were changed to C residues and A residues were changed to G residues as shown in Fig. 7 (designated ATb sense mutant). The inverted orientation of these replacements is designated ATb antisense and should be compared with the intAS replacement.

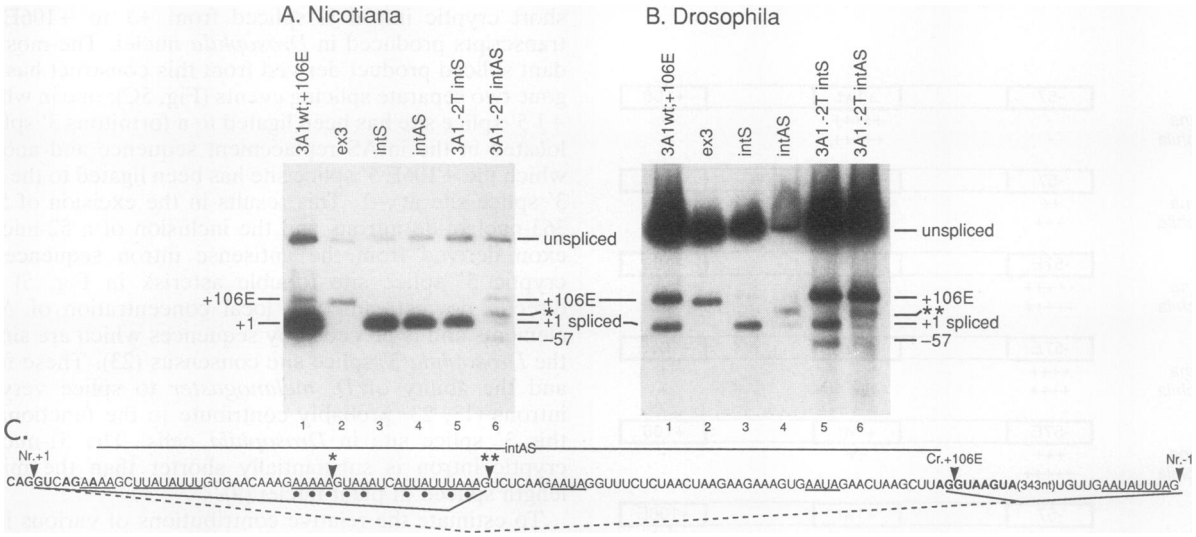

FIG. 5. Splicing of the +1wt to +106E constructs. (A and B) Constructs containing the enhanced +106E cryptic 5′ splice site and either +1wt or mutant (−2T) splice site with the intS (Fig. 7), intAS (C), or ex3 replacement sequences between these two sites were transfected into *N. benthamiana* leaf discs (A) or *Drosophila* S2 cells (B). RNA from each transfection was analyzed by RT-PCR analysis using the 3A1 5′ and 3A1 3′ oligonucleotide primers as described in Materials and Methods. The percentage of accumulated transcripts that was spliced was determined by using a ^{32}P-labeled PCR-amplified probe complementary to *rbcS3A* exon 1 upstream of −57. The construct analyzed is designated above each lane. The positions of the PCR products corresponding to unspliced transcripts, the correctly spliced transcript (+1), and transcripts spliced between the −57 and +106E 5′ splice sites and the normal (−1) 3′ splice site are shown. The single asterisk in panel A designates a PCR product (lane 6) corresponding to a transcript spliced in *Nicotiana* cells between a cryptic 5′ splice site (see panel C) in the intAS intron replacement sequence and the normal (−1) 3′ splice site. The double asterisk in panel B designates a PCR product (lanes 4 and 6) corresponding to a transcript doubly spliced in *Drosophila* cells from the intAS intron replacement as diagrammed in panel C. Because the ex3 replacement sequences slightly decrease the length of its upstream sequences, the molecular weight of the +106E spliced transcript in lane 2 is smaller than that of the wild-type +106E spliced transcripts in lane 1. (C) Sequence of the antisense intron replacement (intAS) between the +1wt site and the +106E enhanced site. AU islands containing four or more contiguous adenosine and uridine residues are underlined. The normal (Nr.+1) and cryptic (Cr.+106E) 5′ splice sites are shown, with consensus nucleotides complementary to U1 snRNA shown in bold. The single asterisk designates the additional cryptic 5′ splice site used for splicing of the 3A1.−2T intAS mutant in *Nicotiana* cells. The splicing patterns of the intAS constructs observed in *Drosophila* cells are summarized below the sequence with the normal (Nr.−1) 3′ splice site and the additional *Drosophila*-specific 3′ splice site (double asterisk). The positions of the A-to-G and U-to-C conversions in the ATb antisense construct are designated with circles positioned below the intAS sequence.

As shown in Fig. 7, the nucleotide substitutions in the ATb sense mutant alter the ability of the intron sequences to direct 5′ splice selection in both *Nicotiana* and *Drosophila* nuclei, though to different degrees. In *Nicotiana* nuclei, the +106E site of the ATb sense mutant was used at approximately 12% efficiency, compared with essentially 0% for the intS parent allele (Fig. 7A, lane 1 versus 2). The shift from the +1 to the +106E 5′ splice site is more dramatic in *Drosophila* nuclei, in which the +106E 5′ splice site was used with 81% (SEM = 1.9%) efficiency in the ATb sense construct (Fig. 7B, lane 2), compared with 11% (SEM = 1.0%) in the parent construct (lane 1). Expression of the ATb antisense allele in both *Nicotiana* and *Drosophila* nuclei results in nearly exclusive usage of the +106E 5′ splice site (Fig. 7A and B, lanes 4). In *Nicotiana* nuclei, this represents an effective switch in the selection of the +1 site to the +106E site; in *Drosophila* nuclei, this represents a switch from selection of the +1 site in singly and doubly spliced transcripts to nearly complete usage of the +106E site. The ATb mutant clearly demonstrates that defined mutations in three internal AU-rich islands affect splice site selection schemes in both *Nicotiana* and *Drosophila* nuclei. The complete switch of the splicing pattern in *Drosophila* nuclei suggests that once this set of elements is disrupted, 5′ splice site selection becomes increasingly dependent on the strength of these two competing sites.

DISCUSSION

Experimental evidence indicates that intronic AU richness is required for efficient splicing (8, 10) and splice site definition (17, 18, 21) in plant nuclei, and a role for AU-rich sequences has been inferred in *C. elegans* (1, 4, 5). Specifically, in vivo *cis*-competition assays in transfected tobacco leaf cells have demonstrated that 5′ splice site selection in the pea *rbcS3A1* intron is dependent on both the number of consensus nucleotides at a potential 5′ splice site and the position of those sequences relative to intronic AU-rich elements (21). In dicot nuclei, 5′ splice sites upstream of the AU-rich intron are favored over those buried within AU-rich sequences. Our initial experiment aimed at extending this functional analysis to *D. melanogaster* demonstrated that the wild-type pea *rbcS3A1* transcript contains all of the elements essential for processing in transfected *Drosophila* S2 cells. This is not surprising in light of the fact that dicot plant and *Drosophila* introns are both AU rich (74% average for dicot introns, 65% for *Drosophila* introns) and contain loosely conserved branchpoint sequences (11, 23). In addition, although the 469-nucleotide *rbcS3A1* intron is significantly longer than the 79-nucleotide median intron length in *D. melanogaster*, it is well within the size range of introns known to be spliced in fly nuclei (23).

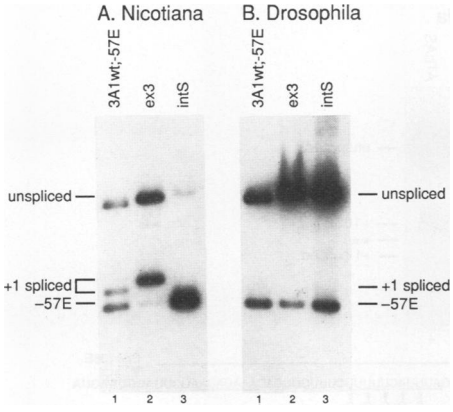

FIG. 6. Splicing of the −57E to +1wt replacement constructs. Constructs containing the +1wt and enhanced +57E cryptic 5′ splice sites with the intS or ex3 replacement sequence between these two sites were transfected into *N. benthamiana* leaf discs (A) or *Drosophila* S2 cells (B). RNA from each transfection was analyzed by RT-PCR analysis using the 3A1 5′ and 3A1 3′ oligonucleotide primers as described in Materials and Methods. The construct analyzed is designated above each lane. The percentage of accumulated transcripts that was spliced was determined by using a ^{32}P-labeled PCR-amplified probe complementary to *rbcS3A* exon 1 upstream of −57. The positions of the PCR products corresponding to unspliced transcripts, the correctly spliced transcript (+1), and the transcript spliced between the −57E 5′ splice site and the normal 3′ splice site (−1) are shown at the right and left. Because the replacement sequences increase the size of exon 1, the molecular weight of the +1 spliced transcripts in lane 2 (panel A) is larger than that of the +1wt spliced transcripts in lane 1.

Masking of intronic 5′ splice sites by AU-rich elements. Our analysis of transcripts containing the +106E 5′ splice site has identified striking similarities between the motifs modulating 5′ splice site recognition in plant and *Drosophila* nuclei. In vivo analysis in tobacco and *Drosophila* nuclei indicates that this downstream 5′ splice site exists within a context recalcitrant to splicing. Neither its wild-type nor its enhanced version functions as efficiently in *cis*-competition assays as does the −57 site. This is most evident in the 3A1.1A;−57E; +106E construct in which the +1 site has been inactivated and both cryptic sites have been enhanced to consensus 5′ splice sites. In this transcript, the +106E site, which is identical to the −57E site, is used at only 22% efficiency in tobacco and 15% efficiency in *D. melanogaster*. In both organisms, the +106E site is strongly activated when AU-poor exonic sequences are substituted for the AU-rich intronic sequences between +1 and +106E and is inactivated when heterologous AU-rich intronic sequences are placed upstream of it. These results suggest that intronic AU-rich elements block usage of downstream cryptic sites even when they contain perfect consensus sequences.

To further define critical elements within the intronic sequences, six U-to-C and three A-to-G transitions were introduced in the intS sense replacement (ATbS mutant). In *D. melanogaster*, these mutant intronic sequences dramatically activate the +106E site to a level comparable to that observed for the exon substitution (ex3) allele. That the ATbS mutant has significantly less effect on selection of the 5′ splice site in tobacco suggests that the AU-rich elements remaining in this mutant intron are sufficient for recognition of the +1wt splice site in plant nuclei. One potential explanation for this disparity is that the *Drosophila* splicing machinery relies more on the strength of competing 5′ splice sites than on the AU richness of adjacent sequences. Disruption of a single set of elements is capable of shifting splicing to the strongest 5′ splice site. The dicot plant splicing machinery, which is less influenced by 5′ splice site strength, requires more extensive disruption of its AU-rich sequences to disturb the balance between sites. Consistent with this, we have recently demonstrated that multiple AU elements must be disrupted to significantly activate the downstream +106E site in tobacco nuclei (21a).

The relative roles of intronic sequence composition and 5′ splice site consensus nucleotides in *D. melanogaster* are also apparent in our analysis of the 3A1.−2T;+106E intS replacement allele. In contrast to tobacco, in which the weak 3A1.−2T site upstream of the intronic sequences is used almost exclusively, the +106E site is preferred in *D. melanogaster*.

A relatively minor fraction of splicing occurs at the weak upstream +1(−2T) site and even weaker −57 splice site. Again, each organism relies to a different degree on AU-rich intronic elements and 5′ splice site strength. The tobacco splicing machinery is more dependent on the position of a site relative to the AU transition point, and the *Drosophila* machinery is more strongly influenced by the strengths of two competing 5′ splice sites.

Usage of cryptic exonic 5′ splice sites. The tobacco and *Drosophila* systems differ significantly in their usage of the cryptic 5′ splice site at −57 in the exon. Firstly, the wild-type −57 site represents the most efficiently used cryptic 5′ splice site in the +1 G-to-A knockout mutant (3A.1A) in *Drosophila melanogaster*, while the wild-type +106 site is favored in tobacco. Secondly, an enhanced −57 site with perfect complementarity to the 5′ end of U1 snRNA (−57E) is used exclusively in competition with the +1wt site in *Drosophila* cells but accounts for only 57% of the splicing activity in tobacco. Finally, this competition is readily manipulated in tobacco nuclei by substituting heterologous AU-rich intronic or AU-poor exonic sequences downstream of the −57E site; the −57E site is used exclusively in *Drosophila* S2 cell nuclei regardless of the downstream sequence composition. These data show a clear preference for the strongest 5′ splice site upstream of the AU-rich intron in *Drosophila* cells and a weaker site closer to the AU transition point in tobacco cells. Interplay between these signal motifs exists in both species, but in *D. melanogaster*, this can readily be overridden by enhancing the number of consensus nucleotides at an upstream 5′ splice site.

Different strengths of sense and antisense intronic elements. The 3A1wt;+106E and 3A1.−2T;+106E alleles in which the intronic replacement sequences between +1 and +106E have been inserted in the antisense orientation (intAS) show different preferences in these two organisms. In tobacco, these antisense constructs are as efficient as sense constructs in promoting splicing at the +1 site when this splice site retains its wild-type sequence but not if it is weakened. The usage of four different 5′ splice sites in the −2T mutant indicates that in the absence of a strong upstream 5′ splice site, the antisense intron sequences do not delineate the 5′ border of the intron as effectively as the sense intron sequences do. The different strengths of sense and antisense intron substitutions are also obvious when the normal *rbcS3A1* intron sequences between +1 and +106E are inverted. In the antisense construct, the +106E site competes more efficiently with the +1 site than in the native sense construct (data not shown). Slight differences in the strengths of the various AU-rich fragments are also apparent

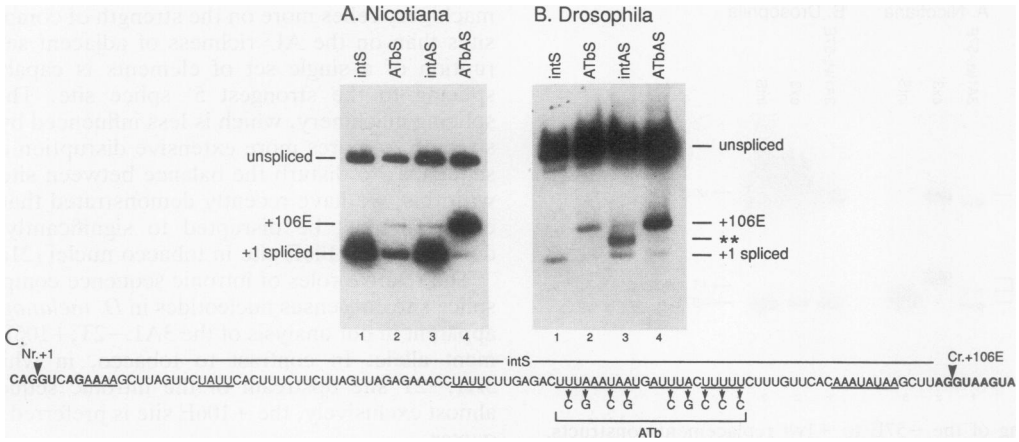

FIG. 7. Splicing of ATb sense and antisense constructs. (A and B) Constructs containing the +1wt and enhanced +106E cryptic 5′ splice sites with the intS or ATbS replacement sequences (C) or the intAS or ATb antisense replacement sequences (Fig. 5) between these two sites were transfected into *N. benthamiana* leaf discs (A) or *Drosophila* S2 cells (B). RNA from each transfection was analyzed by RT-PCR analysis using the 3A1 5′ and 3A1 3′ oligonucleotide primers as described in Materials and Methods. The percentage of accumulated transcripts that was spliced was determined by using a ^{32}P-labeled PCR-amplified probe complementary to *rbcS3A* exon 1 upstream of −57. The construct analyzed is designated above each lane. The positions of the PCR products corresponding to unspliced transcripts, the correctly spliced transcript (+1), and the transcript spliced between the +106E 5′ splice site and the normal (−1) 3′ splice site are shown at the right and left. The double asterisk in panel B designates a PCR product (lane 3) corresponding to the transcript doubly spliced in *Drosophila* cells from the intAS intron replacement as diagrammed in Fig. 5. (C) Sequence of the sense intron replacement (intS) between the +1wt site and the +106E enhanced site. AU islands containing four or more contiguous adenosine and uridine residues are underlined. The normal (Nr.+1) and cryptic (Cr.+106E) 5′ splice sites are shown, with consensus nucleotides complementary to U1 snRNA shown in bold. The A-to-G and U-to-C conversions in the ATb sense construct are designated with arrows.

when intS substitutions are compared with those of the native intron. In both tobacco and *Drosophila* cells, the intS sequences promote splicing at +1 to a greater extent than the wild-type intron sequences do. The different strengths of these intronic sequences may result from the fact that intS sequences are more AU rich (74% AU) than the intron sequences normally present in these positions (69% AU) and thus more effective at redirecting splice site selection patterns. Consistent with this conclusion is the fact that the reduction of AU content at nine additional positions in the intS fragment (to 64% AU) diminishes its activity in tobacco cells and completely eliminates it in *Drosophila* cells. Antisense intron sequences, although identical in AU content to their sense equivalents, have entirely different nucleotide profiles. Thus, the relative strength differences of sense and antisense intron sequences in tobacco nuclei potentially correlate with the fact that the sense intron sequences are U rich and G poor (45% U, 29% A, 16% C, and 10% G for intS) while the antisense intron sequences are neither U rich nor G poor (29% U, 45% A, 10% C, and 16% G for intAS). The marginal activity of the antisense intron substitutions in this A-rich and C-poor intAS background is disrupted more readily than in the alternate U-rich and G-poor intS background.

Activation of a cryptic 3′ splice site in *Drosophila* nuclei. Analysis of the +1 to +106E intAS constructs in *Drosophila* nuclei was complicated by the presence of a cryptic 3′ splice site fortuitously located in the antisense intron replacement sequence. This cryptic 3′ splice site is paired with the +1 5′ splice site, while the +106E 5′ splice site is paired with the normal 3′ splice site at −1 to generate doubly spliced transcripts which account for the majority of accumulated spliced transcripts in the 3A1wt;+106E intAS allele. Additional transcripts correspond to each of the singly spliced transcripts. Although doubly spliced transcripts accumulate at low levels in the 3A1.−2T;+106E and ATb antisense intron constructs, transcripts singly spliced between the +106E 5′ splice site and the normal 3′ splice site at −1 account for most of the spliced product. Thus, mutants which either weaken the 5′ splice site or disrupt sequences immediately preceding the cryptic 3′ splice site reduce excision of the first cryptic intron and effectively redirect splicing to the +106E site.

The usage of the fortuitous 3′ splice site in the intAS replacement in *Drosophila* but not *Nicotiana* nuclei suggests that different intron lengths or 3′ splice site constraints operate in these two types of nuclei. The 51-nucleotide-long intron is the same length as the shortest natural intron known to exist in *D. melanogaster* (23) and below the minimal length (73 nucleotides) capable of being spliced in plant protoplasts (9). Its short length and position immediately downstream from a continuous series of AU-rich elements probably contribute to the functionality of this site in *D. melanogaster*. It is alternatively possible that the presence of a functional 3′ splice site (−1) and/or AU-rich elements downstream from the cryptic 3′ splice site blocks its usage in tobacco nuclei but not *Drosophila* cells. We have already demonstrated that in plant nuclei, cryptic 3′ splice sites within introns are masked by the presence of a functional downstream site (17, 18), but it is not known whether similar mechanisms exist in *Drosophila* cells. These two possibilities are not necessarily mutually exclusive in that an ability to recognize short introns in *Drosophila* cells may position the normal (−1) 3′ splice site beyond the optimal distance for masking.

Conclusions. It is conceivable that the recognition of AU-rich intronic sequences represents a retro-mechanism for defining introns and exons. While this mechanism has remained essential for intron recognition in dicot nuclei, many other organisms have evolved alternative and more highly specialized schemes for intron recognition. The prominence of AU-rich introns and AU transition points at the intron-exon boundaries of *T. thermophila*, *D. melanogaster*, and *C. elegans* introns (6) suggests, however, that remnants of this simple mechanism might remain outside of the plant kingdom. Consistent with this

idea, recent studies have suggested that AU-rich sequences may be important in defining *C. elegans* 5′ and 3′ splice sites (1, 4, 5), and AU-rich transposable elements are spliced from *C. elegans* transcripts by using cryptic splice sites flanking the site of transposon insertion (16). In this report, we demonstrate that 5′ splice site selection in *Drosophila* nuclei is strongly affected by the intrinsic strengths of competing 5′ splice sites and in a significant way by the positions of these sites relative to intronic AU-rich elements. The activity of 5′ splice sites upstream of an AU-rich intron in *Drosophila* cells is correlated with the degree of complementarity to U1 snRNA at each upstream site. In contrast, potential splice sites embedded within AU-rich intron sequences are efficiently selected only after upstream 5′ splice sites and/or upstream AU-rich elements are disrupted. While we do not know the extent to which this mechanism operates in the splicing of natural *Drosophila* introns, our analysis demonstrates that AU-rich intronic elements can be recognized in *D. melanogaster*.

ACKNOWLEDGMENTS

We gratefully acknowledge Lucy Cherbas (Indiana University) for advice on *Drosophila* cell transfections and for supplying the *Drosophila* expression vector. We thank Ron Blackman (University of Illinois) for providing the S2 cell line and Monsanto for supplying the original pMON458 vector. We acknowledge Sherry Xu for cloning and sequencing numerous PCR products and Hua Lou for valuable discussions throughout this project.

This work was supported by National Institutes of Health grant GM39025 RO1 (M.A.S.) and U.S. Department of Agriculture competitive research grant AG 92-37301-7964 (A.J.M.).

REFERENCES

1. **Aroian, R. V., A. D. Levy, M. Kogo, Y. Ohshima, J. M. Kramer, and P. W. Sternberg.** 1993. Splicing in *Caenorhabditis elegans* does not require an AG at the 3′ splice acceptor site. Mol. Cell. Biol. **13**:626–637.
2. **Black, D. L., B. Chabot, and J. A. Steitz.** 1985. U2 as well as U1 small nuclear ribonucleoproteins are involved in pre-messenger RNA splicing. Cell **42**:737–750.
3. **Cherbas, L., K. Lee, and P. Cherbas.** 1991. Identification of ecdysone response elements by analysis of the *Drosophila Eip28/29* gene. Genes Dev. **5**:120–131.
4. **Conrad, R., R. F. Liou, and T. Blumenthal.** 1993. Functional analysis of a *C. elegans trans*-splice acceptor. Nucleic Acids Res. **21**:913–919.
5. **Conrad, R., R. F. Liou, and T. Blumenthal.** 1993. Conversion of a *trans*-spliced *C. elegans* gene into a conventional gene by introduction of a splice donor site. EMBO J. **12**:1249–1255.
6. **Csank, C., F. M. Taylor, and D. W. Martindale.** 1990. Nuclear pre-mRNA introns: analysis and comparison of intron sequences from *Tetrahymena thermophila* and other eukaryotes. Nucleic Acids Res. **18**:5133–5141.
7. **Fluhr, R., P. Moses, G. Morelli, G. Coruzzi, and N.-H. Chua.** 1986. Expression dynamics of the pea *rbcS* multigene family and organ distribution of the transcripts. EMBO J. **5**:2063–2071.
8. **Goodall, G. J., and W. Filipowicz.** 1989. The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell **58**:473–483.
9. **Goodall, G. J., and W. Filipowicz.** 1990. The minimum functional length of pre-mRNA introns in monocots and dicots. Plant Mol. Biol. **14**:727–733.
10. **Goodall, G. J., and W. Filipowicz.** 1991. Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants. EMBO J. **10**:2635–2644.
11. **Goodall, G. J., T. Kiss, and W. Filipowicz.** 1991. Nuclear RNA splicing and small nuclear RNAs and their genes in higher plants. Oxf. Surv. Plant Mol. Cell Biol. **7**:255–296.
12. **Green, M. R.** 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. **7**:559–599.
13. **Guo, M., P. C. H. Lo, and S. M. Mount.** 1993. Species-specific signals for the splicing of a short *Drosophila* intron in vitro. Mol. Cell. Biol. **13**:1104–1118.
14. **Guthrie, C., and B. Patterson.** 1988. Spliceosomal snRNAs. Annu. Rev. Genet. **22**:387–419.
15. **Hanley-Bowdoin, L., J. S. Elmer, and S. G. Rogers.** 1988. Transient expression of heterologous RNAs using tomato golden mosaic virus. Nucleic Acids Res. **16**:10511–10529.
16. **Li, W., and J. E. Shaw.** 1993. A variant transposable element in the nematode *C. elegans* could encode a novel protein. Nucleic Acids Res. **21**:59–67.
17. **Lou, H., A. J. McCullough, and M. A. Schuler.** 1993. Expression of maize *Adh1* intron mutants in tobacco nuclei. Plant J. **3**:393–403.
18. **Lou, H., A. J. McCullough, and M. A. Schuler.** 1993. 3′ splice site selection in dicot plant nuclei is position dependent. Mol. Cell. Biol. **13**:4485–4493.
19. **Luhrmann, R., B. Kastner, and M. Bach.** 1990. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. Biochem. Biophys. Acta **1087**:265–292.
20. **McCullough, A. J., H. Lou, and M. A. Schuler.** 1991. *In vivo* analysis of plant pre-mRNA splicing using an autonomously replicating vector. Nucleic Acids Res. **19**:3001–3009.
21. **McCullough, A. J., H. Lou, and M. A. Schuler.** 1993. Factors affecting authentic 5′ splice site selection in plant nuclei. Mol. Cell. Biol. **13**:1323–1331.

21a. **McCullough, A. J., and M. A. Schuler.** Unpublished data.

22. **Moore, M., C. C. Query, and P. Sharp.** Splicing of precursors to messenger RNAs by the spliceosome. *In* R. Gesteland and J. Atkins (ed.), The RNA world, in press. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
23. **Mount, S. M., C. Burks, G. Hertz, G. D. Stormo, O. White, and C. Fields.** 1992. Splicing signals in *Drosophila*: intron size, information content, and consensus sequences. Nucleic Acids Res. **20**:4255–4262.
24. **Newman, A. J., and C. Norman.** 1991. Mutations in yeast U5 snRNA alter the specificity of 5′ splice-site cleavage. Cell **65**:115–123.
25. **Newman, A. J., and C. Norman.** 1992. U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. Cell **68**:743–754.
26. **Parker, R., P. G. Siliciano, and C. Guthrie.** 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. Cell **49**:229–239.
27. **Reich, C. I., R. W. VanHoy, G. L. Porter, and J. A. Wise.** 1992. Mutations at the 3′ splice site can be suppressed by compensatory base changes in U1 snRNA in fission yeast. Cell **69**:1159–1169.
28. **Seraphin, B., L. Kretzner, and M. Rosbash.** 1988. A U1 snRNA: pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5′ cleavage site. EMBO J. **7**:2533–2538.
29. **Siliciano, P. G., and C. Guthrie.** 1988. 5′ splice site selection in yeast: genetic alterations in base pairing with U1 reveal additional requirements. Genes Dev. **2**:1258–1267.
30. **Steitz, J. A., D. L. Black, V. Gerke, K. A. Parker, A. Kramer, D. Frendewey, and W. Keller.** 1988. Functions of the abundant U-snRNPs, p. 115–154. *In* M. L. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Berlin.
31. **Vallette, F., E. Mege, A. Reiss, and M. Adesnik.** 1989. Construction of mutant and chimeric genes using the polymerase chain reaction. Nucleic Acids Res. **17**:723–733.
32. **Wassarman, D. A., and J. Steitz.** 1992. Interactions of small nuclear RNA's with precursor messenger RNA during *in vitro* splicing. Science **257**:1918–1925.
33. **Woolford, J. L.** 1989. Nuclear pre-mRNA splicing in yeast. Yeast **5**:439–457.
34. **Wu, J., and J. L. Manley.** 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. Genes Dev. **3**:1553–1561.
35. **Zhuang, Y., and A. M. Weiner.** 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell **46**:827–835.
36. **Zhuang, Y., and A. M. Weiner.** 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. Genes Dev. **3**:1545–1552.