The Plant Cell, Vol. 8, 2295–2307, December 1996

Interactions across Exons Can Influence Splice Site Recognition in Plant Nuclei

Andrew J. McCullough,[1] Clair E. Baynton, and Mary A. Schuler[2]

Department of Plant Biology, University of Illinois, Urbana, Illinois 61801

In vivo analyses of *cis*-acting sequence requirements for pre-mRNA splicing in tobacco nuclei have previously demonstrated that the 5′ splice sites are selected by their position relative to AU-rich elements within plant introns and by their degree of complementarity to the U1 small nuclear RNA. To determine whether the presence of adjacent introns affects 5′ splice site recognition in plant nuclei, we have analyzed the in vivo splicing patterns of two-intron constructs containing 5′ splice site mutations in the second intron. These experiments indicated that the splice site selection patterns in plant nuclei are defined primarily by sequences within the intron (intron definition) and secondarily by weak interactions across exons (exon definition). The effects of these secondary interactions became evident only when mutations in the downstream 5′ splice site decreased its functionality and differed depending on the availability of cryptic splice sites close to the mutant site. In β-conglycinin chimeric transcripts containing multiple cryptic 5′ splice sites, the presence of an intact upstream intron significantly increased splicing at the downstream 5′ splice sites in a polar fashion without activating exon skipping. In a natural β-conglycinin transcript, which does not contain cryptic 5′ splice sites, mutation of the first nucleotide of the downstream intron activated an array of noncanonical 5′ and 3′ splice sites and some exon skipping.

INTRODUCTION

Accurate expression of protein-coding sequences in plant systems demands that introns be correctly excised from pre-mRNA transcripts to produce mature, functional mRNAs. Inaccurate recognition of intron and exon sequences can result in the production of cryptically spliced transcripts that introduce missense and frameshift mutations into critical coding sequences. Because translation of truncated and mutated proteins from the inaccurately spliced transcripts can substantially reduce foreign protein expression in transgenic plants, it is essential that the sequence motifs modulating intron recognition and alternative splicing be carefully defined for plant nuclei.

In mammalian and yeast systems, three structural elements in the pre-mRNA are required for proper intron recognition (Woolford, 1989; Moore et al., 1993; Sharp, 1994). The first element located at the 5′ splice site is a series of nine semiconserved nucleotides spanning the 5′ exon–intron border (AG/GUAAGUA) (underlined nucleotides designate absolutely conserved nucleotides; the slash denotes the 5′ exon–intron border). The second element is a branch point sequence, YNYURAY in mammals and UACUAAC in yeast (branch point nucleotide is underlined), that is used for the formation of the intron lariat intermediate (Woolford, 1989; Moore et al., 1993; Sharp, 1994). The final element located at the 3′ splice site (YAG/; the slash denotes the 3′ intron–exon border) is preceded in mammalian introns by an extended polypyrimidine tract primarily containing uridines (Roscigno et al., 1993).

Although highly conserved /GU and AG/ dinucleotides flank all plant introns, other nucleotides spanning the 5′ splice site and preceding the 3′ splice site are more variable than in other eukaryotic systems (Brown, 1986; Hanley and Schuler, 1988; Goodall et al., 1991). Plant introns contain branch point consensus sequences (Liu and Filipowicz, 1996; Simpson et al., 1996) but lack the polypyrimidine tracts found in mammalian introns (Brown, 1986; Hanley and Schuler, 1988; Goodall et al., 1991). In addition, they are significantly richer (11 to 19%) in adenosine and uridine residues than their adjacent exons (Goodall and Filipowicz, 1989; Csank et al., 1990). Although the absolute level of AU richness differs for monocot and dicot introns, similar transitions in AU richness occur for both groups of introns: dicot introns typically contain 74% AU versus 55% AU in adjacent exons; monocot introns contain 56% AU versus 44% AU in adjacent exons (Goodall and Filipowicz, 1989).

Considerable evidence using autonomously replicating Agrobacterium–geminivirus expression systems has indicated that AU-rich elements spread throughout the length of plant introns are critical for 5′ and 3′ splice site recognition in dicot nuclei (Lou et al., 1993a, 1993b; McCullough et al., 1993). On

[1] Current address: Department of Biochemistry, Baylor College of Medicine, Houston, TX 77030.

[2] To whom correspondence should be addressed.

the basis of studies using single intron constructs, we have proposed a model for intron recognition showing that 5′ and 3′ splice sites are selected by their position upstream (5′ splice site) or downstream (3′ splice site) of the AU-rich intron elements (Lou et al., 1993a, 1993b; McCullough et al., 1993). Splicing studies performed with electroporated dicot and monocot protoplasts have generated similar conclusions (Goodall and Filipowicz, 1989, 1991; Carle-Urioste et al., 1994; Luehrsen and Walbot, 1994a, 1994b). Because of these position-dependent selection schemes, inactivation of the normal 5′ splice site in intron 1 of the pea ribulose bisphosphate carboxylase small subunit (*rbcS3A*) gene occurs only by single point mutations of the +1G or +2U (first and second nucleotides of the intron, respectively) in the intron or by double point mutations at other positions (McCullough et al., 1993). In all of these cases, inactivation of the normal 5′ splice site resulted in the activation of multiple cryptic 5′ splice sites in the adjacent exon and intron. Enhancement of these cryptic sites to perfect 5′ splice site consensus sequences and replacement substitutions has demonstrated that weak 5′ splice sites positioned at the normal AU transition point between the exon and intron compete effectively with perfect consensus 5′ splice sites placed within the flanking exon or intron. On the basis of this and an entirely different replacement series having equal strength competing sites (McCullough and Schuler, 1997), we have concluded that 5′ splice sites in plant nuclei are selected by their degree of U1 small nuclear RNA (snRNA) complementarity and by their position relative to the AU transition point.

The presence of multiple introns within most precursor transcripts adds a complexity to the splicing process that has not yet been directly examined in plant systems. Although the experiments described above suggest that plant introns are intron-defined, it remains possible that the splicing of individual introns may be influenced by the presence and/or splicing of adjacent introns. Substantial evidence has indicated that the splicing of vertebrate introns is dramatically affected by the integrity of the flanking exons. For vertebrate transcripts, which contain short internal exons averaging 134 nucleotides in length and extremely long introns (Hawkins, 1988), the primary unit recognized during the early stages of spliceosome assembly is the exon (exon definition) (Robberson et al., 1990; Talerico and Berget, 1990; Berget, 1995). Mutations inactivating the downstream 5′ splice site or lengthening the intervening exon decrease splicing complex assembly on the upstream intron and overall splicing of the transcript (Robberson et al., 1990). Interactive binding of U1 snRNA at the downstream 5′ splice site and U2AF^{65} protein at the upstream 3′ splice site facilitates exon interactions that allow for efficient splicing (Hoffman and Grabowski, 1992). One diagnostic feature for this mode of exon recognition is that in the absence of nearby cryptic 5′ splice sites, mutations in the downstream 5′ splice site reduce splicing of the upstream intron and enhance skipping of the intervening exon (Talerico and Berget, 1990).

To determine whether upstream introns affect recognition of downstream introns to any extent in plant nuclei, we expressed transcripts containing two introns in tobacco nuclei by using an autonomously replicating Agrobacterium–geminivirus expression system (McCullough et al., 1991). The in vivo splicing patterns of these constructs indicated that plant splice site selection patterns are defined primarily by sequences within the intron (intron definition) and secondarily by weak interactions across exons (exon definition). The effects of these secondary interactions became evident only when mutations in the downstream 5′ splice site decreased its functionality and differed depending on the availability of alternate cryptic 5′ splice sites close to the mutant 5′ splice site.

RESULTS

Splicing of β-Conglycinin Intron 4/*rbcS3A* Intron 1 Chimeric Transcripts

The first double intron construct that we analyzed contains the β-conglycinin intron 4 (115 nucleotides; 72% AU) fused upstream of *rbcS3A* intron 1 (469 nucleotides; 73% AU) (Figure 1). The chimeric exon positioned between these introns is 220 nucleotides and has an AU content of 50%. Previous analyses of single intron constructs containing wild-type β-conglycinin intron 4 or *rbcS3A* intron 1 (McCullough et al., 1991) demonstrated that both of these introns are spliced efficiently (>95% for β-conglycinin intron 4; 85% for *rbcS3A* intron 1) in *Nicotiana benthamiana* leaf disc cells by using the autonomously replicating Agrobacterium–geminivirus expression system described by McCullough et al. (1991). Using this same expression system, we monitored the splicing efficiencies for each of these introns by quantitative reverse transcriptase–polymerase chain reaction (RT-PCR) gel blot analysis of total RNA (McCullough and Schuler, 1993) with oligonucleotide primers specific for β-conglycinin intron 4 (Figure 2A), *rbcS3A* intron 1 (Figure 2B), or the entire hybrid transcript (Figure 2C). In the wild-type β-conglycinin intron 4/*rbcS3A* intron 1 construct (115wt/3A1wt) (Figures 2A to 2C, lanes 1), each of the introns is spliced accurately and efficiently. To delineate the rate-limiting step in intron excision and the frequency of exon skipping, we performed RT-PCR analysis across the entire length of the transcript by using primers for the terminal exons (Figure 2C, lane 1) and cloned and sequenced the products. The three products generated in this reaction correspond to precursor transcript, an intermediate transcript lacking the 115wt intron but retaining the 3A1wt intron, and mature transcript lacking both introns (sequence data not shown). Other minor products that might correspond to intermediates retaining the first intron and lacking the second intron were not characterized because they represent <5% of the spliced product. The lack of transcripts containing only the 115wt intron indicates that splicing of β-conglycinin intron 4 precedes splicing of *rbcS3A* intron 1 in the majority of the chimeric transcripts. Small products indicative of exon skipping have never been observed for this wild-type construct or any mutants derived from it (Figure 2C).

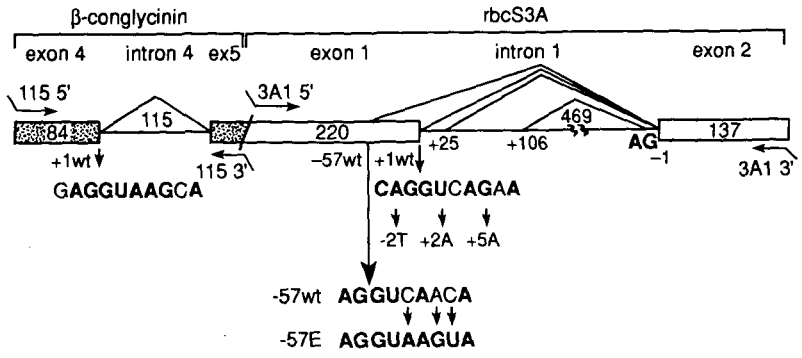

Figure 1. Chimeric β-Conglycinin Intron 4/*rbcS3A* Intron 1 Transcripts Expressed in Transfected Leaf Disc Nuclei.

The β-conglycinin intron 4/*rbcS3A* intron 1 pre-mRNA transcript generated in vivo by transcription from the coat protein promoter of the pMON458 construct is drawn to scale. Open boxes represent exons, and solid lines represent introns. Sequences of the wild-type 5′ splice sites for each intron and the cryptic site at −57 in *rbcS3A* intron 1 are indicated in boldface letters representing nucleotides complementary to the 5′ end of U1 snRNA. Sequences of the mutant 5′ splice sites at positions +1 and −57 used in this study are shown below each wild-type sequence. Lines connecting the 5′ splice sites with each normal 3′ splice site indicate the various splicing patterns documented in this study. Arrows marked 115 5′, 115 3′, 3A1 5′, and 3A1 3′ represent the positions of oligonucleotide primers used for RT-PCR analysis.

To define the extent to which the upstream intron affects splice site selection in this downstream intron, we introduced mutations into the 5′ splice site of *rbcS3A* intron 1. Single intron constructs containing mutations in the canonical +1G or +2U of *rbcS3A* intron 1 activate cryptic 5′ splice sites at position −57 in the upstream exon and at positions +25 and +106 in the intron (Figure 1; McCullough et al., 1993). These wild-type cryptic sites are complementary to U1 snRNA at five of nine, three of nine, and six of nine positions, respectively; the normal 5′ splice site at position +1 is complementary to U1 snRNA at seven of nine positions. In the 3A1.+2A single intron transcript (Figure 2B, lane 2), the strongest +106 wild-type site (+106wt) is used in preference to the −57 wild-type (−57wt) site, and the weak +25 wild-type (+25wt) site was used least efficiently (Figure 3; McCullough et al., 1993).

Several points can be made concerning the comparable multiintron construct (Figure 2B, lane 3, and Figure 3). First, the double intron transcript containing the +2A mutation (/GU mutated to /GA) in *rbcS3A* intron 1 (115wt/3A1.+2A) splices appreciably better (51% splicing efficiency) than does the single intron transcript containing 3A1.+2A (17% splicing efficiency). Second, the cryptic splice sites selected in the double intron transcript are the same as in the single intron transcript, but the usage of each differs (Figure 2B, lanes 2 versus 3). In the presence of an intact upstream intron, usage of each cryptic site is enhanced in a polar fashion: splicing at the −57wt site is enhanced 6.5-fold in the double intron transcript compared with 3.1-fold enhancement at the weak +25wt site and 1.4-fold enhancement at the +106wt site. As a result of this polar effect, usage of the −57wt site predominates over the +106wt site in the double intron construct (Figure 2B, lane 3, and Figure 3) compared with the single intron construct (Figure 2B, lane 2). Third, the increased usage of the −57wt site results from enhanced splicing of the precursor transcript, not decreased usage of downstream sites (i.e., the same percentages of total transcripts are spliced at the +106wt site in single and double intron constructs; Figure 3A). Lastly, increased splicing at the cryptic sites in *rbcS3A* intron 1 has little effect on splicing of the upstream intron (Figure 2A).

For single intron constructs containing *rbcS3A* intron 1, mutation of −2A to a U residue, which decreases the number of contiguous consensus nucleotides in the 5′ splice site, reduces the overall splicing efficiency to 67% of the wild-type control without activating any alternate cryptic sites (McCullough et al., 1993). Double intron constructs containing the −2U mutation in *rbcS3A* intron 1 (115wt/3A1.−2T) splice as efficiently at the mutant +1 site as the original 115wt/3A1wt transcript and marginally at the cryptic −57wt site in the upstream exon (Figure 2B, lane 4). In single intron constructs, mutation of +5G to an A residue weakens the +1 site only when it is placed in competition with a perfect consensus splice site at −57 (−57E) (McCullough et al., 1993). Double intron constructs containing the +5A mutation in *rbcS3A* intron 1 and the cryptic −57wt site (115wt/3A1.+5A) splice as efficiently at the weakened +1 site as in the wild-type 115wt/3A1wt transcript (Figure 2B, lane 5). In contrast to the −2U mutation, the +5A mutation does not activate any cryptic sites. Both of these constructs efficiently splice the upstream 115wt intron (Figure 2A).

Splicing of Chimeric Transcripts Containing an Enhanced Exonic 5′ Splice Site

cis-competitions between the −57E site containing nine contiguous consensus nucleotides and the weaker +1wt site of *rbcS3A* intron 1 have indicated that the number of consensus

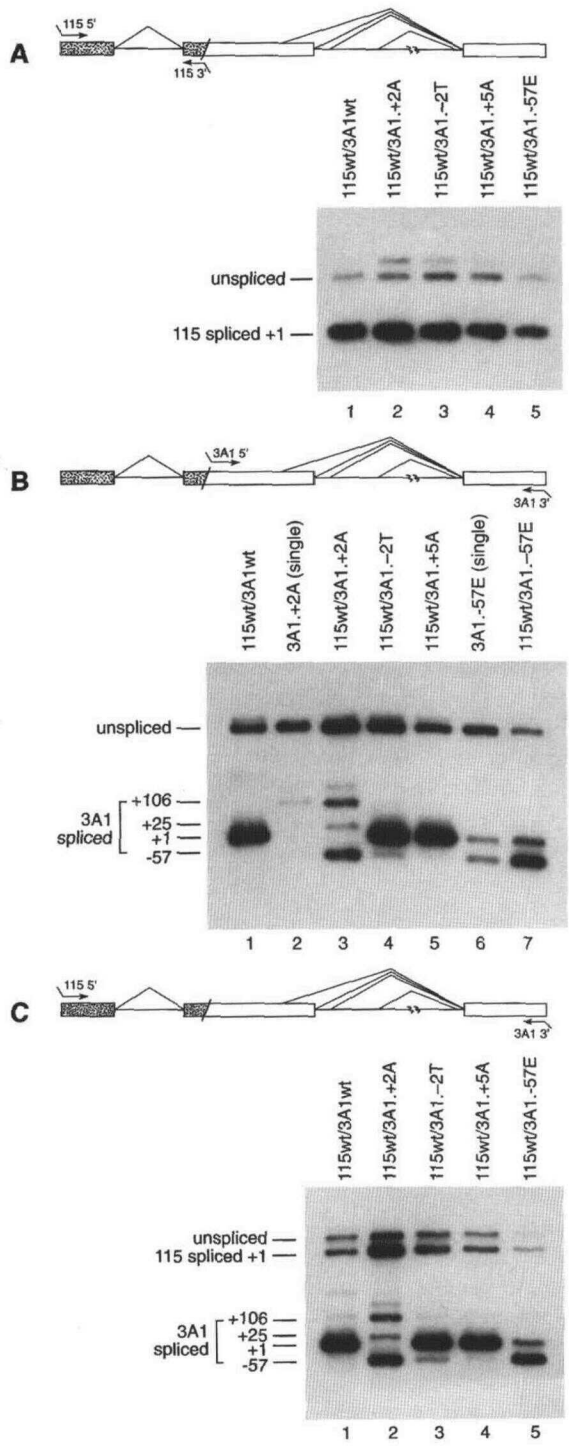

Figure 2. Splicing of the β-Conglycinin Intron 4/*rbcS3A* Intron 1 Constructs.

Constructs containing wild-type β-conglycinin intron 4 (115wt) and a wild-type or mutant version of *rbcS3A* intron 1 (3A1) were transfected into *N. benthamiana* leaf discs, and total RNA was isolated 4 days after transfection. Shaded boxes represent exon sequences derived from β-conglycinin exons 4 and 5. Open boxes represent exon sequences derived from *rbcS3A* exons 1 and 2. The arrows designate oligonucleotide primers used for RT-PCR analysis. The splicing patterns of the transcripts were defined by quantitative RT-PCR gel blot analysis (McCullough and Schuler, 1993) by using oligonucleotide primers specific for β-conglycinin intron 4 **(A)**, *rbcS3A* intron 1 **(B)**, or the entire transcript **(C)**.

The probe used for analysis of β-conglycinin intron 4 splicing **(A)** spanned the entire length of the β-conglycinin gene present in this construct, *rbcS3A* intron 1 splicing **(B)** spanned *rbcS3A* exon 1 sequences upstream of the −57 cryptic 5′ splice site, or the entire transcript's splicing **(C)** spanned *rbcS3A* exon 2 sequences. The construct analyzed is designated above each lane. The positions of PCR products corresponding to unspliced transcripts, correctly spliced transcripts (+1), and transcripts spliced between the −57, +25, and +106 cryptic 5′ splice sites are shown at the left of **(A)** to **(C)**. In lane 3 of **(B)** and lane 2 of **(C)**, the fourth band above the +106 spliced product corresponds to a hybrid formed between the +1 and +106 spliced products during the last PCR cycle. In **(A)**, the band above the unspliced product corresponds to a hybrid formed between unspliced and spliced products during the last PCR cycle.

nucleotides and the position of a prospective splice site relative to the AU transition point modulate 5′ splice site recognition (McCullough et al., 1993). In single intron constructs, these sites are used with efficiencies of 60% (−57E) and 40% (+1wt) (Figure 2B, lane 6, and Figure 4B). In double intron transcripts, the −57E site is used with a splicing efficiency of 78% compared with 22% at the +1wt site (Figure 2B, lane 7). As in the case of the 115wt/3A1.+2A mutant, the increased usage of the −57E site results from enhanced splicing of the precursor transcript and not decreased usage of the +1wt site (Figure 4A). We conclude that even in instances in which the cryptic 5′ splice site closest to the upstream intron is enhanced to a perfect consensus sequence, a weaker site at the normal AU transition point between exon and intron can still be actively selected.

These experiments indicate that an upstream intron can influence the splicing efficiency and 5′ splice site selection patterns of the downstream intron. This effect is polar, with the 5′ splice site nearest the upstream intron enhanced to the greatest extent. Most importantly, the effects of these secondary interactions across exons are observed only after the balance between the normal and cryptic sites is disturbed by weakening the normal +1wt site or strengthening the upstream −57wt cryptic site.

Splicing of Partially Processed Chimeric Transcripts

As a control for these double intron constructs, additional transcripts containing 120 nucleotides of mature β-conglycinin sequences (115M) upstream of the 3A1.−57E and 3A1.+2A

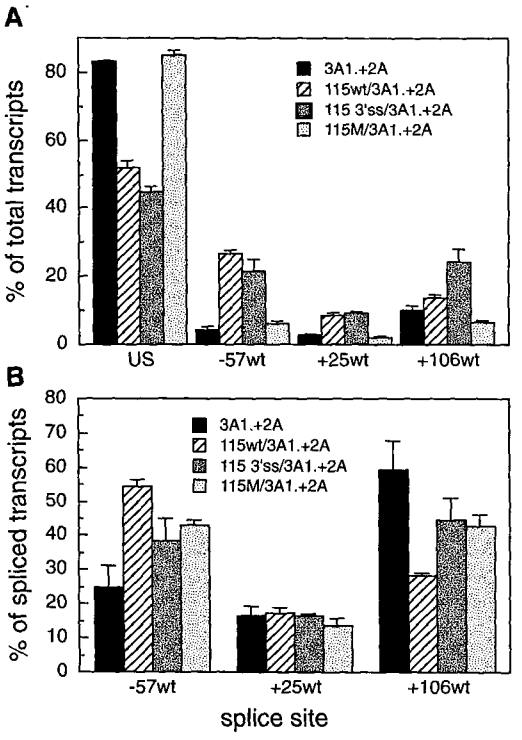

Figure 3. Comparison of the Splicing Efficiencies for the 3A1.+2A Intron Transcripts.

(A) Splicing patterns recorded as the percentage of total transcript accumulated (splicing efficiency). The splicing efficiencies for each cryptic 5′ splice site are compared for transcripts containing one intron (3A1.+2A), two introns (115wt/3A1.+2A), an intact 3′ splice site (115 3′ss/3A1.+2A), or spliced exon sequences (115M/3A1.+2A) upstream of the 3A1.+2A intron. The truncated construct (115 3′ss/3A1.+2A) contains the last 50 nucleotides of β-conglycinin intron 4 fused upstream of the 3A1.+2A intron; the mature spliced construct (115M/3A1.+2A) contains the same β-conglycinin exon sequences shown in Figure 1. Each reported splicing efficiency represents the average of at least three independent transfections with the corresponding standard errors.

(B) Splicing patterns recorded as the percentage of all spliced transcripts. Designations are the same as given in **(A)**.

mutant introns (115M/3A1.−57E and 115M/3A1.+2A) were expressed in vivo. The 115M sequences present in both of these constructs correspond exactly to the exon sequences present in the 115wt/3A1wt double intron construct (Figure 1), with a precise deletion of β-conglycinin intron 4. Removal of the upstream intron from the 115wt/3A1.+2A transcript reduces the overall splicing efficiency of the 115M/3A1.+2A transcript and usage of each cryptic site to the same level as that observed for the single intron 3A1.+2A construct (Figure 3A and Figure 5C, lanes 1 versus 2). Relative to the overall collection of spliced transcripts, usage of the −57wt site in 115M/3A1.+2A is intermediate between the double and single intron constructs and equivalent to usage of the +106wt site (Figure 3B). Usage of the +25wt site remains approximately the same.

Removal of the upstream intron from the 115wt/3A1.−57E transcript reduces the overall splicing efficiency of the 115M/3A1.−57E transcript to a level comparable to that of the single intron 3A1.−57E transcript (Figure 4A and Figure 5C, lanes 4 versus 5). The usages of the −57E and +1wt splice sites are approximately the same as those observed for the single intron transcript (3A1.−57E) (Figure 4B) and substantially different from that observed for the double intron construct (115wt/3A1.−57E).

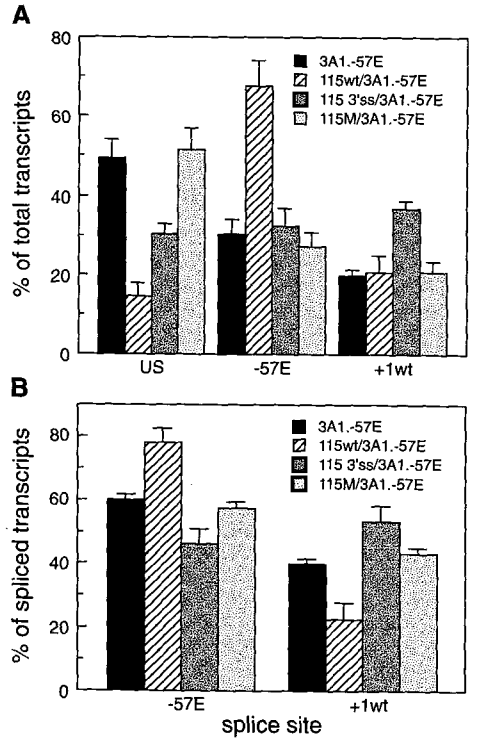

Figure 4. Comparison of the Splicing Efficiencies for the 3A1.−57E Intron Transcripts.

(A) Splicing patterns recorded as the percentage of total transcript accumulated (splicing efficiency). The splicing efficiencies for the +1wt and −57E 5′ splice sites in the 3A1.−57E mutant are compared for transcripts containing one intron (3A1.−57E), two introns (115wt/3A1.−57E), an intact 3′ splice site (115 3′ss/3A1.−57E), or spliced exon sequences (115M/3A1.−57E) upstream of the 3A1.−57E intron. Each reported splicing efficiency represents the average of at least three independent transfections with the corresponding standard errors.

(B) Splicing patterns recorded as the percentage of all spliced transcripts. Designations are the same as given in **(A)**.

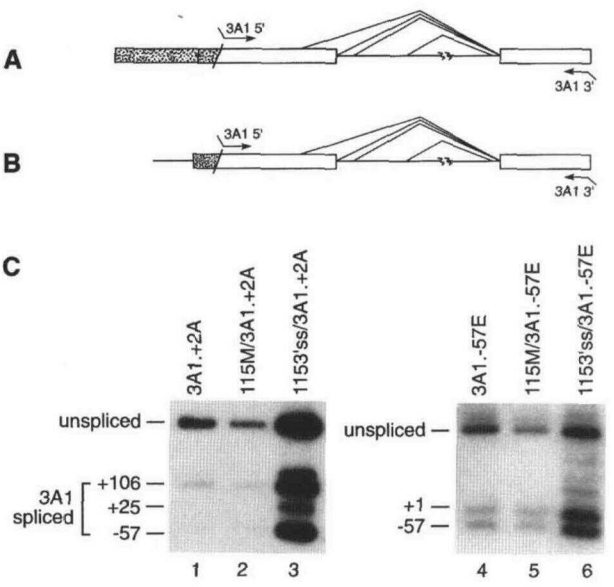

Figure 5. Splicing of Mature and Truncated β-Conglycinin/*rbcS3A* Fusion Constructs.

(A) Schematic diagram of the 115M/3A1.+2A and 115M/3A1.−57E constructs containing 120 nucleotides of mature (spliced) β-conglycinin sequences upstream of *rbcS3A* intron 1. Shaded boxes represent exon sequences from β-conglycinin exons 4 and 5. Open boxes represent exon sequences derived from *rbcS3A* exons 1 and 2. The arrows represent the oligonucleotide primers used for RT-PCR analysis.

(B) Schematic diagram of the truncated 115 3'ss/3A1.+2A and 115 3'ss/3A1.−57E constructs containing the last 50 nucleotides of β-conglycinin intron 4 fused upstream of *rbcS3A* intron 1. Shaded box represents exon sequences from β-conglycinin exon 5. Open boxes represent exon sequences derived from *rbcS3A* exons 1 and 2.

(C) Constructs containing the mature or truncated β-conglycinin sequences upstream of mutant versions of *rbcS3A* intron 1 (3A1.+2A or 3A1.−57E) were transfected into *N. benthamiana* leaf discs, and total RNA was isolated 4 days after transfection. The splicing patterns of the transcripts were defined by quantitative RT-PCR gel blot analysis using oligonucleotide primers specific for *rbcS3A* intron 1 and a probe spanning *rbcS3A* exon 1 sequences upstream of the −57 cryptic 5′ splice site. The construct analyzed is designated above each lane. The positions of PCR products corresponding to unspliced transcripts, correctly spliced transcripts (+1), and cryptically spliced transcripts (−57, +25, and +106) are shown to the left of each gel.

Truncation of the Upstream Intron

Although this analysis demonstrates that the presence of an upstream intron can influence the recognition of the downstream intron when the balance between the normal (+1wt) and cryptic (−57E) site is altered, it is not clear whether enhanced recognition of the downstream 5′ splice site requires splicing of the upstream intron or simply the presence of a functional 3′ splice site. To address this issue, we expressed in vivo truncated transcripts containing the last 50 nucleotides of β-conglycinin intron 4 fused upstream of *rbcS3A* intron 1 mutants (as shown in Figure 1).

The truncated intron construct containing the +2A mutation in *rbcS3A* intron 1 (115 3′ss/3A1.+2A) splices as efficiently as the double intron construct carrying this mutation (115wt/3A1.+2A) (Figure 3A, and Figure 5C, lane 3, versus Figure 2B, lane 3). The splicing preferences at each wild-type cryptic site (−57wt, +25wt, and +106wt) are similar to the transcripts containing a spliced upstream sequence (115M/3A1.+2A) and intermediate between single and double intron constructs (Figure 3B). The truncated intron construct carrying the enhanced −57E site and weaker +1wt site (115 3′ss/3A1.−57E) splices with an overall efficiency intermediate between the single and double intron constructs (Figure 4A, and Figure 5C, lane 6, versus Figure 2B, lanes 6 and 7). Usage of the −57E site is slightly reduced and usage of the +1wt site is slightly increased in this truncated construct in comparison with its single intron homolog (Figure 4B). We conclude that in the case of the 3A1.+2A mutation, which splices at a series of suboptimal cryptic 5′ splice sites, the presence of an upstream 3′ splice site enhances the overall splicing efficiency of the downstream intron to the same extent as a full-length upstream intron. In the case of the 3A1.−57E mutation, which contains two strong 5′ splice sites (one of which is a consensus site), the presence of an upstream 3′ splice site enhances splicing but not to the same extent as an intact intron.

Splicing of β-Conglycinin Intron 3/Intron 4 Transcripts

To determine whether mutations in naturally adjacent introns affect splicing of one another, we expressed in vivo a second double intron transcript (Figure 6) extending from the 3′ end of intron 2 to the 3′ end of exon 5 of the β-conglycinin gene. This transcript contains 37 nucleotides preceding the 3′ splice site of intron 2, β-conglycinin introns 3 (85 nucleotides; 67% AU) and 4 (115 nucleotides; 72% AU), as well as the full lengths of exons 3, 4, and 5. The natural exon positioned between introns 3 and 4 is 297 nucleotides and has an AU content of 55%. Previous analyses demonstrated that multiintron constructs containing β-conglycinin introns 3, 4, and 5 are spliced efficiently in *N. benthamiana* nuclei (McCullough et al., 1991; data not shown).

In our study, the splicing efficiencies of the introns contained within double intron constructs were monitored by RT-PCR analysis with primers specific for the terminal sequences (Figure 6). Amplification of wild-type β-conglycinin intron 3/intron 4 transcripts (85wt/115wt) expressed in vivo generated three products (Figure 7, lane 1). Cloning and sequence analysis demonstrated that these products correspond to the precursor transcript, an intermediate lacking the 115wt intron and retaining the 85wt intron, and the mature transcript lacking both introns (sequencing data not shown). PCR products representing intermediate transcripts lacking the first (85wt) intron and retaining the second (115wt) intron were not observed in DNA gel blot analyses or subsequent clonings. Quantita-

tion of the splicing efficiencies for each intron indicated that the 115wt intron had been removed from 83.1% (standard error of 1.1%) of the transcripts, the 85wt intron had been removed from 48.2% (standard error of 3.1%) of the transcripts, and both introns are retained in 17% of the transcripts (Figure 8A). The absence of intermediates lacking only the first (85wt) intron and its lower splicing efficiency indicated that splicing of intron 3 rate limits production of mature mRNA from this double intron transcript.

To determine whether mutations in the second intron of this construct activate cryptic 5′ splice sites, we introduced a +1G-to-A mutation (/GU to /AU) into β-conglycinin intron 4 (Figure 6). The double intron construct carrying this mutation and a wild-type upstream intron (85wt/115.+1A) accumulated a significant amount of unspliced precursor transcript and three classes of spliced products (Figure 7, lane 2). (Note that dead-end lariat intermediates, such as those accumulating in mammalian +1G-to-A mutants, cannot be detected in this RT-PCR assay.) The largest of the spliced products detected in our assay corresponds to an intermediate lacking the 85wt intron and retaining the 115.+1A intron. The second largest set represents an array of transcripts lacking the first (85wt) intron and different sections of the second (115.+1A) intron. Sequencing of multiple RT-PCR clones spanning this region indicated that in surprising contrast to most other introns we have analyzed, the 115.+1A intron is excised from this transcript by using multiple sets of noncanonical 5′ and 3′ splice sites rather than cryptic canonical 5′ /GU and 3′ AG/ splice sites (Figure 6).

The first of these noncanonical products, designated C in Figure 6, represents a transcript spliced between the mutant /AU at +1 (+1/AU) and an AC positioned nine nucleotides (AC/−9) upstream from the normal 3′ splice site. The second product, designated D, represents a transcript spliced between the same mutant /AU at +1 and an AA positioned seven nucleotides (AA/−7) upstream from the 3′ splice site. A third product, designated B, represents a transcript spliced between the /UU dinucleotide 27 nucleotides (+27/UU) downstream from the normal 5′ splice site and the normal 3′ AG (AG/−1). Products corresponding to the full-length intron spliced between the mutant +1/AU and the normal 3′ AG/−1 (product A) were not detected. An independent series of leaf disc transfections demonstrated that all three of these noncanonical splice sites are utilized in single intron transcripts containing the +1G-to-A mutation in intron 4 (C.E. Baynton, J.M. Limon, S.R. Smith, and M.A. Schuler, manuscript in preparation). The smallest

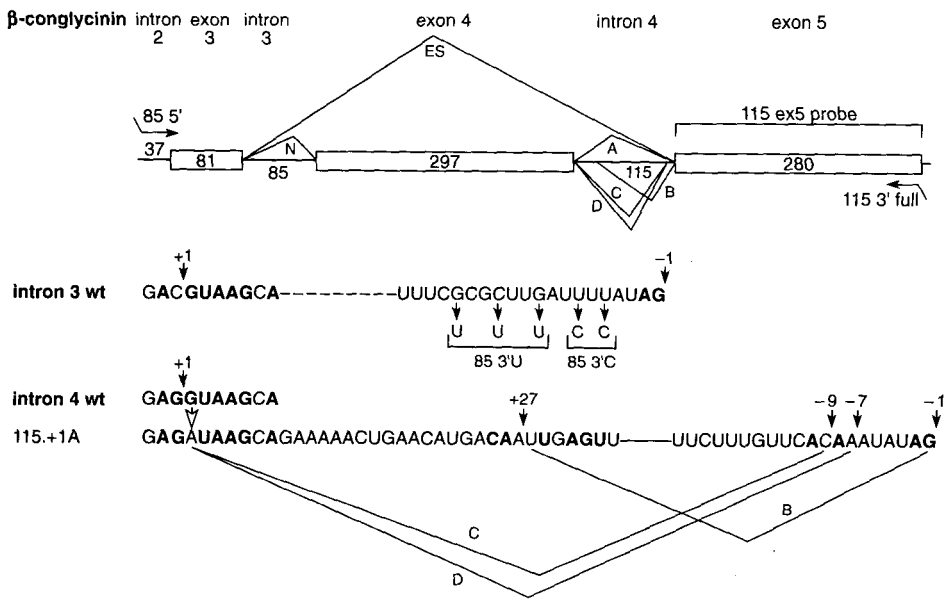

Figure 6. β-Conglycinin Intron 3/Intron 4 Transcripts Expressed in Transfected Leaf Disc Nuclei.

The β-conglycinin intron 3/intron 4 (85wt/115wt) pre-mRNA transcript generated in vivo by transcription from the coat protein promoter on the pMON458 construct is drawn to scale. Open boxes represent exons, and solid lines represent introns. Sequences of the wild-type 5′ splice sites for each intron are indicated with boldface letters representing nucleotides complementary to the 5′ end of U1 snRNA. Mutations introduced into the 3′ splice site of intron 3 or the 5′ splice site of intron 4 are shown below the wild-type sequences. The splicing patterns summarized at the top are designated as follows: ES, exon skipped; N, normal splicing between /AU . . . AG/ in wild-type intron 3; A, normal splicing between /AU . . . AG/ in wild-type intron 4; B, C, and D, noncanonical splicing in 115.+1A mutants as diagrammed at the bottom. Nucleotides in each 5′ splice site complementary to U1 snRNA and canonical nucleotides in the 3′ splice site are shown in boldface.

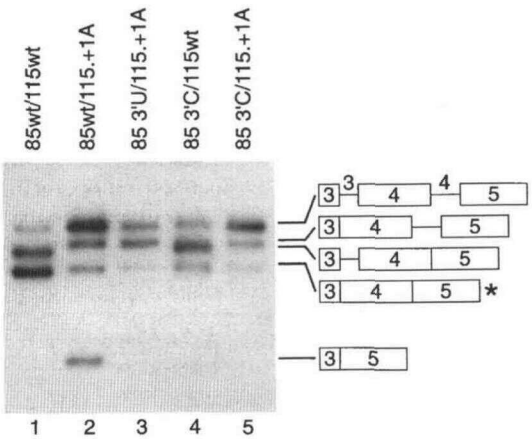

Figure 7. Splicing of the β-Conglycinin Intron 3/Intron 4 Constructs.

Constructs containing wild-type and mutant β-conglycinin introns 3 and 4 were transfected into *N. benthamiana* leaf discs. The splicing patterns of the transcripts were defined by quantitative RT-PCR gel blot analysis (McCullough and Schuler, 1993), using the 85 5′ and 115 3′ full oligonucleotide primers complementary to either end of the precursor transcript and a probe complementary to β-conglycinin exon 5 (Figure 6). The construct analyzed is designated above each lane, and the identities of the PCR products are shown at right. The spliced transcript designated with an asterisk is spliced at canonical splice sites in the 115wt constructs (lanes 1 and 4) and at noncanonical splice sites in the 115.+1A constructs (lanes 2, 3, and 5).

product derived from expression of the 85wt/115.+1A transcript in vivo corresponds to an exon-skipped transcript in which exon 3 has been ligated directly to exon 5.

Quantitation for the three different size classes of these products (Figure 8) was recorded without differentiating between the multiple, noncanonical splice sites used for excision of the 115.+1A intron because collectively all noncanonical splicing events result in intron 4 excision. The data recorded in Figure 8A reflect the frequency at which intron 3 and intron 4 are retained and exon 4 is skipped. The data recorded in Figure 8B reflect the frequency at which the 5′ splice sites in introns 3 and 4 are selected. For the 5′ splice site in intron 3, these data represent the sum of exon skipping and intron 3 excision.

Several points can be made from these comparisons. First, the overall splicing efficiency of the 85wt/115.+1A transcript is substantially reduced relative to its wild-type counterpart (85wt/115wt) (Figure 7, lanes 1 versus 2). The increased accumulation of precursor transcript, which is apparent on these gels, is primarily due to decreased splicing of the mutant β-conglycinin intron 4 (115.+1A). The splicing efficiency for the 85wt intron is only slightly reduced in the presence of a mutant downstream 5′ splice site (Figure 8A). Second, splicing of the mutant 115.+1A intron (including all of the noncanonical sites) occurs in 16.2% (standard error of 7.1%) of all transcripts and represents 19% of the level at which splicing occurs in the wild-type 115wt intron. Third, even though splicing of the 85wt intron is slightly reduced in the 85wt/115.+1A mutant (Figure 8A), the 85wt 5′ splice site is actually selected at a frequency slightly higher than normal (Figure 8B) due to usage of this site in the exon-skipped product.

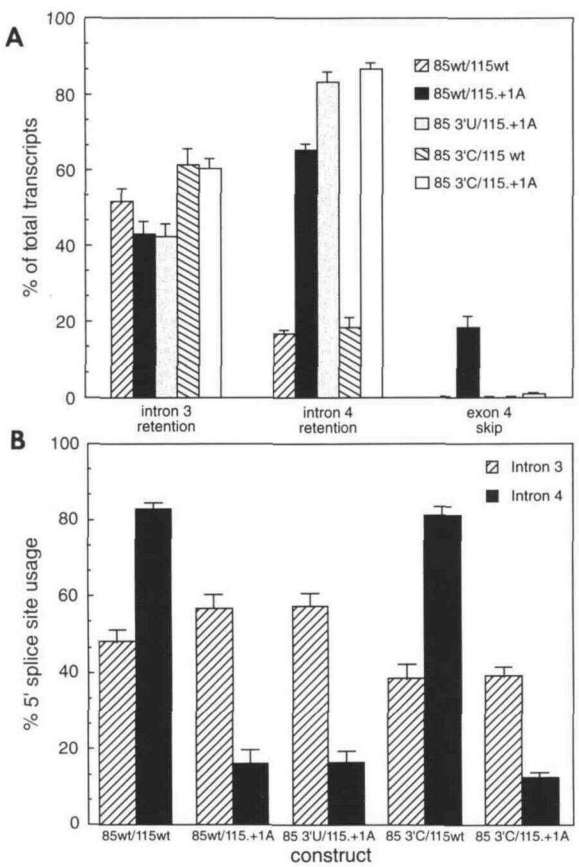

Figure 8. Comparison of the Splicing Efficiencies in β-Conglycinin Intron 3/Intron 4 Constructs.

(A) The relative proportion of intron excision versus exon skipping is defined as the percentage of total accumulated transcript retaining intron 3 or intron 4 versus transcript lacking both introns and exon 4 (exon 4 skip). In this tabulation, the proportion of transcript retaining intron 3 does not include the exon 4 skipped product. Each reported splicing efficiency represents the average of at least three independent transfections with the corresponding standard errors.

(B) The relative usage of the 5′ splice site in intron 3 is defined as the frequency of (intron 3 excision plus exon 4 skipping) divided by the total accumulated transcript. The relative usage of the 5′ splice site in intron 4 is defined as the frequency of (intron 4 excision [using either canonical or noncanonical splice sites]) divided by the total accumulated transcript.

Uridine Enrichment of 3′ Splice Site in Upstream Intron

The occurrence of exon 4 skipping in the 85wt/115.+1A transcript suggests that recognition of the 5′ splice site of intron 4 is important for efficient recognition of the 3′ splice site of intron 3. To explore further this relationship, we strengthened the 3′ splice site of intron 3 by increasing its uridine content at three positions in the region immediately preceding the 3′ splice site (85 3′U) (Figure 6). Transcripts containing this enhanced 3′ splice site and a mutation in the downstream 5′ splice site (85 3′U/115.+1A) completely repressed exon skipping (Figure 7, lane 3, and Figure 8A) while maintaining a high level of intron 3 splicing. As the result of the repression of exon skipping, the intron 4 retention product represents the predominant intermediate accumulating in vivo (Figure 8A). Despite the apparent accumulation of the intron 4 retention product (Figure 7, lane 3), the 5′ splice site usage profiles (Figure 8B) indicate that both the upstream (85 3′U) and the downstream (115.+1A) 5′ splice sites are recognized as efficiently as in 85wt/115.+1A transcripts that skip exon 4. We conclude that recognition and splicing of the 85 3′U intron are increased as a result of the enhanced uridine content near its 3′ splice site and that enhanced recognition of intron 3 precludes recognition of the 3′ splice site in the downstream intron 4 (i.e., exon skipping).

The 3′ splice site of intron 3 was weakened by decreasing its uridine content at two positions without changing its overall pyrimidine content (85 3′C) (Figure 6). Transcripts containing this mutant 3′ splice site and a wild-type downstream intron (85 3′C/115wt) maintain efficient splicing of intron 4 and repress splicing of intron 3 (Figure 7, lane 4) when compared with wild-type 85wt/115wt transcripts (lane 1). The 5′ splice site usage profiles for this mutant (Figure 8B) indicate that reduced recognition of the upstream 3′ splice site (85 3′C) has little effect on recognition of the wild-type 5′ splice site in the downstream intron (115wt).

Transcripts containing the 85 3′C site and a mutation in the downstream 5′ splice site (85 3′C/115.+1A) repress intron 3 splicing to the same extent as in the construct containing a wild-type downstream intron (85 3′C/115wt). The efficiency at which the downstream 5′ splice site (115.+1A) is recognized is largely unaffected by mutations in the preceding 3′ splice site. The fact that exon skipping does not occur with this weaker 3′ splice site suggests that this 3′ splice site strongly modulates the efficiency at which its preceding 5′ splice site (at the other end of the intron) can be recognized. Mutations in the downstream 5′ splice site that enhance the splicing efficiency of the upstream intron above its normal level (85wt/115.+1A) promote low-level exon skipping and are compensated by 3′ splice site mutations that enhance recognition of the upstream intron (85 3′U). Mutations in the upstream 3′ splice site that reduce the splicing efficiency of the upstream intron do not activate exon skipping even when the downstream 5′ splice site is impaired (85 3′C/115.+1A). We conclude again that splicing efficiencies are defined primarily via interactions across introns and only peripherally by interactions across exons.

DISCUSSION

Intron versus Exon Definition

These experiments emphasize the primary importance of interactions spanning introns (intron definition) rather than exons (exon definition) in plant pre-mRNA recognition (Figure 9). For both of our double intron constructs, it is apparent that interactions across exons can influence splice site selection in adjacent introns. In general, though, these interactions are of secondary importance compared with those established within individual introns. Most importantly, these secondary interactions across exons are observed only after the balance between normal and cryptic sites in the adjacent intron is disturbed by weakening (or inactivating) the normal site or strengthening cryptic sites in the intervening exon.

When this balance is altered, mutations that inactivate the 5′ splice site in the downstream intron appear to exert two distinct effects. In the case of introns containing multiple cryptic 5′ splice sites close to the mutant 5′ splice site (i.e., β-conglycinin intron 4/*rbcS3A* intron 1), the presence of an upstream intron enhances usage of the cryptic 5′ splice sites in a polar fashion. The polarity of this effect is especially obvious when the normal 5′ splice site in the downstream intron (*rbcS3A* intron 1) is completely inactivated by mutation of the conserved +2U-to-A or when the cryptic −57 site between these two introns is enhanced to a perfect 5′ splice site consensus (−57E). In both instances, the cryptic 5′ splice site nearest the upstream intron (−57) becomes the most dominant site. Although less obvious, usage of the cryptic −57wt site is also enhanced when the normal +1 site is weakened by mutation at the −2 position.

Comparison of the transcripts containing spliced or unspliced versions of this upstream intron indicates that one of its primary effects is to increase splicing of the downstream mutant intron. Transcripts containing the upstream exon sequences but lacking the upstream intron (115M/3A1.+2A and 115M/3A1.−57E) splice at an efficiency equivalent to their single intron counterparts (3A1.+2A and 3A1.−57E, respectively). Transcripts containing all or part (the 3′ end) of the upstream intron splice more efficiently than do transcripts lacking the upstream intron.

Comparison of the transcripts containing truncated and intact versions of the upstream intron indicates that an intact upstream intron has the ability to alter 5′ splice site selection patterns in a downstream intron. When an intact upstream intron is present, splicing is increased at the most proximal downstream 5′ splice site without altering splicing at the more distal sites (e.g., +1wt in 115wt/3A1.−57E and +106wt in 115wt/3A1.+2A). As a result, the overall splicing efficiency of the downstream intron is higher than that obtained with the comparable single intron construct, and the 5′-most cryptic site becomes dominant. In the presence of a truncated upstream intron (115 3′ss/3A1.+2A and 115 3′ss/3A1.−57E), the overall splicing efficiency of the downstream intron is increased without significantly altering the splice site selection patterns

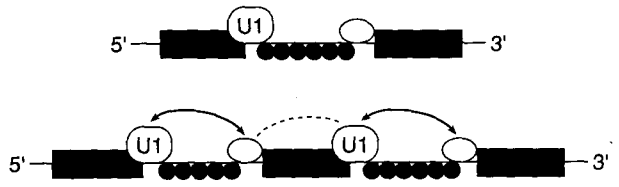

Figure 9. Intron Recognition Domains in Plant Pre-mRNA Splicing.

(Top) This model for plant intron recognition diagrams potential primary interactions involved in the excision of an intron from a single intron-containing transcript.

(Bottom) This model for plant intron recognition diagrams potential primary interactions involved in the excision of introns from multiple intron-containing transcripts.

The interactions diagrammed include those involving U1 small nuclear ribonucleoprotein (U1 ovals) with the 5′ splice site, an unspecified 3′ splice site recognition factor with the region preceding the 3′ splice site (open ovals), and putative AU-binding proteins (black circles) (Lou et al., 1993a; McCullough et al., 1993), with AU islands spread throughout the length of the intron. Additional secondary interactions between factors associated with the upstream intron and others associated with the downstream intron affect splice site recognition when the normal 5′ and 3′ splice site sequences are mutated.

from those obtained in transcripts containing only upstream exon sequences (115M). This dependence of splice site selection on the presence of a functional intron suggests that intra-intronic bridging interactions, such as those described for neural cell adhesion molecule transcripts (Cote et al., 1995), may be important in establishing priorities for these downstream 5′ splice sites.

In the case of an intron that does not contain cryptic sites close to the mutant 5′ splice site (i.e., β-conglycinin intron 3/intron 4), the presence of a functional upstream intron stimulates a low level of exon 4 skipping when splicing of the downstream intron is impaired. In this respect, the 85wt/115.+1A construct displays one of the most characteristic features of exon recognition. Even so, the level of exon skipping is substantially lower (18.4%; standard error of 5.8%) than in vertebrate transcripts containing comparable 5′ splice site mutations. For example, in the splicing of adenovirus major late transcripts in vitro or other natural transcripts in vivo, mutations in downstream 5′ splice sites block splicing completely or activate exon skipping to the near exclusion of other spliced products (Talerico and Berget, 1990). The magnitude of the difference in exon skipping between these systems might be accounted for by the fact that β-conglycinin exon 4 is substantially longer (297 nucleotides) than is the average length of vertebrate exons (175 nucleotides) (Berget, 1995) and potentially close to the limit at which exon-bridging interactions can form in mammalian systems. Nevertheless, it represents a natural plant exon whose length is not unusual among the many that have been cloned.

Another characteristic feature of exon recognition in vertebrate transcripts is that inactivation of the downstream 5′ splice site represses splicing of the upstream intron. For the adenovirus major late transcript processed in vitro, splicing of the upstream intron is inhibited at least 20-fold (Talerico and Berget, 1990). In contrast to these exon-defined vertebrate transcripts, splicing of the upstream intron in our double intron plant constructs is not inhibited by mutations in the downstream intron. When the proportions of exon skipping and intron 3 excision in the β-conglycinin intron 3/intron 4 transcript are summed, splicing of the upstream intron (β-conglycinin intron 3) even appears slightly enhanced by a mutation in the downstream 5′ splice site. Usage of the mutant downstream 5′ splice site (115.+1A) is approximately the same in all of the constructs containing variations in the upstream 3′ splice site (85wt, 85 3′U, and 85 3′C), indicating that for the most part, the introns in this transcript are recognized independently of one another. This conclusion also holds for the wild-type 115wt intron. Its splicing efficiency is maintained irrespective of the sequence of the upstream 3′ splice site (85wt versus 85 3′C). Although these data argue for independent recognition of these introns using intra-intronic interactions, the fact that exon skipping can occur in this mutant and the fact that it is abolished by strengthening the preceding 3′ splice site indicate that some degree of collaboration across the intervening exon facilitates recognition of these introns.

Uridines Enhance 3′ Splice Site Recognition

The observation that uridines in the region preceding the 3′ splice site enhance recognition of β-conglycinin intron 3 and that cytosines repress recognition of intron 3 underscores the importance of uridines rather than pyrimidine content in defining splicing efficiency. Our previous analyses of 3′ splice site recognition motifs in maize alcohol dehydrogenase 1 (*Adh1*) intron 3 and pea *rbcS3A* intron 1 (Lou et al., 1993a; Baynton et al., 1996) and studies by Carle-Urioste et al. (1994) indicate that uridines represent an essential recognition motif in the region immediately preceding the 3′ splice site. In these earlier single intron constructs, extensive U-to-A or U-to-C mutations in the region preceding these 3′ splice sites activated intronic (maize *Adh1* intron 3; Lou et al., 1993a) or exonic (pea *rbcS3A* intron 1; Baynton et al., 1996) cryptic 3′ splice sites and substantially repressed usage of the normal 3′ splice site. In this study, the more limited set of two U-to-C mutations (decreasing the uridine content to six of 18 nucleotides) that we introduced into the 3′ splice site region of β-conglycinin intron 3 reduced splicing slightly, without activating cryptic splice sites or exon skipping. Apparently, in these minimal mutations, the 3′ splice site at position −1 remains the best option for pairing with the normal 5′ splice site. The introduction of three uridines in this 3′ splice site (increasing the uridine content to 11 of 18 nucleotides) enhanced recognition of the 3′ splice site in β-conglycinin intron 3 to a point at which its use precludes

exon skipping and selection of the weaker 3′ splice site in intron 4.

Noncanonical Splice Sites Are Functional in Tobacco Nuclei

Surprisingly, noncanonical splice sites are utilized when the +1G of β-conglycinin intron 4 is mutated to A. Splicing of these noncanonical splice sites, /AU . . . AA/, /AU . . . AC/, and /UU . . . AG/, in tobacco nuclei contrasts dramatically with the behavior of 5′ splice site mutants in other systems. In vivo, mutation of the conserved /GU dinucleotide in a mammalian transcript to /AU blocks splicing at the second catalytic step and, as a result, decreases the overall efficiency of splicing (Aebi et al., 1986); the presence of an AC/ near the 3′ splice site inefficiently restores splicing between the noncanonical /AU . . . AC/ nucleotides (Scadden and Smith, 1995). In yeast (*Saccharomyces cerevisiae*), the /GU to /AU mutation causes a similar block in intron excision. Splicing of this transcript proceeds through the first cleavage reaction and accumulates lariat intermediates in vivo and in vitro without activating alternate cryptic 5′ splice sites (Newman et al., 1985; Vijayraghavan et al., 1986). The second-step block in yeast splicing created by the presence of an /AU at the 5′ end of the intron is partially suppressed (10% of the wild-type efficiency) by the presence of an AC/ at the normal 3′ splice site and less efficiently suppressed (3% of the wild-type efficiency) by the presence of an AA/ at the normal 3′ splice site (Parker and Siliciano, 1993).

Our data demonstrate that a variety of noncanonical 5′ and 3′ splice sites can efficiently complete both steps of pre-mRNA splicing in plant cells. The first two sites, /AU . . . AC/ and /AU . . . AA/, are identical to the noncanonical sites used in yeast (Parker and Siliciano, 1993); the first of these sites, /AU . . . AC/, is identical to the site used inefficiently in mammalian cells (Scadden and Smith, 1995). The last, /UU . . . AG/, represents a novel site not known to function in any other organism. That the tobacco splicing machinery pairs particular noncanonical 5′ sites with an alternate noncanonical 3′ site rather than with the canonical 3′ site indicates that specific interactions between these sites preclude pairing of the /AU with AG/ and facilitate pairing of /AU with the AA/ or AC/ dinucleotides that lie immediately upstream.

Although our clone-and-sequence strategy has not allowed us to define the frequency at which each set of noncanonical sites is used, it is clear that this collection of sites is used at a moderate (16.2%) efficiency when compared with the splicing efficiency of the wild-type intron (83%). This corresponds to a fivefold reduction in overall splicing of β-conglycinin intron 4, in comparison with the splicing of the yeast actin intron 1, which is reduced ~7.5-fold when noncanonical splice sites are present at either end of the intron and 1000-fold when a single noncanonical splice site is present (Parker and Siliciano, 1993). Further analysis indicates that many noncanonical +1 and +2 nucleotides are functional in splicing of β-conglycinin intron 4 (C.E. Baynton, J.M. Limon, S.R. Smith, and M.A. Schuler, manuscript in preparation).

Exon Skipping and Cryptic Splice Site Usage in Other Plant Transcripts

In summary, it is clear that splice site mutations in one intron can affect splicing patterns in an adjacent intron. Although single nucleotide changes in the conserved 5′ splice site sequence activate alternative splicing patterns, exact predictions as to the phenotype of these mutations remain dependent on the topology of the introns and the relative strengths of the upstream and downstream splice sites being examined. Despite the fact that alternative 5′ or 3′ splicing has been identified in a few plant transcription units (Werneke et al., 1989; Dietrich et al., 1990; Rundle and Zielinski, 1991; Kopriva et al., 1995), examples of cryptic splicing or exon skipping resulting from single or multiple point mutation(s) within an intron have not yet been recorded. The recent demonstration that a +5G-to-A in the 5′ splice site of the deetiolated *DET1* transcript inactivates splicing (Pepper et al., 1994) emphasizes the importance of individual 5′ splice site nucleotides for correct mRNA expression, but because this mutation occurs in the first intron of this transcription unit, no inferences can be drawn on the predominance of exon skipping in natural transcription units.

The most prominent examples of exon skipping in plant transcripts are those resulting from the insertion of transposable elements in the maize *bronze*, *waxy*, and *Adh1* genes (reviewed in Purugganan and Wessler, 1992; Varagona et al., 1992). In most of these cases, large transposon insertions (5.2 to 6.2 kb) activate excision of the sequences derived from the transposable element and/or skipping of one or more adjacent exons without ever completely blocking recognition of the normal 3′ splice site (Varagona et al., 1992). The technologies available now for analysis of improperly spliced transcripts within a population of correctly spliced products and for the identification of genetic defects in ethyl methanesulfonate–induced mutants will undoubtedly provide many examples of natural and induced splicing defects in plant genes.

METHODS

Plant Intron Constructs

The first double intron construct (β-conglycinin intron 4/*rbcS3A* intron 1; Figure 1) used in this study contains exon 4 (84 nucleotides), intron 4 (115 nucleotides; 72%AU), and 36 nucleotides of exon 5 derived from the soybean β-conglycinin α′ subunit gene (Schuler et al., 1982) fused via a BglII-BamHI linker with 177 nucleotides of exon 1, intron 1 (469 nucleotides; 73%AU), and exon 2 (137 nucleotides) of the pea *rbcS3A* gene (Fluhr et al., 1986). The second double intron construct (β-conglycinin intron 3/intron 4; Figure 5) extends from the 3′ end of intron 2 (37 nucleotides preceding the 3′ splice site) to the end of exon 5 in

the β-conglycinin α′ subunit gene (Schuler et al., 1982) and contains intron 3 (85 nucleotides; 67%AU) and intron 4 (115 nucleotides; 72%AU).

Mutations were introduced into these constructs by polymerase chain reaction (PCR)–mediated mutagenesis, and wild-type and mutant alleles were transferred from pBluescript II SK+ subclones into the BglII-KpnI site of the pMON458 vector, as previously described (McCullough et al., 1991). Mutant alleles are designated according to the position of their substitution (with +1 representing the first nucleotide of the intron) and the nucleotide substituted (i.e., 3A1.+1A, position +1 in *rbcS3A* intron 1 mutated to A). The 115M/3A1.+2A and 115M/3A1.−57E constructs were prepared by fusing the spliced β-conglycinin mature transcript (84 nucleotides of exon 4 and 36 nucleotides of exon 5) via a EcoRI-BamHI linker to the BamHI site at the 5′ end of the pea *rbcS3A* exon 1. The truncated intron construct was prepared by fusing the last 50 nucleotides at intron 4 and 36 nucleotides of exon 5 via a BglII-BamHI linker to the BamHI site at the 5′ end of pea *rbcS3A* gene, as described above. Recombinant pMON458 constructs were introduced into *Nicotiana benthamiana* leaf discs as outlined in McCullough et al. (1991), and leaf discs were harvested 4 days after transfection. Total RNA samples were prepared and DNase-treated as outlined in McCullough and Schuler (1993).

Analysis of Transcripts

DNase-treated total RNA samples were reverse transcribed and PCR amplified as outlined in McCullough and Schuler (1993), using 15 cycles of PCR. The processing patterns for individual introns were defined using primers specific for the exon sequences flanking the intron(s) of interest, as shown in Figures 1 and 6. PCR products were fractionated on 2% agarose gels containing 1× Tris-borate-EDTA buffer, transferred to GeneScreen (Du Pont), and probed with random hexamer ^{32}P-labeled exon-specific probes, as described in McCullough et al. (1993). The probe used for analysis of β-conglycinin intron 4 splicing (Figure 2A) in the β-conglycinin intron 4/*rbcS3A* intron 1 construct spanned the entire fragment of the β-conglycinin gene present in this construct. The probe used for quantitation of *rbcS3A* intron 1 splicing (Figure 2B) in the β-conglycinin intron 4/*rbcS3A* intron 1 construct contained *rbcS3A* exon 1–specific sequences upstream of the −57 cryptic 5′ splice site. The probe used to define the overall splicing patterns (Figure 2C) of this double intron construct contained *rbcS3A* exon 2–specific sequences. The probe used for quantitation of splicing in the β-conglycinin intron 3/intron 4 construct contained β-conglycinin exon 5–specific sequences (Figure 6).

Hybridization signals were quantified by using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Coamplification of spliced and unspliced *rbcS3A* intron 1 transcripts indicates that this assay quantitatively amplifies precursor and spliced products over a range of RNA concentrations and precursor/product ratios (McCullough and Schuler, 1993). Each reported splicing efficiency represents the average of at least three independent transfection experiments with the corresponding standard errors. The splicing efficiencies cited for each transcript are defined either as the percentage of spliced/[precursor plus spliced] transcript accumulated in each transfection (splicing efficiency) or as the percentage of spliced/[total spliced] transcript (percentage of spliced transcript). The array of spliced transcripts generated with each construct was monitored by cloning and sequencing the collection of PCR products generated with each mutant by using Sequenase 2.0 (U.S. Biochemical) and the 115 3′S, 85-5′, or 3A1 exon 2 primers positioned near each of the relevant junctions.

ACKNOWLEDGMENTS

We acknowledge Sherry Xu for diligent technical assistance. This work was supported by National Institutes of Health Grant No. RO1 GM39025 (to M.A.S.) and U.S. Department of Agriculture Competitive Research Grant No. AG92-37301-7964 (to A.J.M).

Received May 3, 1996; accepted September 25, 1996.

REFERENCES

Aebi, M., Hornig, H., Padgett, R.A., Reiser, J., and Weismann, C. (1986). Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. Cell **47,** 555–565.

Baynton, C.E., Potthoff, S.J., McCullough, A.J., and Schuler, M.A. (1996). U-rich tracts enhance 3′ splice site recognition in plant nuclei. Plant J. **10,** 703–711.

Berget, S. (1995). Exon recognition in vertebrate splicing. J. Biol. Chem. **270,** 2411–2414.

Brown, J.W.S. (1986). A catalogue of splice junction and putative branch point sequences from plant introns. Nucleic Acids Res. **14,** 9549–9559.

Carle-Urioste, J.C., Ko, C.H., Benito, M.-I., and Walbot, V. (1994). In vivo analysis of intron processing using splicing-dependent reporter gene assays. Plant Mol. Biol. **26,** 1785–1795.

Cote, J., Beaudoin, J., Tacke, R., and Chabot, B. (1995). The U1 small nuclear ribonucleoprotein/5′ splice site interaction affects U2AF^{65} binding to the downstream 3′ splice site. J. Biol. Chem. **270,** 4031–4036.

Csank, C., Taylor, F.M., and Martindale, D.W. (1990). Nuclear pre-mRNA introns: Analysis and comparison of intron sequences from *Tetrahymena thermophila* and other eukaryotes. Nucleic Acids Res. **18,** 5133–5141.

Dietrich, M.A., Prenger, J.P., and Guilfoyle, T.J. (1990). Analysis of the genes encoding the largest subunit of RNA polymerase II in *Arabidopsis* and soybean. Plant Mol. Biol. **15,** 207–223.

Fluhr, R., Moses, P., Morelli, G., Coruzzi, G., and Chua, N.-H. (1986). Expression dynamics of the pea *rbcS* multigene family and organ distribution of the transcripts. EMBO J. **5,** 2063–2071.

Goodall, G.J., and Filipowicz, W. (1989). The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. Cell **58,** 473–483.

Goodall, G.J., and Filipowicz, W. (1991). Different effects of intron nucleotide composition and secondary structure on pre-mRNA splicing in monocot and dicot plants. EMBO J. **10,** 2635–2644.

Goodall, G.J., Kiss, T., and Filipowicz, W. (1991). Nuclear RNA splicing and small nuclear RNAs and their genes in higher plants. In Oxford Surveys of Plant Molecular and Cell Biology, Vol. 7, B.J. Miflin, ed (Oxford, UK: Oxford University Press), pp. 255–296.

Hanley, B.A., and Schuler, M.A. (1988). Plant intron sequences: Evidence for distinct groups of introns. Nucleic Acids Res. **16,** 7159–7176.

Hawkins, J.D. (1988). A survey on intron and exon lengths. Nucleic Acids Res. **16,** 9893–9908.

Hoffman, B.E., and Grabowski, P.J. (1992). U1 snRNP targets an essential splicing factor, U2AF65, to the 3′ splice site by a network of interactions spanning the exon. Genes Dev. **6,** 2554–2568.

Kopriva, S., Cossu, R., and Bauwe, H. (1995). Alternative splicing results in two different transcripts for H-protein of the glycine cleavage system in the C_4 species *Flaveria trinervia.* Plant J. **8,** 435–441.

Liu, H.-X., and Filipowicz, W. (1996). Mapping of branchpoint nucleotides in mutant pre-mRNAs expressed in plant cells. Plant J. **9,** 381–389.

Lou, H., McCullough, A.J., and Schuler, M.A. (1993a). 3′ splice site selection in dicot plant nuclei is position-dependent. Mol. Cell. Biol. **13,** 4485–4493.

Lou, H., McCullough, A.J., and Schuler, M.A. (1993b). Expression of maize *Adh1* intron mutants in tobacco nuclei. Plant J. **3,** 393–403.

Luehrsen, K.R., and Walbot, V. (1994a). Intron creation and polyadenylation in maize are directed by AU-rich RNA. Genes Dev. **8,** 1117–1130.

Luehrsen, K.R., and Walbot, V. (1994b). Addition of A- and U-rich sequence increases the splicing efficiency of a deleted form of a maize intron. Plant Mol. Biol. **24,** 449–463.

McCullough, A.J., and Schuler, M.A. (1993). AU-rich intronic elements affect pre-mRNA 5′ splice site selection in *Drosophila melanogaster.* Mol. Cell. Biol. **13,** 7689–7697.

McCullough, A.J., and Schuler, M.A. (1997). Intronic and exonic sequences modulate 5′ splice site selection in plant nuclei. Nucleic Acids Res., in press.

McCullough, A.J., Lou, H., and Schuler, M.A. (1991). In vivo analysis of plant pre-mRNA splicing using an autonomously replicating vector. Nucleic Acids Res. **19,** 3001–3009.

McCullough, A.J., Lou, H., and Schuler, M.A. (1993). Factors affecting authentic 5′ splice site selection in plant nuclei. Mol. Cell. Biol. **13,** 1323–1331.

Moore, M.J., Query, C.C., and Sharp, P.A. (1993). Splicing of precursors to messenger RNAs by the spliceosome. In The RNA World, R. Gesteland and J. Atkins, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory), pp. 303–357.

Newman, A.J., Lin, R.-J., Cheng, S.-C., and Abelson, J. (1985). Molecular consequences of specific intron mutations on yeast mRNA splicing *in vivo* and *in vitro.* Cell **42,** 335–344.

Parker, R., and Siliciano, P.G. (1993). Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. Nature **361,** 660–662.

Pepper, A., Delaney, T., Washburn, T., Poole, D., and Chory, J. (1994). *DET1*, a negative regulator of light-mediated development and gene expression in Arabidopsis, encodes a novel nuclear-localized protein. Cell **78,** 109–116.

Purugganan, M., and Wessler, S.R. (1992). The splicing of transposable elements and its role in evolution. Genetica **86,** 295–303.

Robberson, B.L., Cote, G.J., and Berget, S.M. (1990). Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell. Biol. **10,** 84–94.

Roscigno, R.F., Weiner, M., and Garcia-Blanco, M.A. (1993). A mutational analysis of the polypyrimidine tract of introns: Effects of sequence differences in pyrimidine tracts on splicing. J. Biol. Chem. **268,** 11222–11229.

Rundle, S.J., and Zielinski, R.E. (1991). Organization and expression of two tandemly oriented genes encoding ribulose bisphosphate carboxylase/oxygenase activase in barley. J. Biol. Chem. **266,** 4677–4685.

Scadden, A.D.J., and Smith, C.W.J. (1995). Interactions between the terminal bases of mammalian introns are retained in inosine-containing pre-mRNAs. EMBO J. **14,** 3236–3246.

Schuler, M.A., Schmitt, E.S., and Beachy, R.N. (1982). Closely related families of genes code for the α and α′ subunits of the soybean 7S storage protein complex. Nucleic Acids Res. **10,** 8225–8244.

Sharp, P.A. (1994). Split genes and RNA splicing. Cell **77,** 805–815.

Simpson, C.G., Clark, G., Davidson, D., Smith, P., and Brown, J.W.S. (1996). Mutation of putative branchpoint consensus sequences in plant introns reduces splicing efficiency. Plant J. **9,** 369–380.

Talerico, M., and Berget, S.M. (1990). Effect of 5′ splice site mutations on splicing of the preceding intron. Mol. Cell. Biol. **10,** 6299–6305.

Varagona, M.J., Purugganan, M., and Wessler, S.R. (1992). Alternative splicing induced by insertion of retrotransposons into the maize *waxy* gene. Plant Cell **4,** 811–820.

Vijayraghavan, U., Parker, R., Tamm, J., Iimura, Y., Rossi, J., Abelson, J., and Guthrie, C. (1986). Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome. EMBO J. **5,** 1683–1695.

Werneke, J.M., Chatfield, J.M., and Ogren, W.L. (1989). Alternative mRNA splicing generates the two ribulose bisphophate carboxylase/oxygenase activase polypeptides in spinach and *Arabidopsis.* Plant Cell **1,** 815–825.

Woolford, J.L. (1989). Nuclear pre-mRNA splicing in yeast. Yeast **5,** 439–457.