The EMBO Journal vol.14 no.2 pp.377–388, 1995

Effects of secondary structure on pre-mRNA splicing: hairpins sequestering the 5′ but not the 3′ splice site inhibit intron processing in *Nicotiana plumbaginifolia*

Hong-Xiang Liu, Gregory J.Goodall[1], Ryszard Kole[2] and Witold Filipowicz[3]

Friedrich Miescher-Institut, PO Box 2543, 4002 Basel, Switzerland
[1]Present address: Institute of Medical and Veterinary Science, Adelaide, SA 5000, Australia
[2]Present address: Department of Pharmacology and Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599, USA

[3]Corresponding author

Communicated by W.Filipowicz

We have performed a systematic study of the effect of artificial hairpins on pre-mRNA splicing in protoplasts of a dicot plant, *Nicotiana plumbaginifolia*. Hairpins with a potential to form 18 or 24 bp stems strongly inhibit splicing when they sequester the 5′ splice site or are placed in the middle of short introns. However, similar 24 bp hairpins sequestering the 3′ splice site do not prevent this site from being used as an acceptor. Utilization of the stem-located 3′ site requires that the base of the stem is separated from the upstream 5′ splice site by a minimum of ~45 nucleotides and that another 'helper' 3′ splice site is present downstream of the stem. The results indicate that the spliceosome or factors associated with it may have a potential to unfold secondary structure present in the downstream portion of the intron, prior to or at the step of the 3′ splice site selection. The finding that the helper 3′ site is required for utilization of the stem-located acceptor confirms and extends previous observations, obtained with HeLa cell *in vitro* splicing systems, indicating that the 3′ splice site may be recognized at least twice during spliceosome assembly.

Key words: introns/plant gene expression/pre-mRNA processing/RNA secondary structure/splicing

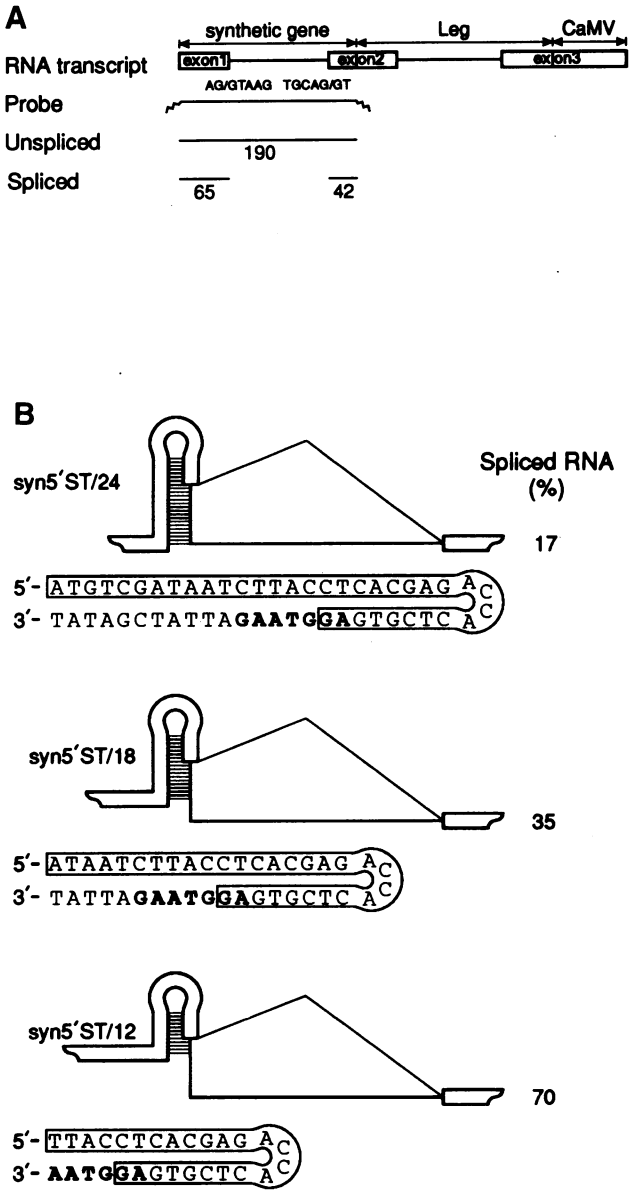
Introduction

Splicing of nuclear pre-mRNAs is mediated by the spliceosome, a large RNA–protein complex, which is assembled in a step-wise fashion from small ribonucleoprotein particles (snRNPs) U1, U2, U4/U6 and U5 and numerous protein factors in addition to substrate RNA. At an early step, U1 snRNP recognizes the 5′ splice site (5′ ss) by a base-pairing interaction and probably also contacts the branch point region of the intron. These interactions commit the pre-mRNA to splicing. U2 snRNP then binds to the commitment complex, with U2 RNA base-pairing to the branch point sequence, forming complex A or the pre-spliceosome. Upon association of U4–U5–U6 triple snRNP, complex A is converted to splicing complex B (spliceosome), which catalyses the splicing reaction (reviewed by Green, 1991; Rymond and Rosbash, 1992; Moore *et al.*, 1993).

The spliceosome is a very dynamic machine as evidenced by the many structural rearrangements in the interactions between pre-mRNA and U-snRNAs and between different U-snRNAs, which occur in the course of spliceosome assembly and function (reviewed by Moore *et al.*, 1993; Madhani and Guthrie, 1994b). Examples of such rearrangements include the interactions at the 5′ ss where early base-pairing with U1 snRNA is replaced by the interaction with U6 snRNA in complex B (Kandels-Lewis and Seraphin, 1993; Lesser and Guthrie, 1993; Sontheimer and Steitz, 1993; and references therein) and the disruption of the U4–U6 RNA complex to allow U6 sequences to interact with U2 snRNA in the catalytically competent spliceosome (Madhani and Guthrie, 1992, 1994a; and references therein). These and other conformational changes are probably assisted by protein splicing factors containing sequence motifs conserved among the DEAD/H-box family of ATP-dependent RNA helicases (reviewed by Madhani and Guthrie, 1994b; Plumpton *et al.*, 1994). However, although RNA-stimulated ATPase activity has been demonstrated for two of the five yeast DEAD/H-box PRP proteins (PRP2 and PRP16; Schwer and Guthrie, 1991; Kim *et al.*, 1992), none of these splicing factors have, to date, been shown to have an RNA helicase activity. This may be due to a stringent substrate specificity of the proteins: they could be active only when bound to specific ligands within the spliceosome.

In addition to the RNA–RNA interactions described above, secondary structure of pre-mRNA also plays an important role in splicing, contributing to splice site selection and to the efficiency of intron excision. For example, in budding yeast, long-range base pair interactions within some natural introns have a positive effect on splicing by shortening the effective distance between the branch point and the splice sites to the required minimum (Newman, 1987; Deshler and Rossi, 1991; Goguel and Rosbash, 1993). On the other hand, stable hairpins which sequester 5′ ss or branch point regions, strongly inhibit splicing in yeast, both *in vivo* and *in vitro* (Yoshimatsu and Nagawa, 1989; Goguel *et al.*, 1993). In mammalian systems, the effects of potential secondary structure on pre-mRNA splicing are usually less pronounced *in vivo* than *in vitro* (Solnick, 1985; Solnick and Lee, 1987), but secondary structure can strongly affect utilization of alternative splice sites in mammalian cells (Eperon *et al.*, 1986, 1988; Fu and Manley, 1987; Clouet d'Orval *et al.*, 1991; Libri *et al.*, 1991).

Splicing of nuclear pre-mRNAs in higher plants, although less well characterized, is similar to splicing in vertebrates (reviewed by Goodall *et al.*, 1991; Filipowicz *et al.*, 1994). Plant introns vary in length from ~70 nt to



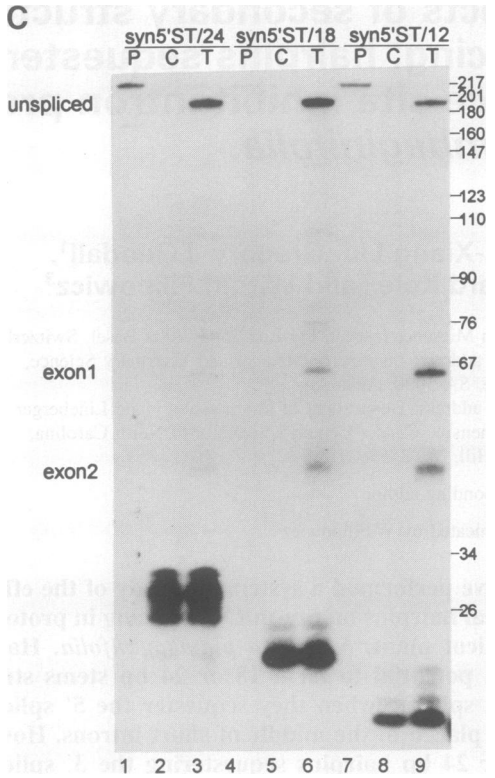


Fig. 1. Splicing of introns with sequestered 5′ ss. (**A**) Experimental system. The origins of the various parts of the recombinant gene are shown above the transcript. Leg and CaMV represent sequences originating from the soybean leghemoglobin gene and cauliflower mosaic virus, respectively. The probe used for RNase mapping and the expected protected fragments are displayed below. Probe sequences not complementary to the pre-mRNA are shown as wavy lines. (**B**) Schematic representation of the syn5′ST constructs. The sequences of potential hairpins (the 5′ ss consensus nucleotides are in bold) and splicing efficiencies are indicated. (**C**) RNase protection analysis. Constructs used for protoplast transfections are indicated at the top. Lane P, aliquot of undigested probes; lanes C and T, analyses with RNA from mock-transfected or plasmid-transfected protoplasts, respectively. Positions of size markers (pBR322 DNA digested by *Hpa*II) and bands representing unspliced RNA and spliced exons 1 and 2 are indicated.

several thousand nt and the 5′ and 3′ ss consensus sequences, AG/GTAAGT and TGCAG/GT respectively, resemble the vertebrate consensus. However, plant introns neither contain nor require the 3′ ss-proximal poly-pyrimidine tracts characteristic of most vertebrate introns (Goodall and Filipowicz, 1989, 1991); instead they are distinctly AU-rich and this property is essential for their processing (Goodall and Filipowicz, 1989, 1991; Lu *et al.*, 1993; McCullough *et al.*, 1993; Luehrson and Walbot, 1994). Like vertebrate introns, plant introns do not contain conserved branch point sequences similar to those found in yeast. However, indirect evidence of branching was obtained for some artificial introns by mutating the putative branch site, present 30 nt upstream of the 3′ ss, which in certain contexts led to a reduction in splicing efficiency (Goodall and Filipowicz, 1989). Moreover, a GT to AT mutation at the 5′ ss of the rubisco activase gene intron of *Arabidopsis* (Orozco *et al.*, 1993) or of the artificial intron (our unpublished results), results in accumulation *in vivo* of RNA which may correspond to the lariat intron-exon 2 intermediate. All spliceosomal U-snRNAs in plants resemble their vertebrate counterparts in their size and structure. They contain all of the conserved sequence motifs, including the U2 sequence which base-pairs with the branch point region, known to participate in different RNA−RNA interactions occurring during splicing in vertebrates and yeast (reviewed by Solymosy and Pollak, 1993; Filipowicz *et al.*, 1994).

We have shown previously that in transfected protoplasts of the dicot plant *Nicotiana plumbaginifolia*, as in yeast but unlike in mammalian cells, intron secondary structure can strongly inhibit pre-mRNA splicing *in vivo* (Goodall and Filipowicz, 1991). In this work, we have performed a more systematic study of the effect of artificial stem−loop structures, introduced at different intron locations, on pre-mRNA splicing in *N.plumbaginifolia*. We have found that hairpins of 18 or 24 bp have a strong inhibitory effect on splicing when they sequester the 5′ ss or are placed in the middle of short introns. Surprisingly, similar hairpins sequestering the 3′ ss, do not prevent this site from being used as an acceptor. The results suggest that productive interaction of U1 and U2 snRNAs with pre-mRNA allows

the spliceosome to unfold secondary structure present in the downstream portion of the intron.

Results

Sequestering of the 5′ ss but not of the 3′ ss in a stem strongly inhibits splicing

In order to gain more insight into how secondary structure affects splicing in plant cells, we have investigated the effect of sequestering the 5′ ss and 3′ ss sequences within hairpins on splicing of a model intron, syn7. This intron is 85 nt long, 75% A+U-rich and contains 5′ and 3′ splice sites conforming to the consensus derived for dicot plant introns (Goodall and Filipowicz, 1990; Filipowicz *et al.*, 1994). Syn7 forms part of a chimeric pre-mRNA gene in which synthetic sequences are fused to a soybean leghemoglobin gene fragment containing the natural intron (Goodall and Filipowicz, 1989, see Figure 1A). Splicing of synthetic introns was analysed and quantified by RNase A/T1 mapping using RNA isolated from transfected *N.plumbaginifolia* protoplasts and ^{32}P-labelled RNA probes complementary to the unspliced form of RNA. In this assay, unspliced RNA yields the protected fragment of 190 nt and spliced RNA generates fragments of ~65 and 42 nt, corresponding to exon 1 and exon 2, respectively (see Figure 1A).

In constructs syn5′ST/24, syn5′ST/18 and syn5′ST/12, sequences in exon 1 were altered such that resulting pre-mRNAs can form hairpins which sequester the 5′ ss in stems of 24, 18 or 12 bp, respectively (Figure 1B). When the 5′ ss is sequestered in the 24 bp stem, the synthetic intron is spliced with an efficiency of only 17% (Figure 1C); under similar conditions, the control syn7 RNA, which has no potential to form a hairpin, is spliced at 85% (Goodall and Filipowicz, 1989, 1991; see also Figure 2A). However, when the stem was shorter, the efficiency of processing increased: the syn5′ST/18 and syn5′ST/12 introns were spliced with efficiencies of 35 and 70%, respectively (Figure 1B and C). This indicates that the observed inhibition of splicing is probably due to the formation of secondary structure (see also below). Accurate splice site selection in syn5′ST RNAs was confirmed by RT-PCR analysis and sequencing of cloned products (data not shown).

The construct syn3′ST contains an inverted palindrome with the potential to form a 24 bp stem which sequesters the 3′ ss TGCAG/GT (Figure 2B). Mapping of syn3′ST-specific RNA revealed that it is spliced with 87% efficiency, as well as syn7 RNA. However, the exon 1 and exon 2 protected fragments were not present in equimolar amounts (the exon 2 fragment of ~42 nt is under-represented; Figure 2A, compare lanes 3 and 6), suggesting that in addition to the consensus 3′ ss located in the stem, another site is also selected as an acceptor. Using an additional probe, extending further downstream into exon 2, it was found that 19% of transcripts are spliced to the alternative 3′ ss, TGCAG/Gg (the nucleotides matching the 3′ ss consensus, TGCAG/GT, are shown in upper case), that occurs fortuitously 6 nt downstream of the hairpin, whereas 67% are spliced to the stem-located 3′ ss (Figure 2A, lane 9; Figure 2B). Selection of these two 3′ sites was confirmed by RT-PCR analysis and sequencing of the cloned PCR products (data not shown).

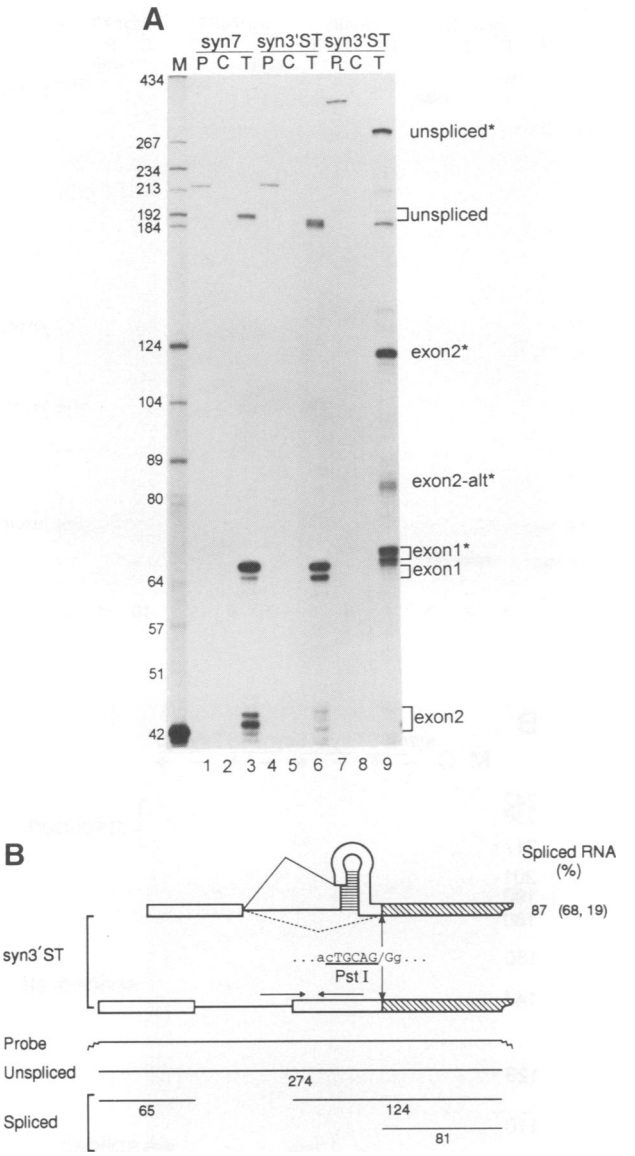
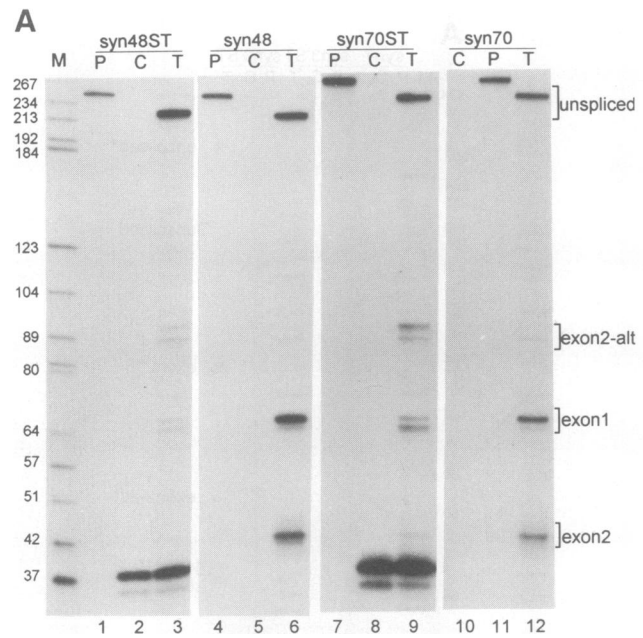
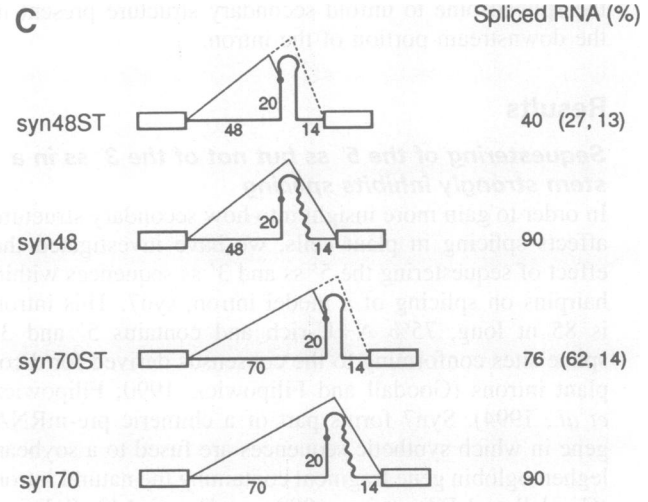


Fig. 2. Analysis of syn3′ST RNA splicing. (**A**) RNase protection analysis. Lanes 1–6, mapping of syn7 and syn3′ST RNAs with complementary probes similar to those shown in Figure 1A. Lanes 7–9, mapping of syn3′ST RNA with a longer probe (P_L; fragments protected by this probe are marked with asterisks) extending into the leghemoglobin gene sequence (see B). The difference in length of exon 1-specific fragments protected by the two probes is due to differences in the 3′-terminal sequence of the probes. Lanes P and P_L, undigested probes; lanes C and T, mapping of RNA from control and transfected protoplasts, respectively. Lane M, size markers (pBR322 DNA digested with *Hae*III, supplemented with the end-labelled 42 nt oligodeoxynucleotide). (**B**) Schematic representation of syn3′ST and its mapping with the probe P_L. Two drawings of syn3′ST are shown to facilitate the display of results and mapping. The sequence of the hairpin is: TATGATATCATGCAGGTACGAGCGttttCGCTCGTAC-CTGCATGATATCATA (inverted palindromes are in capital letters and the stem-located 3′ ss is underlined; the *Pst*I site shown in the scheme immediately follows the descending arm of the stem. The hatched box represents exon sequences originating from the leghemoglobin gene. The sequence around the alternative 3′ ss is shown with the *Pst*I site underlined. Nucleotides matching the 3′ ss consensus are shown in upper case. The splicing efficiency is indicated on the right. Numbers in the bracket represent efficiency of splicing to the stem-located 3′ ss and the alternative 3′ ss (indicated by a broken line), respectively.





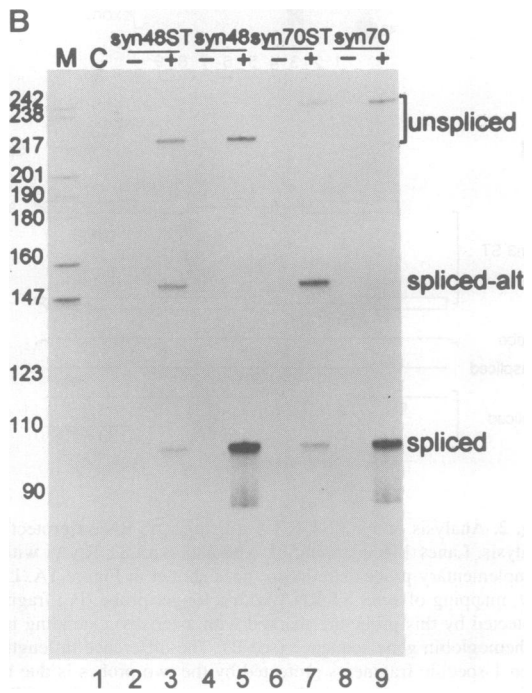


Fig. 3. Utilization of the 3′ ss located in the stem of intra-intronic hairpins. (**A**) RNase protection analysis. The ~90 nt fragment representing RNA spliced to the 3′ ss located in the stem, is marked as exon 2-alt. For other details, see legend to Figure 1C. (**B**) RT-PCR analysis. RNA isolated from either control (lane 1) or transfected (lanes 2–9) protoplasts was subjected to RT-PCR analysis, either in the absence (−) or in the presence (+) of reverse transcriptase, using exon 1 and exon 2 specific primers. Positions of amplified DNA fragments representing unspliced RNA (218 bp for syn48 and syn48ST and 240 bp for syn70 and syn70ST) and RNAs spliced to the downstream consensus 3′ ss (104 bp) and the alternative 3′ ss in the stem (spliced-alt; 150 bp) are indicated. (**C**) Schematic representation of constructs syn48ST, syn48, syn70ST and syn70. The sequence of the potential hairpin is AAAATTATGGAAAATGTTAGCCAAttttTT-GGCTAACATTTTCCATGGTTTT (inverted palindromes are in capital letters and the stem-located AG, indicated by the black dot in schemes, is underlined). Distances between the 5′ and 3′ splice sites and the base of the stem are indicated below the intron. Wavy lines in constructs syn48 and syn70 represent sequences which are not complementary to the ascending arm of the stem. Splicing efficiencies are shown on the right, with the first number in the bracket indicating splicing to the stem-located 3′ ss and the second number representing splicing to the downstream consensus 3′ ss (indicated by broken lines in syn48ST and syn70ST). The splicing efficiency of syn48ST was previously calculated as 10 rather than 40% (Goodall and Filipowicz, 1991); the reason for the discrepancy is not clear. Calculated ΔG value for the hairpin present in syn48ST and syn70ST is −27.5 kcal/mol. For hairpins present in constructs syn5′ST, syn5′ST/18, syn3′ST, syn44ST-16 and syn40ST/18 (see below), the values are −31.7, −22.8, −36,4, −31.6 and −22.9 kcal/mol, respectively.

Taken together, these results indicate that, in contrast to the 5′ ss, the 3′ ss located in the potential stem can be efficiently selected for splicing.

Preferential usage of the 3′ ss located within the intra-intronic double stranded stem

Preferential utilization of the 3′ ss located in the stem was also observed with other derivatives of the syn7 intron: introns syn48ST and syn70ST (Figure 3). Syn48ST and syn70ST have the potential to form an intra-intronic 24 bp hairpin which contains a poor quality 3′ ss gttAG/cc located in the stem (the AG is located at positions 19/20 on the ascending arm of the stem). The distance from the 5′ ss to the base of the stem is 48 and 70 nt in syn48ST and syn70ST, respectively. The consensus 3′ ss, TGCAG/GT, is positioned 14 nt downstream of the stem (Figure 3C).

RNase A/T1 mapping of syn48ST and syn70ST RNAs (Figure 3A; see also Figure 5A) yielded protected RNA fragments representing the spliced exon 1 (~65 nt band) and two alternative forms of exon 2 (~88 and ~42 nt bands), diagnostic of RNAs spliced to the stem-located gttAG/cc and to the downstream TGCAG/GT sites, respectively. The splicing efficiency of syn48ST was 40% and that of syn70ST 76%; for either construct the 3′ ss located within a stem was utilized in preference to the

downstream site (Figure 3A and C). That the 88 and 42 nt exon 2-specific fragments indeed originate from RNAs spliced to the stem-located and the downstream sites respectively, was confirmed by S1 nuclease protection assays using as probes the oligodeoxynucleotides which are complementary to splice junctions formed between the 5′ ss and each of the two 3′ sites (data not shown) and by cloning and sequence analysis of the RT-PCR products (Figure 3B and data not shown). The 150 and 104 bp RT-PCR products (Figure 3B) were found to correspond to RNAs spliced to the stem-located 3′ ss gttAG/cc and to the downstream 3′ ss TGCAG/AG, respectively. (Results of RT-PCR experiments are generally in agreement with the RNase protection data although RT-PCR seemed to overestimate the percentage of spliced RNA. This could be due to the hairpins, present in precursor RNAs but not in spliced RNAs, affecting the reverse transcription or PCR-amplification steps, or because of more efficient amplification of shorter spliced RNAs.)

Short stretches of U residues located within introns contribute significantly to intron recognition in plants (M.Gniadkowski and W.Filipowicz, unpublished data). To eliminate the possibility that the hairpin loop sequence, TTTT, present in introns syn48ST or syn70ST has an effect on their processing, we analysed splicing of a derivative of syn48ST in which this sequence is changed to ACCA. The efficiency and pattern of processing of the modified intron was identical to that of syn48ST. We have also assayed processing of a syn48ST variant in which the stem-located 3′ ss, gttAG/cc, is present not in the ascending but in the descending arm of the stem. In this case too, the hairpin site was utilized for splicing (data not shown).

We next examined whether the upstream 3′ ss of syn48ST and syn70ST introns would still be selected if it were no longer sequestered in a hairpin. Introns syn48 and syn70 represent modified versions of syn48ST and syn70ST in which the descending arm of the palindrome was modified such that its sequence is no longer complementary to the ascending arm (Figure 3C). The RNase mapping (Figure 3A) and RT-PCR (Figure 3B) analyses revealed that splicing of syn48 and syn70 occurred only to the consensus downstream 3′ ss TGCAG/GT, with the intronic AG not being used. Hence, a sequestration of the upstream 3′ ss into a hairpin not only did not prevent its utilization as an acceptor but was in fact required for this event to take place.

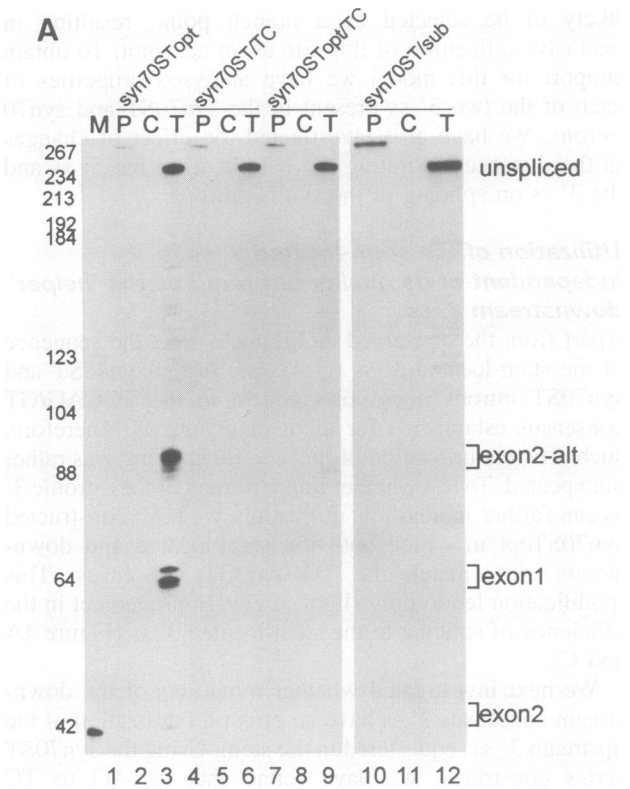
The results presented thus far may be explained by the following model. The upstream 3′ site in constructs syn48ST and syn70ST and in syn3′ST, remains sequestered in the hairpin at the onset of spliceosome assembly. Since in each of these constructs the single-stranded sequence separating the hairpin from the downstream 3′ ss (14 nt in syn48ST and syn70ST, 6 nt in syn3′ST) is probably too short to allow U2 snRNP binding, a region upstream of the hairpin is chosen as a branch point. During subsequent steps of spliceosome aseembly, possibly during scanning for the closest AG positioned downstream of the selected branch point (see Discussion), the hairpin would be unfolded and the stem-located AG used for splicing. In the absence of the hairpin, in constructs syn48 and syn70, a sequence upstream of the consensus 3′ ss is likely to be selected as a branch point, resulting in exclusive utilization of this site as an acceptor. To obtain support for this model we have analysed properties of each of the two 3′ ss present in the syn70ST and syn70 introns. We have also investigated the effect of changes in the distance separating the hairpin from the 5′ ss and the 3′ ss on splicing of the synST introns.

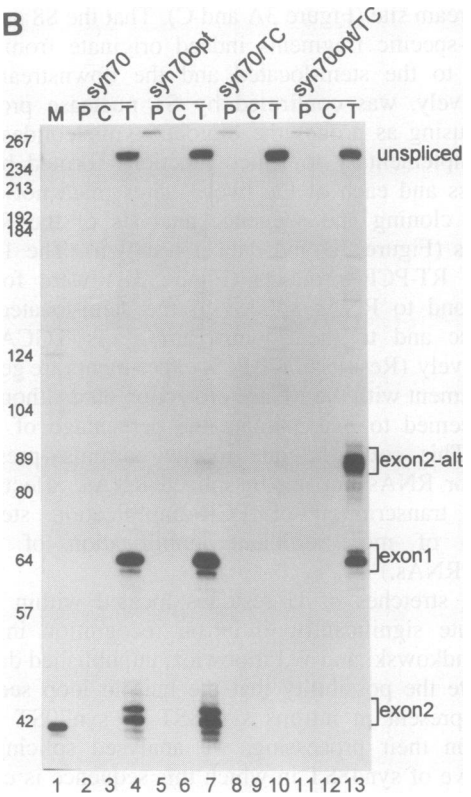
Utilization of the stem-located 3′ ss is independent of its quality but requires the 'helper' downstream 3′ ss

Apart from the conserved AG dinucleotide, the sequence of the stem-located 3′ ss (gttAG/cc) in the syn48ST and syn70ST introns does not conform to the TGCAG/GT consensus established for dicot plant introns. Therefore, such efficient utilization of this site for splicing was rather unexpected. To test whether improvement of the intronic 3′ ss can further increase its utilization, we have constructed syn70STopt in which both the stem-located and downstream sites match the TGCAG/GT consensus. This modification led to only slight, if any, improvement in the efficiency of splicing to the stem-located 3′ ss (Figure 4A and C).

We next investigated whether mutations of the downstream consensus 3′ ss have an effect on utilization of the upstream 3′ ss sequestered in the stem. Using the syn70ST series constructs, we have found that an AG to TC mutation in the TGCAG/GT consensus 3′ ss completely eliminates splicing to both the mutated downstream site and to the stem-located site, irrespective of whether the stem-located 3′ ss is suboptimal or optimal (Figure 4A and C, introns syn70ST/TC and syn70STopt/TC). The AG to TC mutation in the downstream 3′ ss also completely inactivated splicing in the context of the syn48ST intron (data not shown). Surprisingly, mutation of the consensus site of syn70ST to the suboptimal TGtAG/cc also eliminated splicing to either of the two sites (Figure 4A and C, intron syn70ST/sub). These data indicate that the downstream site is essential for utilization of the stem-located AG and that the 'helper' function can only be provided by the more optimal 3′ ss.

The sequence requirements of the stem-sequestered 3′ ss differ from the requirements of the same site present in the region which has no potential to form a hairpin. In the context of the 'single-stranded' syn70 intron which, like syn70ST, contains the consensus downstream site, the intronic site is not used irrespective of whether its sequence corresponds to gttAG/cc (Figure 4B and C, intron syn70; see also Figure 3) or TGCAG/GT (Figure 4B and C, intron syn70opt). However, when the downstream TGCAG/GT 3′ ss is inactivated by the AG to TC mutation, the intronic site becomes utilized, but only when it conforms to the consensus and not when its sequence is gttAG/cc (Figure 4B and C, compare constructs syn70opt/TC and syn70/TC). These results demonstrate that the intronic upstream site has different properties in the syn70 and syn70ST series constructs. In single-stranded syn70 introns, its usage is prevented by the presence of the downstream 3′ ss (this is consistent with the findings that splice sites located at transition points from AU- to GC-rich sequences are preferentially selected for splicing in plants; see Discussion). In contrast, in syn70ST introns the downstream 3′ ss is required for the intron site to be utilized.





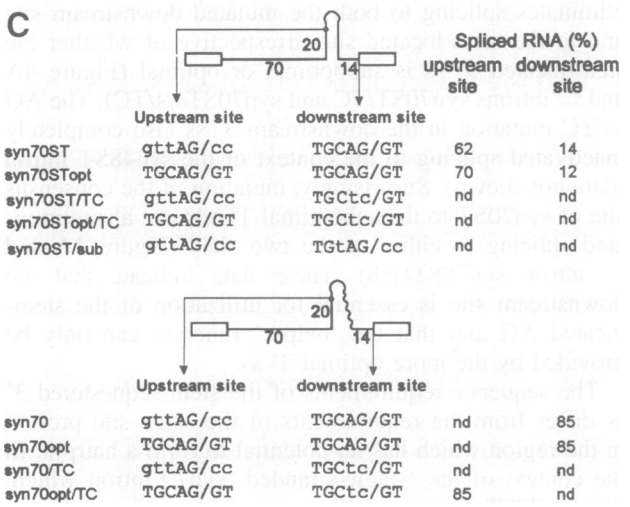


Fig. 4. Effect of 3′ ss mutations on processing of the syn70ST and syn70 constructs. (**A**) and (**B**) RNase protection analysis. For the sizes of the protected fragments and other details, see legends to Figures 1 and 3. (**C**) Schematic representation of different 3′ ss mutants. Sequences of both the upstream 3′ ss and the downstream 3′ ss are indicated; positions matching the 3′ ss consensus are in upper case. For other details, see legend to Figure 3C.

The most plausible explanation for the intronic site being inactive in syn70STopt/TC is its sequestration by the secondary structure.

Use of the stem-located 3′ ss requires a minimum distance between the 5′ ss and the base of the stem

We have previously found that intron syn19ST, which differs from introns syn48ST and syn70ST only with respect to the length of the single-stranded region separating the 5′ ss and the stem, does not undergo processing in transfected protoplasts of *N.plumbaginifolia* (Goodall and Filipowicz, 1991). Differences in the efficiency of splicing of intron syn19ST and between introns syn48ST and syn70ST (see above), suggested that sufficient separation of the 5′ ss from the base of the stem is important for intron processing, possibly to allow the positioning of U1 and U2 snRNPs. We have therefore determined the efficiency of the synST intron processing as a function of this distance (Figure 5A and B). When the 5′ ss was separated from the stem by 19 or 36 nt, no splicing occurred and when this distance was 40 nt splicing was very inefficient. However, further increases in the 5′ ss to the base-of-stem distance resulted in a progressively stronger activation of splicing with the construct syn56ST being processed at a level similar to that of syn70ST. In introns syn44ST through syn70ST, splicing to the stem-located 3′ ss was always more efficient than splicing to the consensus downstream 3′ ss (Figure 5A and B).

Processing to the cryptic 3′ ss in syn44ST resulted in excision of an intron of 64 nt, which is below the minimum length required for processing of other synthetic introns

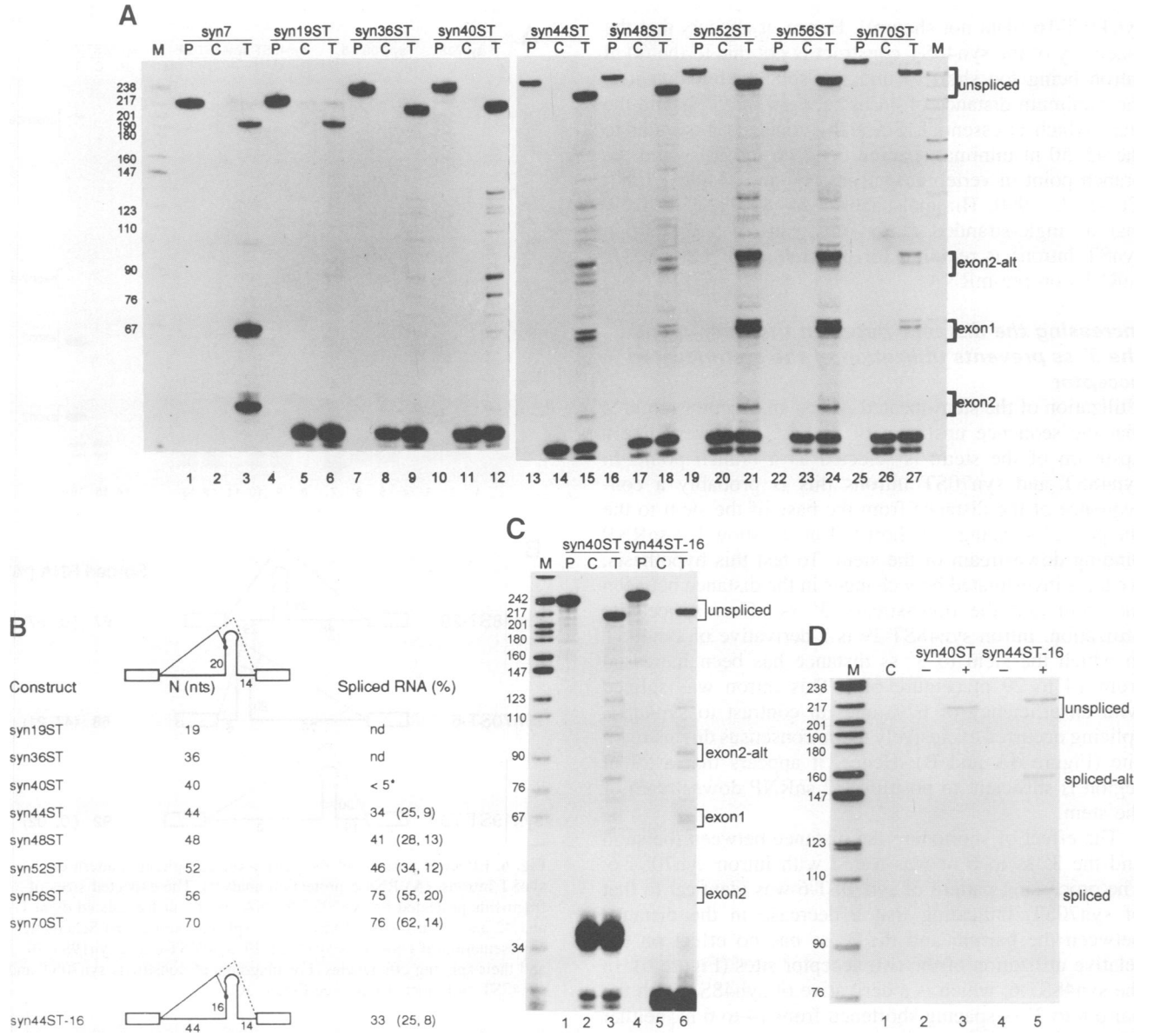


Fig. 5. Use of the stem-located 3′ ss requires a minimum distance between the 5′ ss and the stem. (**A**) RNase protection analysis. For details, see legends to Figures 1C and 3A. (**B**) Scheme of constructs analysed in (A), (C) and (D). N represents the length of the distance separating the 5′ ss from the base of the stem. Splicing efficiencies are indicated on the right. nd, splicing not detectable. *The background bands seen in the mapping of syn40ST (see A) are due to either probe degradation or incomplete RNase digestion. Splicing efficiency of syn40ST was too low to determine the relative utilization of the two 3′ splice sites by RNase mapping; based on the results of RT-PCR (D), residual splicing occurred mainly to the downstream 3′ ss. For other details, see legend to Figure 3C. (**C**) RNase mapping of syn40ST and syn44ST-16 RNAs. The expected protected fragments are: for unspliced RNA, 213 and 217 nt for syn40ST and syn44ST-16, respectively; for spliced exon 1, 65 nt; for exon 2 spliced at the consensus 3′ ss, 42 nt; for exon 2 spliced at the stem-located 3′ ss, 88 and 92 nt for syn40ST and syn44ST-16 respectively. (**D**) RT-PCR analysis of syn40ST and syn44ST-16 RNAs (lanes 2–5), performed either in the absence (−) or in the presence (+) of reverse transcriptase. Lane 1, RT-PCR performed with RNA from control protoplasts. For syn44ST-16 the expected sizes of the amplified DNA fragments are 104 and 154 bp for RNAs spliced to the consensus and the stem-located 3′ splice sites, respectively.

tested in plant protoplasts (Goodall and Filipowicz, 1990) and at the limit of natural plant intron length (the shortest known plant intron is 64 nt; reviewed by Filipowicz *et al.*, 1994). On the other hand, the 60 nt intron was not efficiently spliced from the syn40ST RNA (Figure 5A and C; see also below). In order to determine whether it is the total length of the intron or the distance between the 5′ ss and the stem which makes processing of syn40ST so inefficient, we prepared the construct syn44ST-16, a derivative of syn44ST. The syn44ST-16 intron also has a potential to form the 24 bp hairpin, but the stem-located gttAG/cc site is moved 4 nt upstream, towards the base of the stem (Figure 5B, bottom scheme). As a result, the 60 nt distance between the 5′ ss and the stem-located 3′ ss in syn44ST-16 is identical to that in syn40ST, but the two constructs differ in the length of the 'single-stranded' region separating the 5′ ss from the base of the stem. Analysis of processing of syn40ST and syn44ST-16 by RNase A/T1 mapping (Figure 5C) and by RT-PCR (Figure 5D) has demonstrated that only RNA from the latter construct was relatively efficiently spliced and that the pattern of its processing was similar to that of syn44ST [see Figure 5B; RT-PCR analysis has also revealed that the splicing pattern of syn44ST is similar to that of

syn44ST-16 (data not shown)]. Hence, it appears that the inactivity of the syn40ST construct is not due to the 60 nt intron being too short to undergo splicing but that it is the minimum distance of 44 nt between the 5′ ss and the stem which is essential. Since this distance is similar to the 45–50 nt minimal spacing between the 5′ ss and the branch point in vertebrate introns (Fu and Manley, 1987; Ge *et al.*, 1990; Himmelspach *et al.*, 1991), it is likely that a single-stranded region upstream of the stem in synST introns is required for positioning of U1 and U2 snRNPs on pre-mRNA.

Increasing the distance between the hairpin and the 3′ ss prevents utilization of the stem-located acceptor

Utilization of the stem-located AG as an acceptor requires that the sequence upstream of this AG, and very likely upstream of the stem, is selected as a branch point. In syn48ST and syn70ST introns this is probably a consequence of the distance from the base of the stem to the 'helper' 3′ ss being too short (14 nt) to allow U2 snRNP binding downstream of the stem. To test this hypothesis, we have investigated how changes in the distance between the stem and the downstream 3′ ss affect splice site utilization. Intron syn48ST-29 is a derivative of syn48ST in which the stem to 3′ ss distance has been increased from 14 to 29 nt (Figure 6B). This intron was spliced with an efficiency of 67% and, in contrast to syn48ST, splicing occurred exclusively to the consensus downstream site (Figure 6A and B). Hence, it appears that a 29 nt region is sufficient to position U2 snRNP downstream of the stem.

The effect of shortening the distance between the stem and the 3′ ss to 6 nt was tested with intron syn70ST-6. The processing pattern of syn70ST-6 was identical to that of syn70ST, indicating that a decrease in the distance between the hairpin and the 3′ ss has no effect on the relative utilization of the two acceptor sites (Figure 6). In the syn48ST-6, which is a derivative of syn48ST with the hairpin to 3′ ss spacing shortened from 14 to 6 nt, neither the stem-located nor the downstream 3′ ss was used (data not shown). The effective single-stranded length of this intron is probably too short to allow its processing.

We have also tested splicing of syn19ST-70 which contains a 24 bp hairpin 19 nt downstream from the 5′ ss (Figure 6B). This intron was efficiently spliced and as expected, only the downstream consensus 3′ ss was utilized (Figure 6A). Taken together, the efficient processing of syn48ST-29 and syn19ST-70 suggest that the presence of the stem between the 5′ ss and the region likely to be utilized as a branch point is not inhibitory to splicing. This is reminiscent of the situation in yeast, where hairpins located between the 5′ ss and the branch point do not inhibit splicing. On the contrary, secondary structure in this region may facilitate splicing by shortening the effective length of the intron to the required minimum (Newman, 1987; Goguel and Rosbash, 1993).

Inhibition of splicing by internal hairpins in syn36ST and syn40ST introns depends on the length of the stem

We have sought additional evidence that the effects of inverted palindromes present in the middle of the short

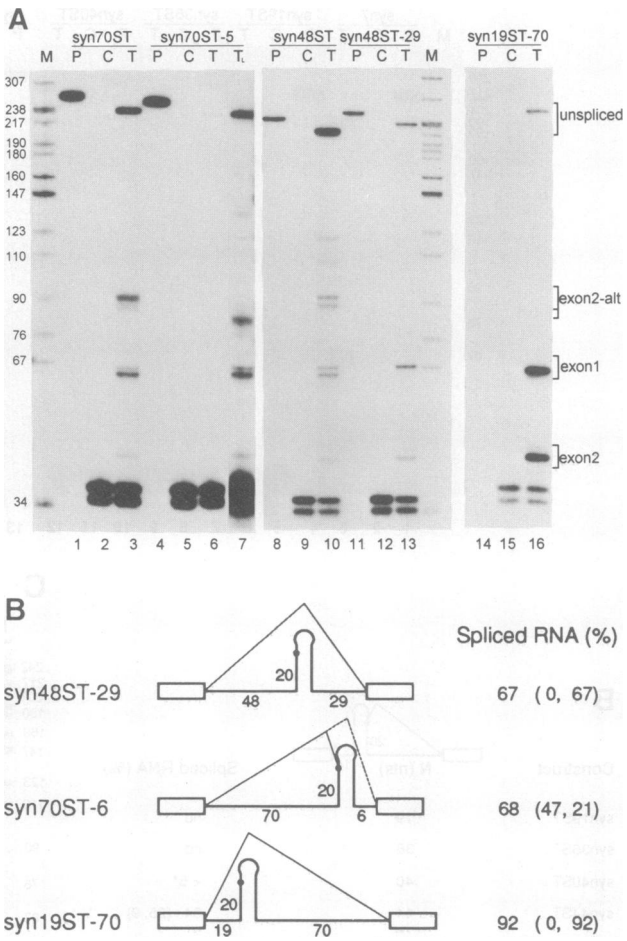


Fig. 6. Effect of position of the hairpin on the splicing pattern of synST introns. (**A**) RNase protection analysis. The expected sizes of fragments protected by syn70ST-6 RNA are: 65 nt for spliced exon 1 and 42 and 79 nt for two forms of the spliced exon 2. (**B**) Schematic representation of constructs syn48ST-29, syn70ST-6 and syn19ST-70 and their splicing efficiencies. For diagrams of constructs syn70ST and syn48ST and other details, see Figure 3C.

intron are indeed due to the secondary structure formed in that region. To this end the variants of introns syn36ST and syn40ST having the potential to form hairpin structures with 18 and 12 bp long stems were constructed (Figure 7B). It was found that introns syn36ST/18 and syn40ST/18, having a potential to form an 18 bp stem, were inactive but shortening of the stem to 12 bp allowed the introns to be efficiently spliced (Figure 7). It is possible that the 12 bp stem in these introns, as in the case of the syn5′ST/12 intron (see Figure 1) is too weak to be maintained *in vivo*. As a result, the effective 'single-stranded' length of syn36ST/12 and syn40ST/12 introns becomes sufficient to allow their processing (in the absence of secondary structure the lengths of these introns are 78 and 82 nt, respectively).

Discussion

The results presented in this work confirm and extend our previous findings (Goodall and Filipowicz, 1991) indicating that pre-mRNA secondary structure can have a profound effect on splicing in dicot plants. Inclusion into the pre-mRNA of hairpins with a potential to sequester

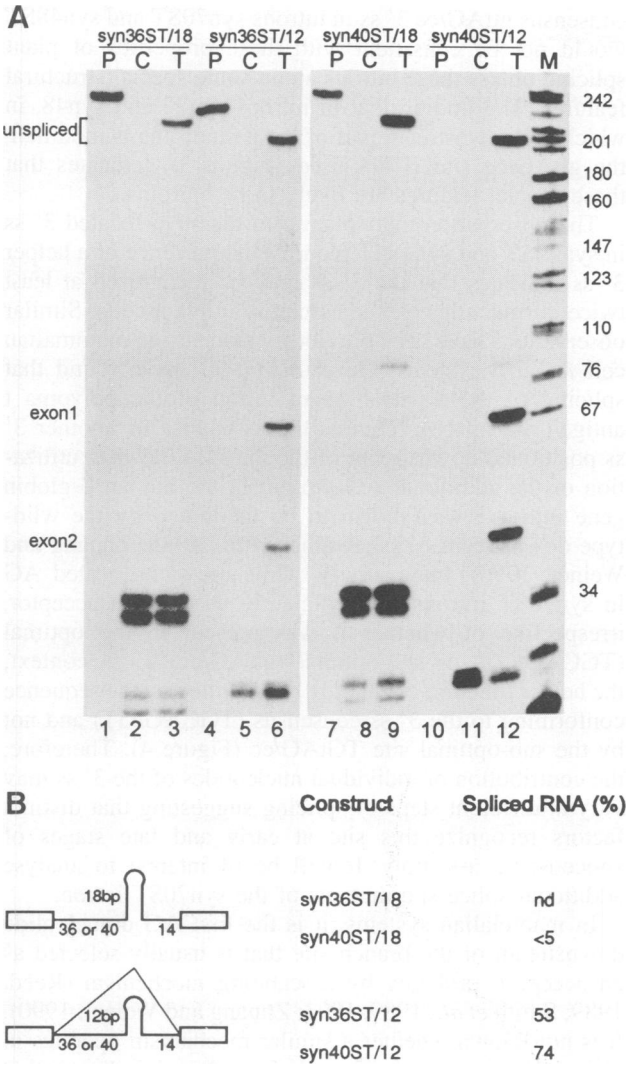


Fig. 7. Inhibition of splicing by internal hairpins in syn36ST and syn40ST introns depends on the length of the stem. (**A**) RNase protection analysis. (**B**) Schematic representation of constructs used for transfections. Sequences of the hairpins are: ATGGAAAATGTTA-GCCAAttttTTGGCTAACATTTTCCAT for 18 bp stem and AATGTTAGCCAAttttTTGGCTAACATT for 12 bp stem. For other details, see Figures 1C and 3C.

the 5′ ss or to shorten the length of single stranded intron sequence below the required minimum results in a strong inhibition of splicing in *N.plumbaginifolia* protoplasts. Inhibition is dependent on the length of double-stranded stems and hairpins with different sequences and positioned at different locations are also inhibitory. These observations indicate that the inhibition is indeed due to secondary structure formation rather than general inhibition of splicing by particular sequences. The finding that the same 24 bp hairpin inhibits splicing of some introns (e.g. syn19ST through syn40ST) but not of others (e.g. syn56ST, syn70ST or syn19ST-70) also makes the latter possibility unlikely.

Secondary structure effects on pre-mRNA splicing in *N.plumbaginifolia* resemble those reported for the yeast, *Saccharomyces cerevisiae* (Yoshimatsu and Nagawa, 1989; Goguel *et al.*, 1993). Goguel *et al.* (1993) have found that hairpins which sequester the 5′ ss or the branch point region have a strong inhibitory effect on splicing both *in vivo* and *in vitro*. Analysis of splicing complexes formed *in vitro* revealed that 5′ ss hairpins inhibit U1 snRNP binding and prevent formation of the early commitment complex, CC1. Sequestration of the branch point allows formation of the CC1 complex, but interferes with the following steps which involve addition of U2 snRNP and formation of the pre-spliceosome. The differences in effects of secondary structure on *in vivo* splicing observed between mammals (see Introduction) and yeast and plants may be due to the abundance of heterogeneous nuclear RNP (hnRNP) proteins in mammalian cells; the binding of these proteins to nascent pre-mRNAs may prevent formation of secondary structure (Dreyfuss *et al.*, 1993; Eperon *et al.*, 1988). Homologues of hnRNP proteins or structures similar to the mammalian hnRNP particles have not yet been characterized in either yeast or plants (reviewed by Dreyfuss *et al.*, 1993; Filipowicz *et al.*, 1994)

The most important finding of this work is that sequestration of a 3′ ss within a double-stranded stem does not prevent its use as an acceptor. On the contrary, for some introns, sequestration of the 3′ ss into a stem sequence is required for its utilization (Figure 3). In view of the evidence that hairpins are readily formed on pre-mRNAs expressed in plant protoplasts (see also below) and taking into account that 3′ ss hairpins in syn3′ST or syn70ST and syn48ST introns have stabilities similar to the internal or 5′ ss hairpins which strongly inhibit splicing (e.g. hairpins in introns syn5′ST/24, syn5′ST/18, syn40ST or syn40/18; for predicted ΔG values, see legend to Figure 3), two possible explanations of these results can be offered. The first is that the AG dinucleotide can act as the splicing acceptor while present within the double-stranded helix. Although not formally excluded, this possibility is unlikely. Prior to and/or during the second chemical step of splicing, the bases which are immediately adjacent to the 3′ ss junction interact with the intron 5′-terminal guanosine (Parker and Siliciano, 1993) and nucleotides in the conserved loop of U5 snRNA (Sontheimer and Steitz, 1993). Since these interactions generally occur on pre-mRNA substrates which are devoid of elaborate secondary structure, it is very improbable that they could also take place effectively with the 3′ ss being sequestered within the double-stranded stem. We favour another explanation, according to which the 3′ ss hairpin undergoes unfolding in the course of spliceosome assembly either prior to, or at the step of, 3′ ss selection. Results of this work are consistent with such a scenario and imply that the spliceosome or factors associated with it have a potential to unfold pre-mRNA secondary structure.

We have demonstrated that splicing to the stem-located AG requires that the base of the stem is separated from the upstream 5′ ss by a minimum of 44 nt (Figure 5) and that another 'helper' 3′ ss is present downstream of the stem (Figure 4). The minimal 5′ ss to the base-of-stem distance required for splicing of synST introns is similiar to the 45–55 nt minimum which must separate the 5′ ss and the branch point in mammalian introns in order to allow simultaneous binding of U1 and U2 snRNPs (Fu and Manley, 1987; Ge *et al.*, 1990; Himmelspach *et al.*, 1991; and references therein). In view of the similar architecture of mammalian and plant introns (see Introduction), this might indicate that the single-stranded region upstream of the stem is similarly required for a productive

interaction of U1 snRNP at the 5′ ss and U2 snRNP at the branch point of synST pre-mRNAs. Positions of utilized branch points have not yet been mapped in plant introns. However, utilization of the stem-located AG as an acceptor obviously requires that the sequence upstream of this AG and very likely upstream of the stem, is selected as a branch point. This conclusion is supported by the finding that increasing the stem to the 3′ ss distance from 14 to 29 nt in the syn48ST-29 intron eliminates the utilization of the stem-located 3′ ss (Figure 6). The 29 nt 'single-stranded' region is probably sufficient to allow U2 snRNP binding downstream of the stem. Taken together, these results suggest that productive interaction of U1 and U2 snRNPs with pre-mRNA in a region upstream of the hairpin may be one of the prerequisites for its subsequent unfolding.

It is probable that the downstream helper 3′ ss functions early in the assembly of the spliceosome on the syn70ST and syn48ST introns, either by supporting formation of the commitment-type spliceosomal complex or by assisting U2 snRNP in selection of the branch point in a region upstream of the hairpin. Plant introns neither contain nor require conserved branch point sequences or distinct 3′ ss-proximal polypyrimidine tracts (Goodall and Filipowicz, 1989; Filipowicz *et al.*, 1994), which are important for binding of the U2 snRNP to yeast and mammalian introns, respectively (reviewed by Green, 1991; Rymond and Rosbash, 1992). In plants, perhaps the 3′ ss itself plays a major role in the positioning of U2 snRNP at the branch point. Such a role would explain why, compared with vertebrates, more conservation is found at positions −5, −4 and +2 in the plant 3′ ss consensus, TGCAG/GT (Filipowicz *et al.*, 1994). In mammalian introns which lack long polypyrimidine tracts and in some yeast introns (collectively referred to as AG-dependent introns), the 3′ ss plays an essential role in early steps of spliceosome assembly (Reed, 1989; Reich *et al.*, 1992; and references therein).

The helper 3′ ss is required for processing of syn70ST introns irrespective of whether the stem-located upstream 3′ ss conforms to the consensus (TGCAG/GT) or is suboptimal (gttAG/cc; Figure 4). When the syn70ST intron was modified, such that its sequence no longer had the potential to form a hairpin, the upstream site in the consensus form (but not in the sub-optimal, gttAG/cc, form) continued to be efficiently selected for splicing. However, its utilization no longer required and in fact was inhibited by, the downstream helper site (Figure 4). Taken together, these results provide strong additional evidence that the upstream 3′ ss in syn70ST introns is indeed sequestered within a hairpin and as a result, is not available for interactions with spliceosomal factors during early steps of spliceosome assembly. The inability of the upstream TGCAG/GT to act as a splicing acceptor in the single-stranded syn70opt intron may be explained by it being 'buried' in the middle of AU-rich intron sequence. In plants, splice sites located at transition points between AU-rich intron and GC-rich exon sequences are usually selected for splicing, while the sites located within the AU-rich intron sequence are ignored, particularly when functional sites are present at the AU/GC borders (Goodall and Filipowicz, 1989; Lou *et al.*, 1993; McCullough *et al.*, 1993). Preferential utilization of the intronic non-consensus gttAG/cc 3′ ss in introns syn70ST and syn48ST would not be consistent with these properties of plant splicing unless these introns attain some special structural features. The finding that in introns syn70 and syn48, in which the downstream part of the palindrome is modified, the gttAG/cc site is not used (Figure 3), indicates that these special features are likely to be hairpins.

The observation that splicing to the stem-located 3′ ss in syn70ST and syn48ST requires the presence of a helper 3′ ss indicates that the 3′ ss may be recognized at least twice during spliceosome assembly in plant cells. Similar observations have been previously made using mammalian cell *in vitro* systems. Ge *et al.* (1990) have found that splicing of the atypical 48 nt intron of the polyoma t antigen pre-mRNA requires the assistance of another 3′ ss positioned downstream of the intron. Likewise, utilization of the additional AG present in the human β-globin gene intron 1 was shown to be facilitated by the wild-type downstream 3′ ss (Krainer *et al.*, 1985; Zhuang and Weiner, 1990). Interestingly, while the stem-located AG in syn70ST introns was efficiently used as an acceptor, irrespective of whether it was present in the optimal (TGCAG/GT) or sub-optimal (gttAG/cc) 3′ ss context, the helper function could only be provided by the sequence conforming to the 3′ ss consensus (TGCAG/GT) and not by the sub-optimal site TGtAG/cc (Figure 4). Therefore, the contribution of individual nucleotides of the 3′ ss may vary at different steps of splicing suggesting that distinct factors recognize this site at early and late stages of spliceosome assembly. It will be of interest to analyse additional splice site mutants of the syn70ST intron.

In mammalian systems, it is the first AG dinucleotide downstream of the branch site that is usually selected as an acceptor, probably by a scanning mechanism (Reed, 1989; Smith *et al.*, 1989, 1993; Zhuang and Weiner, 1990). It is not known whether a similar mechanism operates in plants although two pieces of evidence indicate that scanning might, at least partially, contribute to 3′ ss selection. Firstly, insertion of an artificial AG, 6 bp upstream of the wild-type AG in intron 1 of castor bean catalase gene, results in its exclusive use (Tanaka *et al.*, 1990); secondly, AG dinucleotides are very rarely found in the region extending 20 nt upstream of the 3′ ss in plant introns (our unpublished observations). If scanning for the acceptor AG does exist in plants, then this may be the most likely step at which the unfolding of pre-mRNA occurs in plants.

It remains to be established whether the potential of the spliceosome to unfold pre-mRNA in the 3′ ss region is a more general phenomenon. In a study addressing the scanning model of the 3′ ss selection, performed in HeLa extracts, Smith *et al.* (1993) found that a 22 bp GC-rich stem−loop placed between the polypyrimidine tract and the 3′ ss of the rat α-tropomyosin intron either inhibits the second step of splicing or, when placed closer to the branch point, is bypassed during the scanning. In the latter case, the PyAG located either in the stem or the loop sequence was ignored, with the more downstream PyAG being selected for splicing (AGs preceded by pyrimidines and not by purines are usually efficiently selected as acceptors during scanning; Reed, 1989; Smith *et al.*, 1989, 1993). In another study (Chebli *et al.*, 1989), a hairpin present between the branch point and the polypyrimidine

tract of the adenovirus intron was found to have a positive effect on splicing in the HeLa cell-free system, possibly by shortening the effective branch point to 3′ ss distance. Also in this case the intervening AG, present in the gAG/cc context, was ignored. Although these data argue against mRNA unfolding in mammals, they have to be interpreted with caution, since they are derived from *in vitro* experiments. Perhaps most pertinent are the results of Deshler and Rossi (1991) who found that correct splicing of the actin gene intron in the budding yeast *Kluyveromyces lactis* requires the formation of a hairpin which effectively reduces the 122 nt distance between the branch point and the 3′ ss. Since this hairpin also inhibited utilization of an intervening YAG as an acceptor, it could be argued that unfolding of mRNA does not occur during 3′ ss selection in this intron.

In summary, the results of our experiments, performed with dicot plant cells, suggest that spliceosomal components have the potential to unfold secondary structure in the downstream portion of the intron and that this process probably occurs at a late stage of spliceosome assembly, prior to or at the step when the AG to be utilized as the acceptor is defined. Elucidation of the mechanism of RNA unfolding and demonstration of whether an RNA helicase could possibly be involved in this process, must await further experimentation.

Materials and methods

Constructs for transient expression in protoplasts and for in vitro transcription

Unless stated otherwise, all techniques for manipulating DNA were as described by Sambrook *et al.* (1989). Plasmids containing synthetic genes syn7 (Goodall and Filipowicz, 1989) and its variants syn19ST and syn48ST (called previously synST19/14 and synST48/14, respectively; Goodall and Filipowicz, 1991) have been described previously. The sequence of the syn19ST intron is GTAAGATTATCGATATTTA-AAAATTATGGAAAATGTTAGCCAATTTTTTGGCTAACATTTTC-CATGGTTTTATGATATCATGCAG (inverted repeats are underlined); syn48ST is a derivative of syn19ST containing the 29 nt insertion, CGAAAAGAAGAAATAAATAATTTAAATAT, between intron positions 10 and 11. All constructs described below are derivatives of these three syn genes. They were obtained by replacing, in one or several steps, the regions between different restriction sites with appropriate new synthetic DNA fragments. Sequences of all constructs were verified by dideoxy sequencing. All modified synthetic genes described above were first generated as derivatives of the pGS-7 plasmid from which the *XbaI*−*PstI* fragment was subsequently cloned into the plant expression vector, pDELb (Goodall and Filipowicz, 1989). pGS derivatives were used for synthesis of complementary RNA probes, while the pDELb derivatives were used for protoplast transfections.

Syn5′ST/24, syn5′ST/18 and syn5′ST/12 are derivatives of syn7 in which exon 1 sequences have been altered (sequences of the hairpins are shown in Figure 1). Syn3′ST was prepared by replacing the sequence of the syn7 exon 2 (positions 10–37) with the sequence TTTTCGCTCGTACCTGCATGATATCATA (see also legend to Figure 2B). In syn44ST, the 12 nt sequence GAAAAGAAGAAA (intron positions 12–23) of syn48ST was replaced with the 8 nt sequence TGATCATT. In syn40ST and syn36ST, the 4 nt sequence ATCT (intron positions 9–12) of syn44ST was deleted; an additional 4 nt ATTT deletion (intron positions 17–20 in syn44ST) then generated syn36ST. An 8 nt sequence (GAAAAGAA) or 12 nt sequence (GAAAAGAA-GAAA) was inserted into syn44ST between intron positions 11–12 to generate syn52ST and syn56ST, respectively. Syn70ST was made by inserting a 22 nt sequence (ATCATTCCTAAAGATCTATACG) into syn48ST between intron positions 41 and 42. Constructs syn48 and syn70 were obtained by replacement of a *Bst*XI−*Eco*RV fragment of syn48ST and syn70ST with the sequence ATGGAATTACAAACG-AATAAACATATGAT. To generate syn44ST-16, a *ClaI*−*NcoI* fragment of syn44ST was replaced by the sequence CGATATTTAAAAATTA-TGGTGTTAGCCATCCAATTTTTTGGATGGCTAACAC (the altered nucleotides are underlined). Syn36ST/18 and syn36ST/12 and syn40ST/18 and syn40ST/12 are derivatives of syn36ST and syn40ST, respectively obtained by replacing their *ClaI*−*Eco*RV fragments with the hairpin sequences shown in the legend of Figure 7B.

Syn48ST-29, syn70ST-6 and syn19ST-70 are derivatives of syn48ST, syn70ST and syn19ST, respectively, obtained by changing the sequences between the *NcoI* and *XhoI* sites. The changed sequences were: a 15 nt (AACTTTACTTCAATA) insertion between intron positions 109 and 110 of syn48ST to obtain syn48ST-29; an 8 nt (ATGATATC) deletion between intron positions 123 and 130 of syn70ST) to give syn70ST-6; a 64 nt insertion (sequence not shown) between intron positions 78 and 79 of syn19ST to give syn19ST-70. The 3′ ss mutants of the syn70ST and syn70 series were obtained by replacing appropriate regions in syn70ST and syn70 with the new modified sequences (for sequences of individual mutants, see Figure 4C). To obtain the plasmid pGS3′l, used for the synthesis of the P_L probe, a fragment from syn3′ST (in the pDELb vector) was amplified by PCR with the primers 5′-CCCGGGG-ATCCTCTAGAG-3′ and 5′-CGAATAGTCACTTTCGTG-3′ and cloned into the *SmaI* site of pGEM1.

Transient expression in protoplasts and analysis of splicing

Transfections of mesophyll protoplasts of *N.plumbaginifolia* were performed as described by Goodall and Filipowicz (1990). RNA was isolated either 7 or 24 h after transfection and treated with DNase I to remove plasmid DNA (Goodall *et al.*, 1990). When the RNA was used for RT-PCR, another round of DNase treatment was performed. RNA probes for RNase A/T1 mapping were prepared by *in vitro* transcription with SP6 polymerase from the pGS series of plasmids, linearized with *Eco*RI. All probes were gel purified before use. RNase A/T1 mapping was performed as described by Goodall and Filipowicz (1989) except that hybridization reactions were performed at 45°C for 12 h and RNase digestions at 34°C for 40 min. RNA was fractionated on 6% polyacrylamide−8 M urea gels. The efficiency of splicing was determined by measuring the radioactivity in the protected bands using a PhosphoImager. Correction was made for the different amounts of label in each protected fragment. The values given are mean values determined from at least three transfections.

RT-PCR analysis was carried out essentially as described by Lou *et al.* (1993). Reaction mixtures (100 µl) contained 1 µg of total protoplast RNA and 50 pmol of primers specific for exon 1 (5′-GGATCCTCTAGA-GCCACGTCCCTCA-3′) and exon 2 (5′-GACCCCTGCAGCAGGCC-TATCGCGATACCCATC-3′) of the syn gene. The latter primer was 5′-end labelled (1000 c.p.m./pmol) in the presence of [γ-^{32}P]ATP by T4 polynucleotide kinase. The cDNA synthesis step was performed for 30 min at 50°C, followed by 20 cycles of PCR (denaturation at 94°C for 1 min, annealing at 55°C for 2 min and polymerization at 72°C for 2 min). PCR products were fractionated on 6% polyacrylamide−8 M urea gels.

Acknowledgements

We thank J.Petruska and A.Glaser for excellent technical assistance and Drs M.Konarska, J.Paszkowski, H.Rothnie and G.Simpson for valuable discussions and/or critical reading of the manuscript.

References

Chebli,K., Gattoni,R., Schmitt,P., Hildwein,G. and Stevenin,J. (1989) *Mol. Cell. Biol.*, **9**, 4852–4861.

Clouet d'Orval,B., d'Aubenton-Carafa,Y., Sirand-Pugnet,P., Gallego, M.E., Brody,E. and Marie,J. (1991) *Science*, **252**, 1823–1828.

Deshler,J.O. and Rossi,J.J. (1991) *Genes Dev.*, **5**, 1252–1263.

Dreyfuss,G., Matunis,M.J., Pinol-Roma,S. and Burd,C.G. (1993) *Annu. Rev. Biochem.*, **62**, 289–321.

Eperon,L.P., Estibeiro,J.P. and Eperon,I.C. (1986) *Nature*, **324**, 280–282.

Eperon,L.P., Graham,I.R., Griffiths,A.D. and Eperon,I.C. (1988) *Cell*, **54**, 393–401.

Filipowicz,W., Gniadkowski,M., Klahre,U. and Liu,H.-X. (1994) In Lamond,A. (ed.), *Pre-mRNA Processing*. R.G.Landes Publishers, Georgetown, TX, in press.

Fu,X.-Y. and Manley,J.L. (1987) *Mol. Cell. Biol.*, **7**, 738–748.

Ge,H., Nobel,J., Colgan,J. and Manley,J.L. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 3338–3342.

Goguel,V. and Rosbash,M. (1993) *Cell*, **72**, 893–901.

Goguel,V., Wang,Y. and Rosbash,M. (1993) *Mol. Cell. Biol.*, **13**, 6841–6848.
Goodall,G.J. and Filipowicz,W. (1989) *Cell*, **58**, 473–483.
Goodall,G.J. and Filipowicz,W. (1990) *Plant Mol. Biol.*, **14**, 727–733.
Goodall,G.J. and Filipowicz,W. (1991) *EMBO J.*, **10**, 2635–2644.
Goodall,G.J., Kiss,T. and Filipowicz,W. (1991) *Oxford Surveys Plant Mol. Cell Biol.*, **7**, 255–296.
Goodall,G.J., Wiebauer,K. and Filipowicz,W. (1990) *Methods Enzymol.*, **181**, 148–161.
Green,M.R. (1991) *Annu. Rev. Cell Biol.*, **7**, 559–599.
Himmelspach,M., Gattoni,R., Gerst,C., Chebli,K. and Stevenin,J. (1991) *Mol. Cell. Biol.*, **11**, 1258–1269.
Kandels-Lewis,S. and Seraphin,B. (1993) *Science*, **262**, 2035–2039.
Kim,S.H., Smith,J., Claude,A. and Lin,R.J. (1992) *EMBO J.*, **11**, 2319–2326.
Krainer,A.R., Reed,R. and Maniatis,T. (1985) In Berg,P. (ed.), *The Robert A.Welch Foundation Conferences on Chemical Research XXIX. Genetic Chemistry: The Molecular Basis of Heredity.* Welch Foundation, Houston, TX, pp. 353–382.
Lesser,C.F. and Guthrie,C. (1993) *Science*, **262**, 1982–1988.
Libri,D., Piseroa,A. and Fiszman,M.Y. (1991) *Science*, **252**, 1842–1845.
Lou,H., McCullough,A.J. and Schuler,M.A., (1993) *Mol. Cell. Biol.*, **13**, 4485–4493.
Luehrsen,K.R. and Walbot,V. (1994) *Genes Dev.*, **8**, 1117–1130.
Madhani,H.D. and Guthrie,C. (1992) *Cell*, **71**, 803–817.
Madhani,H.D. and Guthrie,C., (1994a) *Genes Dev.*, **8**, 1071–1086.
Madhani,H.D. and Guthrie,C. (1994b) *Annu. Rev. Genet.*, in press.
McCullough,A.J., Lou,H. and Schuler,M.A. (1993) *Mol. Cell. Biol.*, **13**, 1323–1331.
Moore,M.J., Query,C.Q. and Sharp,P.A. (1993) In Gesteland R. and Atkins,J. (eds), *The RNA World.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–358.
Newman,A. (1987) *EMBO J.*, **6**, 3833–3839.
Orozco,B.M., Robertson,M.C., Werneke,J.M. and Ogren,W.L. (1993) *Plant Physiol.*, **102**, 227–232.
Parker,R. and Siliciano,P.G. (1993) *Nature*, **361**, 660–662.
Plumpton,M., McGarvey,M. and Beggs,J.D. (1994) *EMBO J.*, **13**, 879–887.
Reed,R. (1989) *Genes Dev.*, **3**, 2113–2123.
Reich,C.I., VanHoy,R.W., Porter,G.L. and Wise,J.A. (1992) *Cell*, **69**, 1159–1169.
Rymond,B.C. and Rosbash,M. (1992) In Jones,E.W., Pringle,J.R. and Broach,J.R. (eds), *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression.* Volume II, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 143–192.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schwer,B. and Guthrie,C. (1991) *Nature*, **349**, 494–499.
Seraphin,B. and Kandels-Lewis,S. (1993) *Cell*, **73**, 803–812.
Smith,C.W.J., Porro,E.B., Patton,J.G. and Nadal-Ginard,B. (1989) *Nature*, **342**, 243–247.
Smith,C.W.J., Chu,T.T. and Nadal-Ginard,B. (1993) *Mol. Cell. Biol.*, **13**, 4939–4952.
Solnick,D. (1985) *Cell*, **43**, 667–676.
Solnick,D. and Lee,S.I. (1987) *Mol. Cell. Biol.*, **7**, 3194–3198.
Solymosy,F. and Pollak,T. (1993) *Crit. Rev. Plant Sci.*, **12**, 275–396.
Sontheimer,E.J. and Steitz,J.A. (1993) *Science*, **262**, 1989–1996.
Tanaka,A., Mita,S., Ohta,S., Kyozuka,J., Shimamoto,K. and Nakamura,K. (1990) *Nucleic Acids Res.*, **18**, 6767–6770.
Yoshimatsu,T. and Nagawa,F. (1989) *Science*, **244**, 1346–1348.
Zhuang,Y. and Weiner,A.M. (1990) *Gene*, **90**, 263–269.

Received on August 29, 1994; revised on October 24, 1994