Processing of chimeric introns in dicot plants: evidence for a close cooperation between 5' and 3' splice sites

Elisabeth Waigmann and Andrea Barta*

Institut für Biochemie, Universität Wien, Währingerstraße 17, A-1090, Vienna, Austria

Received October 9, 1991; Revised and Accepted December 9, 1991

ABSTRACT

Splice sites of vertebrate introns are generally not recognized in plant cells. Several lines of evidences have led to the proposal that the mechanism of 3' splice site selection differs in plants and animals (K. Wiebauer, J.J. Herrero, and W. Filipowicz, Mol. Cell. Biol. 8:2042–2051, 1988). To gain a better insight into the mechanistic differences between plant and animal splicing, we constructed chimeric introns consisting partly of dicotyledonous plant and partly of animal intron sequences. Splicing of these chimeric introns was analyzed in transiently transfected tobacco protoplasts. The results show that there are no principal sequence or structural differences between the 3' splice regions of plants and animals. Furthermore, evidence is provided that cooperation between 5' and 3' splice sites takes place and influences their mutual selection.

INTRODUCTION

The basic mechanism of nuclear pre-mRNA splicing has been elucidated in vertebrates and yeast. In both systems two subsequent transesterification steps lead to the excision of intervening sequences (IVS) as branched lariats and the ligation of neighbouring exons. Consensus sequences at the intron borders (5' and 3' splice sites) and within the intron (branch point sequence and polypyrimidine stretch) have been identified as important recognition signals for splicing (for review see: 1–4). Splicing is facilitated by a number of nuclear factors that assemble together with the pre-mRNA in a large ribonucleoprotein complex, the spliceosome (for review see: 1–3). An important part of these spliceosomal factors are the small nuclear ribonucleoprotein particles (snRNP) U1, U2, U4/U6 and U5 which play a vital role in the recognition of splicing signals on pre-mRNA (for review see: 5).

Much less is known about the plant splicing system. The consensus sequences of 5' and 3' splice sites are nearly identical to the corresponding vertebrate consensus sequences (6, 7) and U-RNAs have also been identified in plants (8–11). Structural features important for snRNP assembly (anti Sm binding site) or pre-mRNA splicing (e.g. sequence complementarity of U1 RNA to the 5' splice region) are conserved in plant U-RNAs

(for review see: 12, 13). On the other hand, mutational analysis of a synthetic plant intron did not reveal any requirement for a particular sequence which makes the presence of even a degenerated branch point sequence unlikely. In addition, there is no requirement for a polypyrimidine stretch anywhere in the intron (14). A remarkable feature of plant introns is their elevated AU-level compared to the surrounding exons (dicotyledonous plants: 73% versus 55%, monocotyledonous plants: 56% versus 44%, see 15). So far the lowest AU-level found in a natural dicotyledonous intron is 60.5% (14). Furthermore, differences in the processing efficiencies of pre-mRNAs in monocotyledonous and dicotyledonous plants have been observed (16–19).

In spite of identical 5' and 3' splice site consensus sequences and similar structures of snRNA's a clear difference in the requirements for splicing in plant and animal *in vivo* systems was observed: Introns in pre-mRNA from the human growth hormone and the human α -globin genes cannot be recognized by the plant splicing machinery (20, 21). Furthermore, in dicot systems human β -globin IVS2 is processed with low efficiency and uses a cryptic 3' splice site in context with the authentic 5' splice site (15). Only the SV 40 small t intron, an unusually small AU-rich animal viral intron, has recently been processed successfully in dicot plants (22). Similar experiments carried out with plant pre-mRNAs in HeLa *in vivo* systems also showed the selection of cryptic 3' splice sites together with authentic 5' splice sites (in the case of soybean leghemoglobin IVS1 and IVS3), but correct and efficient processing of plant introns took place as well (21). Several plant introns have been faithfully and efficiently processed in HeLa *in vitro* systems (7, 23). These results indicated, that the mechanism of 3' splice site selection is different in plants and animals and that selection conditions are more stringent in plants. Details of the processing of mammalian introns in plant *in vitro* systems are however not available, owing to the lack of such systems.

To gain a better insight into the requirements of the plant splicing machinery and the mechanistic differences between plant and animal splicing we constructed several chimeric introns—consisting partly of plant and partly of animal sequences—and studied their processing in tobacco protoplasts *in vivo*. Surprisingly, in such constructs an authentic vertebrate 3' splice site was activated indicating that the structural requirements of the 3' splice sites per se do not differ between animals and plants.

* To whom correspondence should be addressed

In addition, the results suggest that 'compatible' splice sites are necessary for splicing, thus indicating an intimate cooperation between the splice sites.

MATERIALS AND METHODS

Materials

Restriction enzymes were from Böhrringer Mannheim or New England Biolabs. T7 and T3 RNA polymerases were from Böhrringer Mannheim and Stratagene, respectively. Avian myeloma virus reverse transcriptase was supplied by United States Biochemicals and Taq DNA polymerase by Promega. pBluescript KS(−) was from Stratagene.

Construction of plasmids

DNA manipulations were done according to standard methods (24).

Cloning of human metallothionein (hMT): *Bam*HI digestion of pSP64-hMT (25) yielded a fragment spanning 26 bp of exon1, 306 bp of intron1 and 69 bp of exon2 and 10 bp of intron2. This fragment was cloned into the *Bam*HI site of Bluescript(−) and pDH51 (26) giving rise to plasmids Blue-hMT and pDH51-hMT, respectively.

Cloning of the small subunit of ribulose 1,5-bisphosphate carboxylase (RbcS) of *Pisum sativum*: A 563 bp *Eco*RV–*Xba*I fragment of the RbcS gene of pea covering 163 bp of exon1, intron1 (83 bp), exon2 (137 bp), intron2 (85 bp) and 75 bp of

exon3 was cut out of pSP64-RbcS (23) and cloned into the *Sma*I/*Xba*I sites of pDH51 creating plasmid pDH51-RbcS.

Chimeric intron hMT-Leg (Chi-1): A 263 bp *Bam*HI–*Sma*I fragment of the hMT gene containing 26 bp of exon1 and the 5' part of intron1 and a 216 bp *Dra*I–*Pst*I fragment of pSP65LegJi (7) covering the 3' part of the intron and 155 bp of exon2 of the leguminJ gene from *Pisum sativum* were ligated into the *Bam*HI/*Pst*I site of Bluescript(−) yielding plasmid Blue-Chi-1. The chimeric intron and its flanking exons were cut out with *Bam*HI and the 456 bp fragment was cloned into the *Bam*HI site of pDH51, generating plasmid pDH51-Chi-1.

Chimeric intron hMT-RbcS (Chi-3): A 263 bp *Bam*HI–*Sma*I fragment of hMT gene containing 26 bp of exon1 and the 5' part of intron1 and a 321 bp *Hpa*I–*Hind*III fragment of RbcS gene containing the 3' part of intron1 (23 bp), exon2, intron2 and part of exon3 were cloned into the *Bam*HI/*Hind*III site of Bluescript(−) yielding plasmid Blue-Chi-3. A 600 bp fragment spanning part of the hMT exon1, the chimeric intron, RbcS exon2, RbcS intron 2 and part of RbcS exon3 was cut out by digestion with *Bam*HI and *Sal*I and cloned into the *Bam*HI and *Sal*I sites of pDH51, giving rise to plasmid pDH51-Chi-3.

Chimeric intron hMT-elongated RbcS (Chi-3V): Plasmid Blue-Chi-3 was the starting material for the elongation of the RbcS part of the chimeric intron. The elongation was done by polymerase chain reaction with specific primers. In the first step plasmid Blue-Chi-3 was amplified with primer pairs L5 (5' cagtaatgattcaagttgttaaccggttatgttg 3')/K8 (5' aacagctatgaccatg 3') giving rise to a RbcS fragment of 440 bp and L6 (5'

Fig. 1. Sequences of chimeric introns. Plant sequences of the chimeric introns are derived from either the leguminJ gene (Leg) or the RbcS gene of pea (notation in *italics*), animal sequences are contributed by the hMT-IIA gene. Chi-1: The 5' part of intron1 (236 bp) from hMT has been joined to the 3' part of the leguminJ intron (61 bp). Chi-3 and Chi-3V: The same vertebrate 5' part as in Chi-1 is ligated to the 3' part of the RbcS intron1 (23 bp). Chi-3: 44 bp, Chi-3V: 39 bp. Chi-9: The 5' part of intron1 (60 bp) from the plant gene RbcS is joined to the 3' part of intron1 from the hMT gene (69 bp). Chi-10: The 5' part of the leguminJ intron (74 bp) is ligated to the 3' part of the hMT intron1 (69 bp). Sequences of the chimeric introns have been verified by dideoxy sequencing reactions.

aacaaactgaatcattactggggccgagttctc 3')/K7 (5' gtttccagtcacgac 3') which resulted in a hMT fragment of 389 bp length. The fragments were isolated and used directly for a second amplification step with primers K7/K8. The reaction product, a fragment of 808 bp, which already contained the elongated chimeric intron, was digested by *Bam*HI and *Hind*III and cloned into the *Bam*HI/*Hind*III site of Bluescript(-), thereby generating plasmid Blue-Chi-3V. A 620 bp fragment spanning RbcS exon1, the elongated chimeric intron and hMT exon2 was cut out of Blue-Chi-3V with *Bam*HI and *Sal*I and ligated into the *Bam*HI/*Sal*I site of pDH51, giving rise to plasmid pDH51-Chi-3V.

Chimeric intron RbcS-hMT (Chi-9): A 268 bp *Hind*III-*Hpa*I fragment from pSP64-RbcS covering 208 bp of exon1 and the 5' part of intron1 of the RbcS gene and a 148 bp *Sma*I-*Bam*HI fragment of pSP64-hMT that contains the 3' part of intron1 and 79 bp of exon2 of the hMT gene were cloned into the *Hind*III/*Bam*HI sites of Bluescript(-) resulting in plasmid Blue-Chi-9. A 380 bp *Eco*RV-*Xba*I fragment spanning the chimeric intron and surrounding exons was inserted into the *Sma*I/*Xba*I sites of pDH51 creating plasmid pDH51-Chi-9.

Chimeric intron Leg-hMT (Chi-10): A 113 bp *Dra*I-*Eco*RI fragment of pSP65LegJi spanning 69 bp of the 3' part of intron1 and 78 bp of exon2 of the legumin gene and a 148 bp *Sma*I-*Bam*HI fragment of pSP64-hMT covering the 3' part of

intron1 and 79 bp of exon2 of the hMT gene were cloned into the *Eco*RI/*Bam*HI sites of Bluescript(-) giving rise to plasmid Blue-Chi-10. A 273 bp *Eco*RV-*Xba*I fragment containing the chimeric intron and surrounding exons was cloned into the *Sma*I/*Xba*I sites of pDH51 creating plasmid pDH51-Chi-10.

Transfection of protoplasts via electroporation

Plant protoplasts prepared essentially as described (27) from *Nicotiana tabacum* L. cv. Xanthi suspension cultures were transfected with 10 μ g pDH51-DNA and 50 μ g salmon sperm carrier DNA per 400 μ l sample containing 210⁶ protoplasts. Electroporation was carried out in EP-Medium (10 mM HEPES-KOH (N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid) pH 7.2, 150 mM NaCl, 5 mM CaCl₂, 0.2 M Mannitol) with a 250V, 750 μ F discharge pulse. Protoplasts were regenerated in K3 media at 26°C for 2–24 hrs (27).

Preparation of RNA substrates

Plant RNA: Protoplasts were harvested 2–24 hrs after electroporation and RNA was prepared according to the guanidinium-phenol-chloroform method (28). RNA was further purified by proteinase K digestion in 0.3 M NaCl, 10 mM Tris-HCl, pH 7.5, 5 mM EDTA, 0.5% SDS and 300 μ g/ml proteinase K (45 min, 37°C). Traces of DNA were removed by DNase I digestion (0.1 M NaOAc, pH 5.3, 5 mM MgSO₄·7H₂O, 300 U/ml RNasin and 250 U/ml DNase I) for 30 min at 37°C.

Fig. 2. Fig. 2A. Analysis of Chi-3 RNA expression in transiently transfected tobacco protoplasts. Protoplasts were electroporated with pDH51-Chi-3 DNA and total RNA was prepared 4 and 24 hours after transfection. RNA was analyzed by an RNase protection assay using a labeled antisense full length probe. Protected fragments were analyzed on an 8% polyacrylamide gel. Only the authentic plant intron was processed. Lane M: pBr322/*Msp*I; Lane C: RNase protection assay of RNA from mock transfected protoplasts; Lanes 4 and 24: RNase protection assay of Chi-3 RNA prepared 4 and 24 hours after electroporation. Abbreviations: E1:exon1, E2:exon2, E3:exon3, IVS2:intron2. Drawing symbols: unhatched boxes:vertebrate exons; hatched boxes:plant exons; unhatched bars:vertebrate intron sequences; hatched bars:plant intron sequences. Fig. 2B. Analysis of Chi-3V RNA expression in transiently transfected tobacco protoplasts. Protoplasts were electroporated with pDH51-Chi-3V DNA and total RNA was prepared 3h after transfection. RNA was analyzed by an RNase protection assay using a labeled antisense full length probe. Protected fragments were analyzed on an 8% polyacrylamide gel. Only the authentic plant intron was processed. Lane M: pUC8/*Msp*I; Lane C: RNase protection assay of RNA from mock transfected protoplasts. Lane 1: RNase protection assay of Chi-3V RNA. Abbreviations and drawing symbols as in Fig. 2A.

Fig. 3. Analysis of Chi-1 RNA expression in transiently transfected tobacco protoplasts. Protoplasts were electroporated with pDH51-Chi-1 DNA and RNA was prepared 3h after transfection. Panel A: RNase protection assay of total RNA using a labeled antisense full length probe (lane 1) or a 5' probe complementary to the hMT part of the chimeric intron (lane 2). Lane C: RNase protection assay of RNA from mock transfected protoplasts with full length probe or 5' probe. Lane M: pBr322/*Msp*I. Panel B: Primer extension analysis of polyA⁺ selected RNA. Lane RT: Reverse transcription using primer P7 complementary to sequences at the 5' end of exon2. Lanes G,A,T,C: Dideoxy sequencing reactions of Chi-1 DNA using primer P7. Abbreviations and drawing symbols as in Fig. 2A. E1': alternatively spliced exon1. The arrow marks the cryptic 5' splice site. The dashed box represents the additional exon1 sequences when the cryptic 5' splice site is used.

PolyA⁺ selection was carried out with Hybond-map polyU filter paper (Amersham) as described by the supplier.

Preparation of *in vitro* transcripts: Antisense RNA probes were synthesized from linearized DNA templates by T7 or T3 polymerase in the presence of (α -³²P)UTP (Amersham, 800 Ci/mMol) as described (29). Templates used: Blue-hMT/*Spe*I; Blue-Chi-3/*Xba*I; Blue-Chi-1/*Spe*I; Blue-Chi-9/*Hind*III; Blue-Chi-9/*Hinc*II; Blue-Chi-10/*Eco*RI; Blue-Chi-3V/*Xba*I.

Analysis of RNA

RNAse protection assay: The assay was carried out essentially as described (29). 20 μ g total RNA or 0.2 μ g of polyA⁺ RNA in 18 μ l hybridization buffer (80% formamide, 0.4 M NaCl, 40 mM PIPES [piperazine-*N,N'*-bis(2-ethanesulfonic acid)] pH 6.4, 5 mM EDTA) were hybridized to 150 000 cpm antisense RNA in 2 μ l H₂O for 16 h at 50°C after denaturation (1' 90°C, 15' 85°C). RNAse A digestion was performed by adding 200 μ l of RNAse A solution (40 μ g/ml RNAse A, 10 mM Tris-HCl, pH 7.5, 5 mM EDTA, 300 mM NaCl) and incubation at 26°C for 1 h. RNA was analyzed on an 8% denaturing polyacrylamide gel.

Primer extension: 2 μ g polyA⁺ RNA and 100 fmol of 5'-P³²-labeled oligo-nucleotide primer P7 (5' ctgtttctgggtcccaccaagg 3') or primer MT (5' gcaggagccggcgaggtgc 3') in 20 μ l of prePE buffer (10 mM Tris-HCl, pH 8, 50 mM NaCl, 1 mM EDTA) were denatured for 10' at 70°C, gradually cooled down to 42°C (several hours) and annealed for 12 h (30). Primer extension was carried out for 1 h at 42°C by adding 10 μ l of a reaction mixture containing 150 mM Tris-HCl, pH 8.3, 120 mM KCl, 20 mM MgCl₂, 150 μ g/ml actinomycin D, 3 mM dNTPs, 1.5 U/ μ l RNasin and 1.5 U/ μ l of AMV reverse transcriptase (USB). RNA was degraded in 100 μ l RNAse A solution (40 μ g/ml RNAse A, 10 mM EDTA) for 30 min at 37°C. Products were analyzed on an 8% denaturing polyacrylamide gel. For accurate length evaluation a dideoxy DNA sequencing reaction using plasmid Blue-Chi-1 and primer P7 or plasmid pDH51-Chi-10 and primer MT was carried out in parallel (T7 sequencing kit, Pharmacia).

RESULTS

Plant expression vectors containing the chimeric introns

Five chimeric introns consisting partly of plant and partly of animal sequences were cloned into the plant expression vector pDH51 under the control of the CaMV 35S promoter and terminator and in Bluescript KS(−) under the control of T7 or T3 RNA polymerase promoter. The 5' part of the human metallothionein (hMT) intron1 was combined a) with the 3' part of the pea leguminJ intron giving rise to hMT-Leg construct Chi-1, b) with a fragment spanning 23 bp of the 3' part of intron1 of pea ribulose 1,5-bisphosphate carboxylase (RbcS) thereby generating hMT-pRbcS (Chi-3) and c) with a fragment spanning 44 bp of the 3' part of intron1 of pea RbcS thereby generating hMT-pRbcSV (Chi-3V). In the hMT-RbcS constructs Chi-3 and Chi-3V an additional authentic plant intron, the RbcS intron2, was included and was used as an internal positive control. In the RbcS-hMT construct Chi-9 the pea RbcS 5' part of intron1 was fused to the 3' part of hMT intron1 (Fig. 1). The Leg-hMT construct Chi-10 was created by combining the 5' part of intron of the leguminJ gene to the 3' part of hMT intron1 (Fig. 1).

For control experiments, a part of the pea RbcS gene including introns 1 and 2 was cloned into pDH51. The hMT intron1 and its surrounding exons were cloned into pDH51 and Bluescript KS(−).

Processing of hMT intron1

Most of the vertebrate introns tested so far in plant *in vivo* systems were not recognized by the plant splicing machinery (15, 20, 21). Nevertheless, it was necessary to verify the splicing behaviour of hMT intron1, which donates all the vertebrate sequences in our chimeric introns as it possesses an almost perfect match to the proposed plant consensus sequences at the 3' and 5' border of the intron. RNA of protoplasts transfected by pDH51-hMT was analyzed by an RNase mapping assay. Only one fragment corresponding to pre-mRNA was protected, indicating that neither splice site of the hMT intron1 was recognized (data not shown).

Processing of the hMT-pRbcS constructs Chi-3 and Chi-3V

In these constructs the chimeric intron consists of 237 nt of hMT sequences at the 5' site and either 23 nucleotides (Chi-3) or 44 nucleotides (Chi-3V) of pea RbcS at the 3' site (Fig. 1). 44 nucleotides of pea RbcS represent more than half of the original pea intron1 and should therefore contain all signals necessary for a functional 3' splice site. In addition to the chimeric intron, there is also an authentic plant intron present, the pea RbcS intron2, which is expected to be spliced. Total RNA from protoplasts transfected with Chi-3 or Chi-3V was hybridized to a full length anti-sense pre-mRNA probe (template: Blue-Chi-3/*Xba*I or Blue-Chi-3V/*Xba*I, 'full length probe') and analyzed by RNase digestion. Only the authentic plant intron was

Fig. 4. Analysis of Chi-9 RNA expression in transiently transfected tobacco protoplasts. Protoplasts were electroporated with pDH51-Chi-9 and total RNA was prepared 3h after transfection. Lane 1: hybridization of polyA⁺ Chi-9 RNA to full length probe. Lane 2: hybridization of RNA from pDH51-RbcS (gene structure see Table 1) transfected protoplasts to the Chi-9 full length probe used in lane 1, thereby supplying a correctly processed exon1 that can be compared to the exon1 band in lane 1. Lanes M: pBr322/*Msp*I. Abbreviations and drawing symbols as in Fig. 2A. The sequence of the 3' ends of E1' and E1 is given at the bottom of the drawing. E1': alternatively spliced exon1. The arrow marks the cryptic 5' splice site.

excised faithfully giving rise to two bands corresponding to exon3 (E3: 90 nucleotides) and exon1-chimeric intron-exon2 (E1-chimeric intron-E2: 423 nucleotides in the case of construct Chi-3, see Fig. 2A, lanes 4 and 24, or 444 nucleotides in the case of construct Chi-3V, see Fig. 2B, lane 1). In contrast to other authors, who prepare RNA 24 hrs after electroporation (15), in our experiments RNA levels were highest 3–4 hrs after electroporation and considerably lower at 24 hrs (Fig. 2A, lanes 4 and 24). As this was reproducible with other RNA's (data not shown) we routinely isolated RNA 4 hrs after electroporation.

Processing of the hMT-pLeg construct Chi-1

From the *in vivo* processing results of the human β -globin and soybean leghemoglobin in plants and animals, it was inferred that the selection of the 3' splice site may differ in these systems (15). It had therefore been expected that the pea 3' splice site in Chi-3 or Chi-3V would be recognized. As no processing of the chimeric intron took place, we generated another chimeric intron (Chi-1, see Fig. 1) with the same vertebrate 5' part as in the hMT-RbcS constructs and fused it to a different plant 3' part (65 nt of leguminJ intron).

Total RNA from protoplasts transfected with pDH51-Chi-1 was hybridized to a labeled full length anti-sense RNA probe (template: Blue-Chi-1/*Spe*I) and digested with RNase A (Fig. 3 A, lane 1). In addition to the pre-mRNA band, two protected fragments appeared, indicating that the chimeric intron had undergone splicing although not in the predicted way: the upper band approximately corresponded in size to leguminJ exon2 (E2:155 nucleotides), but the lower band (E1') was definitely not derived from the authentic hMT exon1 which consists of only 26 nucleotides (Fig. 3 A, lane 1). To test if the second band resulted from a cryptic 5' splice site, a 5' anti-sense probe complementary to the hMT part of Chi-1 was used for RNase mapping (Blue-hMT/*Spe*I). The probe protected a fragment corresponding to pre-mRNA and a fragment of the same size as the lower band in the previous experiment (Fig. 3 A, lane 2), thereby identifying this fragment (E1': 147 nucleotides) as exon1' generated by usage of a cryptic 5' splice site 121 nucleotides downstream of the authentic metallothionein 5' splice site.

To support this hypothesis, a primer extension experiment was carried out (Fig. 3 B, lane RT). The 5' labeled primer (P7) was complementary to sequences at the 5' end of exon2; to enable an accurate length evaluation the reverse transcription products were compared to a sequence ladder of Chi-1 DNA using the same primer (Fig. 3 B, lanes G,A,T,C). The bands for pre-mRNA and spliced product (191 nucleotides) were of the expected length and confirmed the RNase mapping results.

Taken together, the chimeric intron is processed, however, in place of the authentic vertebrate 5' splice site, a cryptic 5' splice site 121 nucleotides into the hMT intron1 is used in context with the authentic plant 3' splice site.

Processing of the RbcS-hMT construct Chi-9

As previous results had suggested a difference in 3' splice site selection, we constructed two different chimeric introns possessing the same 3' hMT part (Chi-9 and Chi-10). The chimeric intron of Chi-9 is composed of a plant 5' part (60 nucleotides of RbcS intron1) and a vertebrate 3' part (69 nucleotides of hMT intron1). RNase mapping of total RNA from transfected protoplasts resulted in a series of degradation fragments which prevented splicing analysis (data not shown). Therefore, polyA⁺ selection was employed to remove these

obscuring fragments. RNase mapping of polyA⁺ RNA with the full length Chi-9 anti-sense probe (template: Blue-Chi-9/*Hind*III) indeed protected only three fragments: pre-mRNA, the hMT exon2 (E2: 91 nucleotides), and a set of three bands close together and corresponding roughly in size to the authentic RbcS exon1 (Fig. 4, lane 1). As a positive control for the authentic exon1 length, RNA from protoplasts transfected with the original RbcS gene was probed with the Chi-9 anti-sense full length probe. The control experiment showed two adjacent protected fragments for exon1 (E1: 163 nucleotides; the appearance of a doublet is due to RNase nibbling, see Fig. 4, lane 2). This doublet corresponds to bands also present in Fig. 4, lane 1, which means that at least partially the authentic plant 5' splice site was used. However, the main exon1 band (E1': 153 nucleotides) in Fig. 4, lane 1 is not detected in the control experiment and is therefore due to the selection of a cryptic 5' splice site 10 nucleotides upstream from the authentic one. The alternatively processed exon1 fragment is termed exon1' (E1'). Sequences of the 3' ends of exon1 and exon1' are given at the bottom of Fig. 4.

The experiment shows that the upstream cryptic 5' splice site is used preferentially, while the authentic RbcS 5' splice site is recognized only in about 20% of the splicing events. Surprisingly, both sites are used in context with the authentic vertebrate 3' splice site.

Processing of the Leg-hMT construct Chi-10

The chimeric intron of Leg-hMT is composed of 74 nucleotides from the pea leguminJ intron at its 5' side and 69 nucleotides

Fig. 5. Analysis of Chi-10 RNA processing in transiently transfected tobacco protoplasts. Protoplasts were electroporated with pDH51-Chi-10 DNA and total RNA was prepared 3h after transfection. Panel A: Lane 1: RNase protection assay using a labeled antisense full length probe. Lane C: RNase protection assay of RNA from mock transfected protoplasts with full length probe. Lane M: pBr322/*Msp*I. Panel B: Primer extension analysis of polyA⁺ selected RNA using primer MT. Lane C: Reverse transcription of RNA from mock transfected protoplasts. Lane RT: Reverse transcription of Chi-10 RNA. S: spliced product. Lanes G,A,T,C: Dideoxy sequencing reactions of Chi-10 DNA using primer MT. Abbreviations and drawing symbols as in Fig. 2A.

Table 1: Summary of *in vivo* RNA processing behaviour of parental and chimeric constructs in tobacco protoplasts

		SPlicing
PLANT		+
ANIMAL		-
CHIMERIC CONSTRUCTS		
CHI-3		- (IVS2: +)
CHI-3V		- (IVS2: +)
CHI-1		+ 5': cryptic 3': authentic
CHI-9		+ 5': cryptic (80%) authentic (20%) 3': authentic (100%)
CHI-10		+ 5': authentic 3': authentic

Drawing symbols and abbreviations as in Fig. 2A. ↓: splice sites used in tobacco protoplasts

of hMT intron1 on its 3' side. Total RNA was analyzed with a full length anti-sense probe derived from Blue-Chi-10 cut with *Eco*RI. The RNase mapping pattern showed a band of 91 nucleotides corresponding to exon2 and a band of 38 nucleotides corresponding in length to exon1 (Fig. 5 A, lane 1). To verify this, a primer extension experiment was carried out (Fig. 5 B, lane RT). The 5' labeled primer MT was complementary to sequences at the 5' end of exon2; for the accurate length evaluation, the reverse transcription product was compared to a sequence ladder of Chi-10 DNA using the same primer (Fig. 5 B, lanes G,A,T,C). The band for spliced RNA ('S': 82 nucleotides) was of the expected length and confirmed the RNase mapping results. Taken together, the chimeric intron Leg-hMT was spliced efficiently; both the authentic plant 5' splice site and the authentic hMT 3' splice site were used.

DISCUSSION

To further investigate the differences between the plant and animal splicing machinery, we constructed several chimeric introns consisting partly of dicotyledonous plant and partly of animal sequences and analyzed their processing in tobacco protoplasts *in vivo* (Table 1).

Three constructs composed of the same vertebrate 5' part of hMT intron1 but ligated to different plant 3' parts behave in a completely different manner: In two cases, the chimeric introns are not processed at all; in the third case the authentic plant 3' site is used in context with a cryptic 5' site, selected from within the hMT part of the chimeric intron some 120 nucleotides downstream of the authentic vertebrate 5' splice site (Table 1). The chimeric introns of the two reverse constructs consisting of a plant 5' part and a vertebrate 3' part are also processed. Surprisingly, the authentic vertebrate 3' splice site is used in both cases, in one case mainly together with a cryptic 5' splice site close by to the authentic plant 5' splice site which is used to a much lesser extent. In the other case it is used together with the authentic plant 5' splice site and the intron is correctly removed (Table 1).

Several conclusions can be drawn from this complex processing pattern. A vertebrate 3' splice site is principally acceptable for

Table 2: Structural features of parental and chimeric introns

Intron	5'splice site	3'splice site	%AT	Splicing
RbcS IVS1	CAG [↓] GTGAC	TTTAG [↓] G	76	+
Legumin	AGA GTAAG	TGCAG G	75	+
CHI-10 (Leg-hMT) authentic	AGA GTAAG	TGCAG G	62	+
CHI-1 (hMT-Leg) alternative	GGG GTATC	TGCAG G	60	+
CHI-9 (RbcS-hMT) authentic	CAG GTGAC	TGCAG G	60	(+)
alternative	AAA GTGCA	TGCAG G	59	+
CHI-3V (hMT-RbcS) alternative	GGG GTATC	TTTAG G	57	-
authentic	CCG GTAAG	TTTAG G	54	-
CHI-1 (hMT-Leg) authentic	CCG GTAAG	TGCAG G	56	-
CHI-3 (hMT-RbcS) alternative	GGG GTATC	TTTAG G	55	-
authentic	CCG GTAAG	TTTAG G	53	-
hMT	CCG GTAAG	TGCAG G	50	-
Plant consensus sequence	AG GTAAG	TGCAG G		

the dicot plant splicing system, provided that a plant 5' splice site is available. This indicates that the inability of the plant system to process animal introns is not due to structural or sequence differences at the 3' splice site. Note that in context with its own animal 5' splice site the hMT 3' splice site is not recognized, which argues against a chance meeting of the plant requirements.

Furthermore, the presence of two principally active splice sites is not sufficient for the occurrence of a splicing event, as the hMT intron1 offers both a functional 5' splice site—though cryptic—and a functional 3' splice site to the plant system and still cannot be processed.

Clear preferences in splice site pairing were found. In our approach, the same vertebrate 5' part (hMT intron1) was connected to four different 3' parts: 23 nucleotides of the 3' part of RbcS intron1, 44 nucleotides of the 3' part of RbcS intron1, 65 nucleotides of the 3' part of leguminJ intron and the 3' part of hMT intron1. Only fusion to the 3' part of leguminJ resulted in a successful splicing event although both the 3' part of RbcS intron1 and the 3' part of hMT intron1 are able to act as functional 3' splice sites in plants if in context with other 5' splice sites. Moreover, in the RbcS-hMT construct the authentic vertebrate 3' splice site selects mainly a cryptic 5' splice site although the authentic plant 5' splice site is only 10 nucleotides apart. This points to a close cooperation between 5' and 3' splice sites which can exert influence on their mutual selection. Cooperation between 5' and 3' splice sites has been also demonstrated in the selection of alternative splice sites in the case of adenovirus-2 E1A pre-mRNA (31) and the myosin light-chain 1/3 pre-mRNA (32) in a HeLa *in vivo* system.

The frequent usage of cryptic splice sites observed in our experiments was quite unexpected in view of experiments with a synthetic intron introduced into dicot plants by Filipowicz's group. Sequence changes within this synthetic intron only influenced splicing efficiency but did not lead to the selection of cryptic splice sites (14). However, in their experiments the exon environment and the intron borders remained constant.

In a more recent paper Goodall and Filipowicz show that optimal splice site sequences can improve splicing efficiency in monocot but not in dicot plants (19). This is in agreement with our results as the selection of new splice sites in our experiments

cannot be correlated to an enhanced homology to the published plant splice site consensus sequences. In Table 2 all combinations of 5' and 3' splice sites used in our experiments are listed. The intron part of the splice site consensus sequence does not seem to exert any influence on the spliceability of the intron as e.g. the consensus sequences 5'-GTAAT and TGCAG-3' occur both in the leguminJ intron (spliced), the hMT intron1 (not spliced), and the authentic Chi-1 (not spliced). Therefore, the functional unit consists not only of the splice site consensus sequence but also includes an as yet unidentified portion of adjacent sequences.

A prominent feature of plant introns is their unusual high AU-content in contrast to the AU-content of the surrounding exons (73% versus 55% in dicots, 56% versus 44% in monocots). The lowest known AU-content of a naturally occurring dicotyledonous plant intron (Alcoholdehydrogenase of *Arabidopsis thaliana*) is 60,5%. The influence of AU-rich regions in plant introns has been demonstrated by inserting AU-sequences at various places in a synthetic GC-rich intron thereby restoring its spliceability in dicot plants (14). However, a very recent paper indicates that AU-rich introns are not essential for splicing in monocots (19). The spliceability of the introns in our constructs was correlated to their AU-content (Table 2). Interestingly, no processing of an intron below 59% AU occurred which fits well to the observed natural AU-limit in dicotyledonous plants. In contrast, synthetic introns with an AU-content as low as 41% could still be processed in dicot plants (14).

In spite of the correlation between AU-content and spliceability of introns in our approach a simple model based solely on the overall AU-content is not sufficient to define a dicot plant intron. This is illustrated most clearly in Chi-9: A cryptic plant 5' splice site is used preferentially over the authentic plant 5' splice site, although this is accompanied by a slight decrease in the overall AU-content of the removed intron (Table 2). As the plant exon 1 and 5' part of the chimeric intron in Chi-9 have a wildtype plant sequence, the activation of this cryptic site can only be due to the combination with the animal 3' part of the construct. Therefore, the results presented in this paper point to a clear preference in splice site pairing.

Taken together, it seems that the decision whether or not a splicing event takes place is influenced by AU-rich sequences (possibly because secondary structure would be easier to open up in this region or because of the existence of splicing factors requiring AU-rich regions for binding). The exact definition of the intron borders would then be brought about by a close cooperation of the 5' and 3' splice regions, exerting a mutual influence on the selection of splice sites. The nature of this cooperation can be uncovered by defining all the components involved in plant splicing, which will greatly depend on the development of an *in vitro* plant splicing system.

ACKNOWLEDGMENTS

We are grateful to J. Brown for providing plasmid pSP65LegJi, H. Zellhofer for technical assistance and T. Skern and M. Matzke for critically reading the manuscript. We are indebted to K. Hartmuth and Z. Rattler for comments and invaluable discussions. This work was supported by grants from the Österreichische Fonds zur Förderung der wissenschaftlichen Forschung and from the Bundesministerium für Wissenschaft and Forschung to A.B.

REFERENCES

1. Green, M.R. (1986) *Ann. Rev. Genet.*, **20**, 671–708.
2. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S., and Sharp, P.A. (1986) *Ann. Rev. Biochem.*, **55**, 119–150.
3. Ruby, S., and Abelson, J. (1991) *Trends Genetics*, **7**, 79–85.
4. Sharp, P.A. (1987) *Science*, **235**, 766–771.
5. Steitz, J.A., Black, D.L., Gerke, V., Parker, K.A., Kraemer, A., Frendewey, D., and Keller, W. (1987) In Birnstiel, M. (ed.). *Structure and function of major and minor small nuclear ribonucleoprotein particles*. Springer Verlag, Heidelberg, pp. 115–154.
6. Brown, J. (1986) *Nucleic Acids Res.*, **14**, 9549–9559.
7. Brown, J.W.S., Feix, G., and Frendewey, D. (1986) *EMBO J.*, **5**, 2749–2758.
8. Abel, S., Kiss, T., and Solymosy, F. (1989) *Nucleic Acids Res.*, **17**, 6319–6337.
9. Brown, J.W.S., and Waugh, R. (1989) *Nucleic Acids Res.*, **17**, 8991–9001.
10. Hanley, B.A., and Schuler, M.A. (1989) *Nucleic Acids Res.*, **17**, 10106.
11. Vankan, P., Edoh, D., and Filipowicz, W. (1988) *Nucleic Acids Res.*, **16**, 10425–10440.
12. Filipowicz, W., Kiss, T., Marshallsay, C., and Waibel, F. (1990) *Mol. Biol. Reports*, **14**, 125–129.
13. Solymosy, F. (1991) *Acta Biochim. Biophys. Hung.*, in press.
14. Goodall, G.J., and Filipowicz, W. (1989) *Cell*, **58**, 473–483.
15. Wiebauer, K., Herrero, J.J., and Filipowicz, W. (1988) *Mol. Cell. Biol.*, **8**, 2042–2051.
16. Keith, B., and Chua, N. (1986) *EMBO J.*, **5**, 2419–2425.
17. Peterhans, A., Datta, S.K., Datta, K., Goodall, G.J., Potrykus, I., and Paszkowski, J. (1990) *MGG*, **222**, 361–368.
18. Tanaka, A., Mita, S., Ohta, S., Kyojuka, J., Shimamoto, K., and Nakamura, K. (1990) *Nucleic Acids Res.*, **18**, 6767–6770.
19. Goodall, G.J., and Filipowicz, W. (1991) *EMBO J.*, **10**, 2635–2644.
20. Barta, A., Sommergruber, K., Thompson, D., Hartmuth, K., Matzke, M.A., and Matzke, A.J.M. (1986) *Plant Mol. Biol.*, **6**, 347–357.
21. van Santen, V.L., and Spritz, R.A. (1987) *Gene*, **56**, 253–265.
22. Hunt, A.G., Mogen, B.D., Chu, N.M., and Chua, N.-H. (1991) *Plant Mol. Biol.*, **16**, 375–379.
23. Hartmuth, K., and Barta, A. (1986) *Nucleic Acids Res.*, **14**, 7513–7528.
24. Sambrook, J., Fritsch, E., and Maniatis, T. E. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour University Press, Cold Spring Harbour.
25. Patzelt, E., Thalmann, E., Hartmuth, K., Blaas, D., and Kuechler, E. (1987) *Nucleic Acids Res.*, **15**, 1387–1399.
26. Pietrzak, M., Schillito, R., Hohn, T., and Potrykus, I. (1986) *Nucleic Acids Res.*, **14**, 5857–5868.
27. Hirt, H., Casari, G., and Barta, A. (1989) *Planta*, **179**, 414–420.
28. Chomczynski, P., and Sacchi, N. (1987) *Anal. Biochem.*, **162**, 156–159.
29. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K., and Green, M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
30. Streuli, M., and Saito, H. (1989) *EMBO J.*, **8**, 787–796.
31. Ulfendahl, P.J., Kreivi, J.P., and Akusjarvi, G. (1989) *Nucleic Acids Res.*, **17**, 925–938.
32. Gallego, M.E., and Nadal-Ginard, B. (1990) *Mol. Cell. Biol.*, **10**, 2133–2144.