Lariat RNA's as Intermediates and Products in the Splicing of Messenger RNA Precursors

Richard A. Padgett, Maria M. Konarska, Paula J. Grabowski
Steven F. Hardy, Phillip A. Sharp

Elucidation of the mechanism of splicing of messenger RNA (mRNA) precursors has become possible with the recent development of in vitro systems that process exogenously added RNA substrates (*1–3*). This has led to a systematic exploration of the substrate and cofactor requirements, as well as of the intermediates and products of the splicing process.

Abstract. *The splicing of messenger RNA precursors in vitro proceeds through an intermediate that has the 5′ end of the intervening sequence joined to a site near the 3′ splice site. This lariat structure, which has been characterized for an adenovirus 2 major late transcript, has a branch point, with 2′-5′ and 3′-5′ phosphodiester bonds emanating from a single adenosine residue. The excised intervening sequence retains the branch site and terminates in a guanosine residue with a 3′ hydroxyl group. The phosphate group at the splice junction between the two exons originates from the 3′ splice site at the precursor.*

In earlier work, we used a precursor RNA derived from the major late transcription unit of adenovirus 2 and a whole cell extract of uninfected HeLa cells to show that splicing requires Mg^{2+} and ATP (*1*) and that the reaction is inhibited by antisera that recognize small nuclear ribonucleoprotein particles containing U1 RNA (*4*). Splicing of substrate RNA in our system depends on recognition of a 5′ cap structure since, for example, the addition of cap analogs inhibits the reaction (*5*).

Two RNA species with an unusual structure have been isolated from the reaction and characterized (*6*). One of these species, IVS(A), is the excised intervening sequence that accumulates in molar yield with the spliced product. The other RNA species, IVS-L2(A), is a probable intermediate in the splicing reaction (Table 1). We demonstrated by hybridization and direct sequence analysis that the potential intermediate, IVS-L2(A) RNA, resulted from cleavage at the 5′ splice site of precursor RNA. The migration of IVS-L2(A) RNA on polyacrylamide gels of different porosity was aberrant relative to the migration of linear molecular size standards, an indication of an unusual topological structure (Table 1). IVS-L2(A) RNA also contained a modification of an oligonucleotide (46) produced by ribonuclease T_1 digestion and located within the intervening sequence near the 3′ splice site (Fig. 1A). The excised intervening sequence, IVS(A) RNA, shared the unusual electrophoretic behavior and modification of oligonucleotide 46 with the intermediate, IVS-L2(A) RNA. In a two-dimensional electrophoretic and chromatographic analysis (fingerprinting), the 3′ terminal oligonucleotide (48) produced by ribonuclease T_1 digestion of IVS(A) RNA migrated to a location different from that of the equivalent oligonucleotide from the precursor RNA. Limited ribonuclease digestion of IVS(A) RNA resulted in the conversion of this species to a discrete faster migrating form, suggesting that a significant portion of the RNA molecule was circular. We interpreted these results as evidence of a lariat RNA; that is, a tailed circular structure containing an RNA branch point.

We now provide further evidence for the branched lariat structure. Specifically we show (i) that the excised intervening sequence has a 3′ terminal hydroxyl group; (ii) that both IVS(A) and IVS-L2(A) RNA's contain an unusual nuclease-resistant product that appears to have a branched trinucleotide structure containing both a 2′-5′ and a 3′-5′ phosphodiester bond emanating from a single adenosine residue; (iii) that the branch is formed by joining the 5′ end of the intervening sequence to a position near the 3′ splice site; and (iv) that the phosphate group in the newly formed 3′-5′ phosphodiester bond in the spliced RNA is derived from the 3′ splice site.

Intervening Sequence Terminates in Guanosine 3′ Hydroxyl Group

In a two-dimensional fingerprint analysis of ribonuclease T_1 oligonucleotides from IVS(A) and IVS-L2(A) RNA's (*6*), the migration of oligonucleotides 46 and 48, the two ribonuclease T_1 products from the intervening sequence adjacent to the 3′ splice site (see Fig. 1A), was different from that in digests of precursor RNA. Both of these large oligonucleotides can also be readily resolved by polyacrylamide gel electrophoresis. Oligonucleotides 46 and 48 are two of the three most slowly migrating RNA's released after digestion of precursor RNA with ribonuclease T_1 (lane 1 in Fig. 1B). A similar digest of IVS(A) RNA (lane 3 in Fig. 1B) yielded more slowly migrating forms of oligonucleotides 46 and 48, denoted 46′ and 48′, respectively. Digestion of IVS-L2(A) RNA by ribonuclease T_1 (lane 2 in Fig. 1B) yielded a normal oligonucleotide 48 and a modified form of oligonucleotide 46 that comigrated with the corresponding species from IVS(A) RNA. Thus oligonucleotide 46 is probably identically modified in both RNA's, whereas oligonucleotide 48 has a different structure from that of the precursor only in IVS(A) RNA.

Oligonucleotide 48 is generated by ribonuclease T_1 cleavage at the G (guanosine) residue at the 3′ splice site (Fig. 1A). The slower migration of oligonucleotide 48′ from IVS(A) RNA could be caused by excision of the intervening sequence in a form lacking the 3′ terminal phosphate. To test this possibility, we treated the ribonuclease T_1 products of precursor, IVS(A), and IVS-L2(A)

Richard A. Padgett, Maria M. Konarska, Paula J. Grabowski, and Phillip A. Sharp are in the Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge 02139. The present address of Steven F. Hardy is State University of New York at Stony Brook, Long Island 11794. The permanent address of Maria M. Konarska is Institute of Biochemistry and Biophysics, Polish Academy of Sciences, 02-532 Warsaw, Poland.

RNA's with phosphatase to remove 3′ terminal phosphate moieties and analyzed them by electrophoresis in a polyacrylamide gel (Fig. 1C). After phosphatase treatment, oligonucleotide 48 from the precursor RNA (lane 2 in Fig. 1C) comigrated with oligonucleotide 48′ from IVS(A) RNA (lane 3). This is consistent with the above hypothesis. Furthermore, phosphatase treatment of oligonucleotide 48′ from IVS(A) RNA did not shift the mobility of this species while altering the mobility of the other ribonuclease T_1 products (lane 4 in Fig. 1C). This further suggests that the 3′ terminal G of IVS(A) RNA did not have a phosphate moiety. In total, these results strongly indicate that processing at the 3′ splice site follows the "Chambon rule" (*7*) of cleavage after the AG dinucleotide (A, adenosine) and that the terminal G residue has a 3′ hydroxyl group.

Origin of the Phosphate at the Splice Junction

The particular phosphate bond formed during splicing of the L1 and L2 exons cannot be simply identified by comparison of the sequences of precursor and product because of the presence of a G residue before both splice sites (see Fig. 2A). However, the above finding that the

Table 1. Properties of RNA's produced by the in vitro splicing reaction. Intermediates and splicing products contain subsets of the sequences in precursor RNA. L1 is the first leader exon, IVS-1 is the first intervening sequence, L2 is the second leader exon, and IVS-2 is a portion of the second intervening sequence. The sequence of the RNA is shown in Fig. 1A.

RNA	Sequence content	Nucleotides	Apparent molecular size* (nucleotides)		Modified ribonuclease T_1 oligonucleotides
			4 percent gel	10 percent gel	
Precursor	L1 ◇ IVS-1 ◇ L2 ◇ IVS-2	372	380	420	None
IVS-L2(A)	IVS-1 ◇ L2 ◇ IVS-2	331	380	1800	46
IVS(A)	IVS-1	231	270	900	46, 48†
L1-2(A)	L1 ◇ L2 ◇ IVS-2	141	140	150	None

*Relative to single-stranded DNA markers in polyacrylamide gels of the indicated percentage containing 8*M* urea. †Oligonucleotide 48 from IVS(A) RNA lacks a 3′ phosphate and hence migrates differently from the analogous ribonuclease T_1 product from precursor RNA.

A

```
7mGpppACUCUCUUCCG CAUCG CUG UCUG CG AG G G CCAG CUG UUG G G : G UG
L1>                                                          <L1   IVS-1>
            47
AG UACUCCCUCUCAAAAG CG G G CAUG ACUUCUG CG CUAAG AUUG UCAG

UUUCCAAAAACG AG G AG G AUUUG AUAUUCACCUG G CCCG CG G UG AUG CCUUUG

AG G G UG G CCG CG UCCAUCUG G UCAG AAAAG ACAAUCUUUUUG UUG UCAAG

CUUG CUG CACG UCUAG G G CG CAG UAG UCAAG G G UUUCCUUG AUG

            46                48
AUG UCAUACUUAUCCUG UCCCUUUUUUUUCCACAG : CUCG CG G UUG AG G
      (UACUAACA)                <IVS-1   L2>

ACAAACUCUUCG CG G UCUUUCCAG UACUCUUG G AUCG G AAACCCG UCG G CCUCCG

AACG : G UAAG AG CCUAG CAUG UAG AACUG G UUG ACG G CCUG G U
 <L2   IVS-2>
```

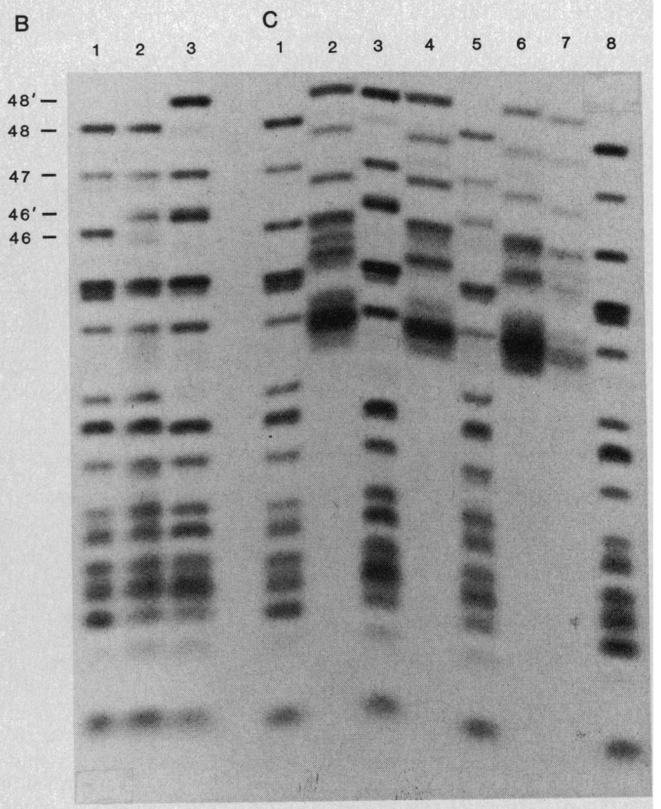

Fig. 1. (A) Sequence of the precursor RNA transcribed from the adenovirus 2 major late promoter. The sequence is separated into the various ribonuclease T_1 oligonucleotides. L1 is the first leader exon, IVS-1 is a deleted form of the first intervening sequence (*6*), L2 is the second leader exon, and IVS-2 is a portion of the second intervening sequence. The splice sites are denoted by the colons. The ribonuclease T_1 oligonucleotides referred to in the text are numbered. The conserved sequence found in yeast mRNA introns (*19*) is shown below oligonucleotide 46. (B) Separation by electrophoresis on a 15 percent polyacrylamide gel of the ribonuclease T_1 digestion products of (lane 1) precursor RNA; (lane 2) IVS-L2(A) RNA; and (lane 3) IVS(A) RNA. The oligonucleotides and their variants referred to in the text are indicated. (C) Separation of ribonuclease T_1 digestion products before or after treatment with bacterial alkaline phosphatase (BAP) to remove 3′ terminal phosphates. (Lane 1) Precursor RNA digested with ribonuclease T_1; (lane 2) precursor RNA digested with ribonuclease T_1 followed by BAP; (lane 3) IVS(A) RNA digested with ribonuclease T_1; (lane 4) IVS(A) RNA digested with ribonuclease T_1 followed by BAP; (lane 5) IVS-L2(A) RNA digested with ribonuclease T_1; (lane 6) IVS-L2(A) RNA digested with ribonuclease T_1 followed by BAP; (lane 7) precursor RNA digested with ribonuclease T_1 followed by BAP; (lane 8) precursor RNA digested with ribonuclease T_1. Methods: Precursor RNA was transcribed from pIVS-1 DNA cleaved with Bgl I in a HeLa whole cell extract in the presence of [α-^{32}P]UTP and purified by preparative polyacrylamide gel electrophoresis. The RNA was spliced by incubation in HeLa whole cell extract, and the resulting IVS-L2(A), IVS(A), and unspliced precursor RNA's were purified by preparative polyacrylamide gel electrophoresis (*6*). The purified RNA's were digested with 1 unit of ribonuclease T_1 in 20 μl of 10 m*M* tris (*p*H 7.5) and 1 m*M* EDTA for 4 hours at 37°C. The digests were loaded onto a 15 percent polyacrylamide, 8*M* urea gel in 90 m*M* tris-borate (*p*H 8.3) and 1 m*M* EDTA. For phosphatase treatment, the ribonuclease T_1–digested RNA was treated with 0.008 unit of bacterial alkaline phosphatase for 30 minutes at 55°C in the presence of 0.1 percent sodium dodecyl sulfate.

excised intervening sequence terminates in a G residue indicates that the phosphodiester bond formed during splicing is the -GpC- (C, cytidine) at the boundary of the two exons (see Fig. 2A). This newly formed phosphodiester bond is contained in the unique ribonuclease A oligonucleotide GpGpG$\overset{*}{p}$Cp (* denotes the phosphate group at the junction). As shown in Fig. 2A, if the splicing of [α-^{32}P]CTP (CTP, cytidine triphosphate) labeled precursor RNA yields radioactivity in this ribonuclease A oligonucleotide, then the junction phosphate linking the two exons must have been retained from the α-phosphate on the C residue at the 3′ splice site. Figure 2, B and C, shows the fingerprint analysis of ribonuclease A digestions of [α-^{32}P]CTP–labeled L1 and L2 RNA's, and L1-2 spliced RNA, respectively. Clearly, the junction oligonucleotide GpGpG$\overset{*}{p}$Cp (indicated by the arrow in Fig. 2C) is labeled in the two-dimensional analysis of L1-2 RNA but is absent from the mixture of L1 and L2 RNA's. Thus, the phosphate moiety in the bond joining the two exons originated from the α position of C of the L2 exon. This conclusion is consistent with the presence of a 3′-terminal hydroxyl group on the excised intervening sequence, IVS(A) RNA.

IVS(A) and IVS-L2(A) RNA's Contain a Branched Structure

We suggested earlier (*6*) that both IVS(A) and IVS-L2(A) RNA's contained a branched RNA structure similar to that described by Wallace and Edmonds (*8*). When ribonuclease T_2 digests of uniformly ^{32}P-labeled precursor, IVS-L2(A) and IVS(A) RNA's were resolved by two-dimensional thin-layer chromatography, a novel oligonucleotide species was observed in both IVS-L2(A) and IVS(A) RNA's (Fig. 3, A and B, respectively). Digestion of IVS(A) and IVS-L2(A) RNA's with nuclease P_1 also produced a novel oligonucleotide species that migrated more slowly than the nucleoside monophosphates. The ribonuclease T_2– and nuclease P_1–resistant products were eluted from the cellulose plates and further analyzed by DEAE–thin-layer chromatography to determine their net charges (Fig. 4A). The ribonuclease T_2 product (lane 3) had a net charge of −6 as predicted for a branched structure containing two phosphodiester bonds and two 3′ phosphomonoesters (Fig. 4A). The net charge on the nuclease P_1 product (lane 4) was −4 as predicted for the 5′ phosphate containing structure (Fig. 4A). All of the phosphates contained in these structures participate in phosphodiester bonds in the intact RNA molecule, as no change in the nuclease-resistant products is seen when the RNA is treated with bacterial alkaline phosphatase before nuclease digestion. The fact that the ribonuclease T_2 and nuclease P_1 products have different net charges is strong evidence for a structure with a different number of 3′ ends as compared to 5′ ends—hence a branched structure.

To further establish that the nuclease-resistant product is a branched trinucleotide, we treated the nuclease P_1–resistant product with sodium periodate and then with aniline to remove the 2′ and 3′ nucleosides by β-elimination. This should yield a 2′,3′,5′ nucleoside triphosphate. The products of such a reaction were analyzed by two-dimensional thin-layer chromatography in the presence of unlabeled markers (Fig. 4B). The labeled species comigrated with the 2′,3′,5′ adenosine triphosphate marker (*9*) (obtained from U. L. RajBhandary). The chromatography system used can discriminate between nucleotides phosphorylated on the 2′ versus the 3′ position (*10*). This system will also distinguish 2′,3′,5′ nucleoside triphosphates containing different bases, and thus the nuclease P_1–resistant product is a branched trinucleotide with an A residue at the base of the branch.

Both two-dimensional fingerprint (*6*) and one-dimensional gel electrophoretic analyses of ribonuclease T_1 digests indicated that oligonucleotide 46′ of IVS(A) RNA was modified as compared to the equivalent oligonucleotide in precursor RNA (see Fig. 1B). Secondary ribonuclease T_2 digestion of purified oligonucleotide 46′ from IVS(A) RNA released the branched oligonucleotide in anticipated amounts as shown in Fig. 3C. In parallel control experiments, secondary ribonuclease T_2 digestion of other large oligonucleotides from IVS(A) RNA and oligonucleotide 46 from precursor RNA produced only the expected nucleoside monophosphates. Oligonucleotide 46 contains sequences from −19 to −32 nucleotides from the 3′ splice site, and thus the branched structure must lie within these limits.

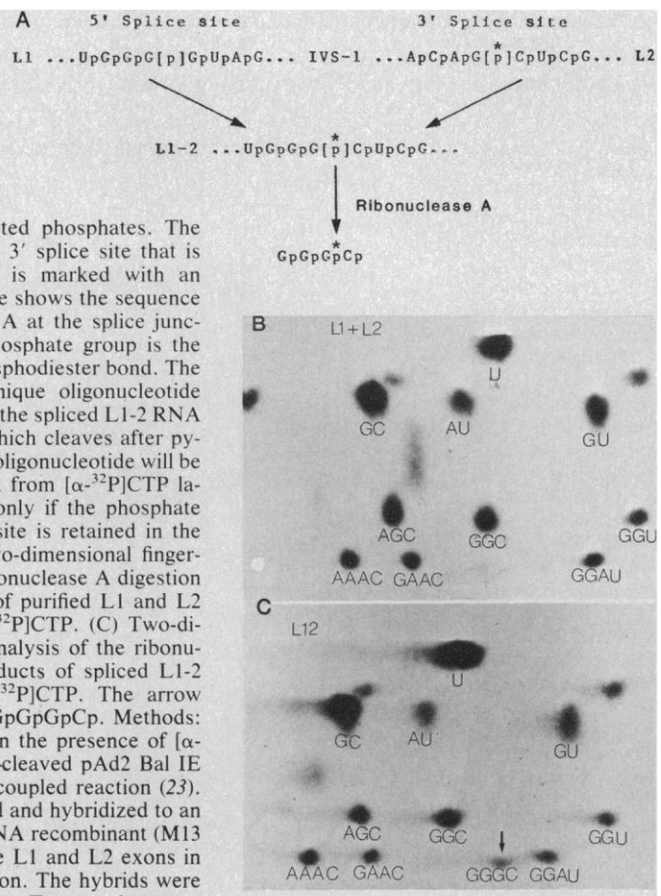

Fig. 2. (A) Strategy for the determination of the origin of the phosphate group at the splice junction. The top line shows the sequences at the 5′ and 3′ splice sites, with the splice sites denoted by the bracketed phosphates. The phosphate group at the 3′ splice site that is labeled by [α-^{32}P]CTP is marked with an asterisk. The middle line shows the sequence of the spliced L1-2 RNA at the splice junction. The bracketed phosphate group is the newly formed 3′-5′ phosphodiester bond. The last line shows the unique oligonucleotide released by digestion of the spliced L1-2 RNA with ribonuclease A, which cleaves after pyrimidine residues. This oligonucleotide will be labeled in spliced RNA from [α-^{32}P]CTP labeled precursor RNA only if the phosphate group at the 3′ splice site is retained in the splice junction. (B) Two-dimensional fingerprint analysis of the ribonuclease A digestion products of a mixture of purified L1 and L2 exons labeled with [α-^{32}P]CTP. (C) Two-dimensional fingerprint analysis of the ribonuclease A digestion products of spliced L1-2 RNA labeled with [α-^{32}P]CTP. The arrow shows the position of GpGpGpCp. Methods: RNA was transcribed in the presence of [α-^{32}P]CTP from Bam HI–cleaved pAd2 Bal IE DNA and spliced in a coupled reaction (*23*). The RNA was extracted and hybridized to an M13 complementary DNA recombinant (M13 pJAW43) containing the L1 and L2 exons in their spliced configuration. The hybrids were digested with ribonuclease T_1 to produce protected L1 and L2 exons from unspliced RNA and protected L1-2 spliced RNA. These RNA's were released from the DNA and separated by preparative polyacrylamide gel electrophoresis (*23*). The purified spliced L1-2 RNA and a mixture of unspliced L1 and L2 exons were digested with 0.6 μg of ribonuclease A for 1 hour at 37°C. The oligonucleotides were separated in two dimensions by electrophoresis at *p*H 3.5 followed by homochromatography on DEAE cellulose plates (*24, 25*).

The 5′ End of the Intervening Sequence Is Linked to Oligonucleotide 46

Limited ribonuclease digestion showed that a significant portion of the IVS(A) RNA molecule has a circular topology (*6*). As concluded above, IVS(A) RNA also has a free 3′ terminus and contains a branched structure. Therefore, it is reasonable to postulate that the 5′ end of the intervening sequence is joined by a 2′-5′ phosphodiester bond to a nucleotide within oligonucleotide 46′ to produce a lariat type of structure. Furthermore, since IVS-L2(A) and IVS(A) RNA's (i) both contain the same branched structure in oligonucleotide 46′, (ii) both show unusual electrophoretic mobilities, and (iii) both lack the L1 exon, it is probable that both RNA's have a lariat structure differing only in the length of the 3′ tail. To test whether oligonucleotide 46′ was joined to the 5′ end of the intervening sequence, we hybridized IVS-L2(A) and IVS(A) RNA's to a single-stranded M13 recombinant (XH11) that was complementary to sequences in the L1 exon and 151 nucleotides at the 5′ end of the intervening sequence (Fig. 5A). The hybrids were then treated with ribonuclease T_1 to degrade single-stranded RNA. The protected RNA was released from the hybrid, gel-purified, and digested to completion with ribonuclease T_1. The products of the final ribonuclease T_1 digestion were resolved on a polyacrylamide gel and assayed for the retention of oligonucleotide 46′. The M13 DNA should protect from ribonuclease T_1 digestion the 5′ end of the intervening sequence and any RNA linked to it in a ribonuclease T_1–resistant bond (Fig. 5A). If the lariat model is correct, oligonucleotide 46′ from the 3′ part of the

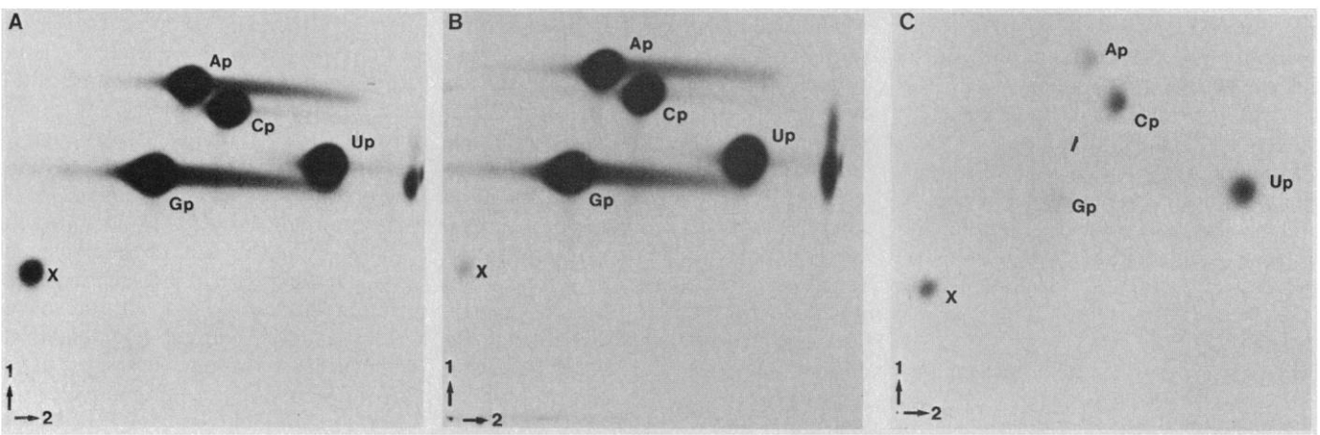

Fig. 3. Two-dimensional thin-layer chromatographic analysis of ribonuclease T_2 digestions of (A) IVS(A) RNA; (B) IVS-L2(A) RNA; (C) ribonuclease T_1 oligonucleotide 46′ from IVS(A) RNA. The positions of the nucleoside 3′ monophosphates and the ribonuclease T_2-resistant structure (X) are marked. Methods: IVS(A) and IVS-L2(A) RNA's labeled with all four α-^{32}P nucleoside triphosphates were prepared as in Fig. 1. Oligonucleotide 46′ was eluted with 1*M* triethanolamine bicarbonate from a two-dimensional fingerprint of ribonuclease T_1–digested IVS(A) RNA. The RNA's were digested with 1 unit of ribonuclease T_2 for 16 hours at 37°C in 20 μl of 20 m*M* ammonium acetate (*p*H 4.5) and separated on cellulose plates in a mixture of isobutyric acid, concentrated NH_4OH, and H_2O (45:3:30) in the first dimension and *t*-butyl alcohol, concentrated HCl, and H_2O (70:15:15) in the second dimension (*24*).

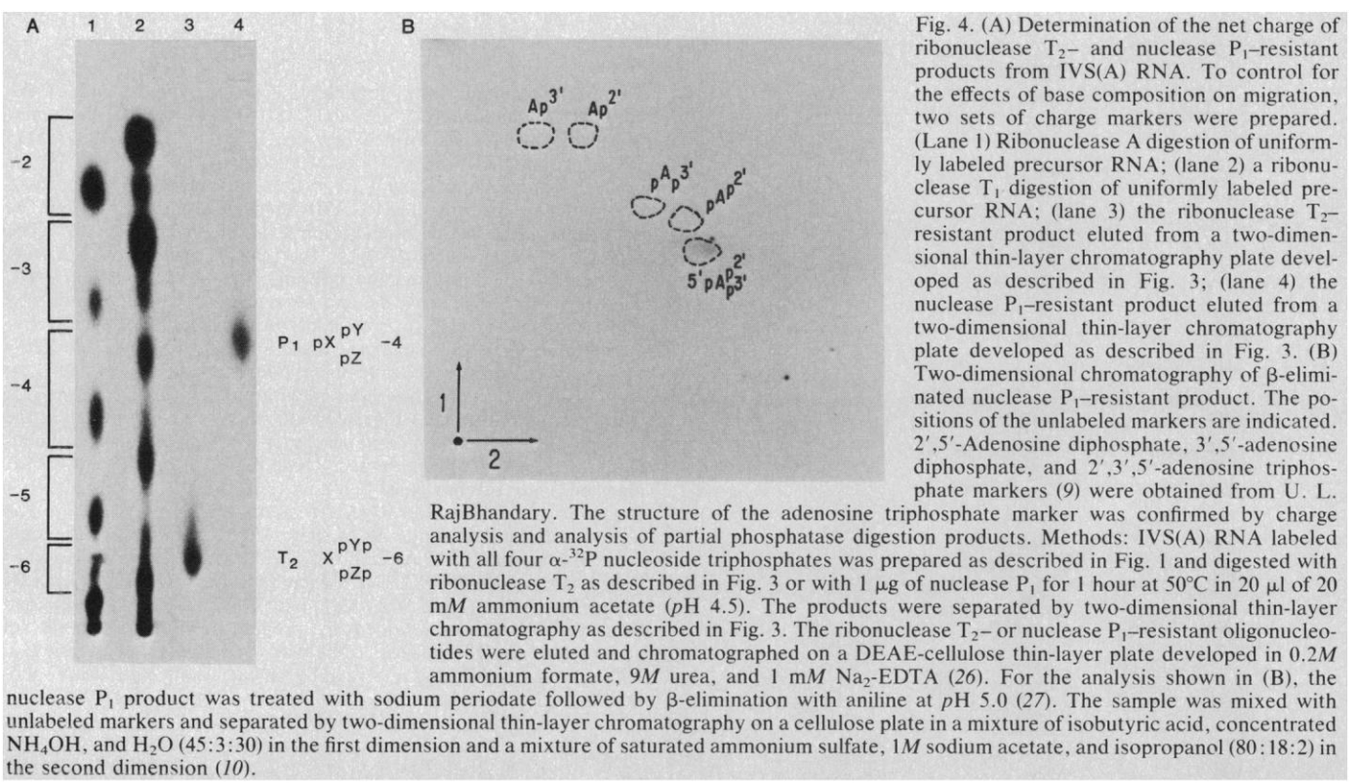

Fig. 4. (A) Determination of the net charge of ribonuclease T_2– and nuclease P_1–resistant products from IVS(A) RNA. To control for the effects of base composition on migration, two sets of charge markers were prepared. (Lane 1) Ribonuclease A digestion of uniformly labeled precursor RNA; (lane 2) a ribonuclease T_1 digestion of uniformly labeled precursor RNA; (lane 3) the ribonuclease T_2–resistant product eluted from a two-dimensional thin-layer chromatography plate developed as described in Fig. 3; (lane 4) the nuclease P_1–resistant product eluted from a two-dimensional thin-layer chromatography plate developed as described in Fig. 3. (B) Two-dimensional chromatography of β-eliminated nuclease P_1–resistant product. The positions of the unlabeled markers are indicated. 2′,5′-Adenosine diphosphate, 3′,5′-adenosine diphosphate, and 2′,3′,5′-adenosine triphosphate markers (*9*) were obtained from U. L. RajBhandary. The structure of the adenosine triphosphate marker was confirmed by charge analysis and analysis of partial phosphatase digestion products. Methods: IVS(A) RNA labeled with all four α-^{32}P nucleoside triphosphates was prepared as described in Fig. 1 and digested with ribonuclease T_2 as described in Fig. 3 or with 1 μg of nuclease P_1 for 1 hour at 50°C in 20 μl of 20 m*M* ammonium acetate (*p*H 4.5). The products were separated by two-dimensional thin-layer chromatography as described in Fig. 3. The ribonuclease T_2– or nuclease P_1–resistant oligonucleotides were eluted and chromatographed on a DEAE-cellulose thin-layer plate developed in 0.2*M* ammonium formate, 9*M* urea, and 1 m*M* Na_2-EDTA (*26*). For the analysis shown in (B), the nuclease P_1 product was treated with sodium periodate followed by β-elimination with aniline at *p*H 5.0 (*27*). The sample was mixed with unlabeled markers and separated by two-dimensional thin-layer chromatography on a cellulose plate in a mixture of isobutyric acid, concentrated NH_4OH, and H_2O (45:3:30) in the first dimension and a mixture of saturated ammonium sulfate, 1*M* sodium acetate, and isopropanol (80:18:2) in the second dimension (*10*).

intervening sequence should be present in protected fragments of IVS-L2(A) and IVS(A) RNA's but absent from the protected fragment of the precursor RNA. As shown in Fig. 5B, this is the case. Oligonucleotide 46′ was protected in IVS(A) RNA (lane 6) and IVS-L2(A) RNA (lane 9) but was not protected in precursor RNA (lane 3). As an internal control, oligonucleotide 48 or 48′ was not protected by hybridization to the M13 recombinant. These results clearly establish that the 5′ end of the intervening sequence in both IVS(A) and IVS-L2(A) RNA's is attached to a sequence in oligonucleotide 46, almost certainly through a branched structure. At the moment, the results do not reveal the distribution of nucleotides at the branch, but the simplest hypothesis would be that the 5′ terminal G of the intervening sequence is covalently linked by a 2′-5′ phosphodiester bond to an A residue in oligonucleotide 46. This is supported by preliminary analysis suggesting that the branched trinucleotide has a composition of A, G, and U (U, uridine).

Scheme for Messenger RNA Splicing

A general scheme for the splicing of mRNA precursors is suggested by these results (Fig. 6). This scheme is primarily based upon studies in vitro but is fully consistent with results obtained in vivo. The precursor RNA is first probably organized into an RNA-protein structure that is currently ill-defined but most likely has U_1 ribonucleoprotein bound to the 5′ splice site (*11*). We propose that formation of this complex in vitro accounts for both the pronounced lag in the reaction (*1*) and the requirement for cap recognition (*5*). A number of in vivo studies suggest that the structure formed by this complex selects both the 5′ and 3′ splice sites before cleavage at the former site [see, for example, (*12*)].

The metastable intermediate produced by reaction 1 of Fig. 6 contains a free 5′ exon (L1) and a second RNA with the 5′ end of the intervening sequence joined through a 2′-5′ phosphodiester bond to a site adjacent to the 3′ splice site. The 2′-5′ phosphodiester bond produces a ribonuclease T_2– and nuclease P_1–resistant trinucleotide, which constitutes the branch point of a lariat structure. Whether formation of the 2′-5′ phosphodiester bond and cleavage at the 5′ splice site are concerted reactions is not clear. It is possible that the 2′ hydroxyl group directly participates in a transesterification reaction producing the RNA's diagrammed in Fig. 6B. Alternatively, cleavage at the 5′ splice site and ligation to the 2′ position could occur through a series of reactions too rapid to produce detectable intermediates. The accumulation of an intermediate composed of two RNA's is surprising. These two RNA's must be tightly held in a complex for subsequent efficient joining of the exons. The nature of this structure would be compatible with the reaction of the free 5′ exon with a 3′ splice site on another precursor RNA in a rare transplicing reaction. Two reservations should be stated concerning this portion of the scheme: (i) only kinetic evidence suggests that the RNA species diagrammed in Fig. 6B are intermediates and (ii) the rate-limiting step in the splicing reaction in vivo may be different from that in vitro.

The second step diagrammed in Fig. 6, which joins the two exons, is the formation of a bond between the 3′ hydroxyl group of the first exon and the phosphate group at the 3′ splice site. This step could also consist of either a series of reactions or a concerted transesterification reaction. The fact that the spliced product contains the phosphate group from the 3′ splice site distinguishes this reaction from the transfer RNA splicing processes (*13–15*), but is consistent with the ribosomal RNA self-catalyzed splicing process seen in *Tetrahymena* (*16*). In this latter process, the 3′ hydroxyl group from the 5′ exon attacks the phosphate group at the 3′ splice site in a transesterification reaction to produce the spliced exons and the excised intervening sequence. This reaction is similar to the second step in Fig. 6. Alternatively, step 2 of Fig. 6 could be catalyzed by enzymes similar to those characterized in bacterial systems.

The excised intervening sequence RNA homogeneously terminates in $-AG_{OH}$, and thus splicing occurs by cleavage and joining at the precise

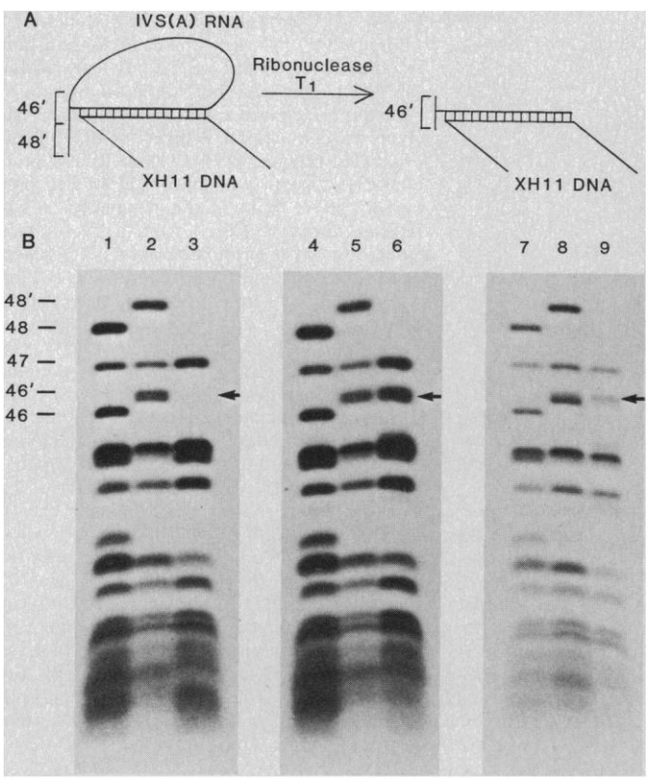

Fig. 5. Attachment of oligonucleotide 46′ to the 5′ end of the intervening sequence in IVS(A) and IVS-L2(A) RNA's. (A) Schematic showing the hybrid formed between IVS(A) and the XH11 M13 DNA recombinant. The locations of oligonucleotides 46′ and 48′ are marked. After ribonuclease T_1 digestion, oligonucleotide 46′ should be retained on the protected RNA fragment since the DNA covers the G residue nearest the branch point in oligonucleotide 46′. The presence of oligonucleotide 46′ in the purified protected RNA fragment can be demonstrated by releasing the RNA from the DNA, digesting with ribonuclease T_1, and separating the oligonucleotides by polyacrylamide gel electrophoresis. (B) Analysis of the RNA oligonucleotides protected from nuclease digestion by hybridization to XH11 DNA. The positions of the oligonucleotides are indicated. (Lanes 1, 4, and 7) Total precursor RNA digested with ribonuclease T_1; (lanes 2, 5, and 8) total IVS(A) RNA digested with ribonuclease T_1; (lane 3) the protected RNA fragment from precursor RNA digested with ribonuclease T_1; (lane 6) the protected RNA fragment from IVS(A) RNA digested with ribonuclease T_1; and (lane 9) the protected RNA fragment from IVS-L2(A) RNA digested with ribonuclease T_1. The position of oligonucleotide 46′ is indicated by the arrows. Methods: Purified precursor, IVS(A), and IVS-L2(A) RNA's were prepared as in Fig. 1, hybridized to XH11 DNA, and treated with ribonuclease T_1 to digest single-stranded RNA (*23*). The protected RNA was released from the hybrid and purified by preparative polyacrylamide gel electrophoresis. The purified RNA was then digested to completion with ribonuclease T_1 and analyzed by polyacrylamide gel electrophoresis as in Fig. 1.

boundaries of the intervening sequence as defined by the "Chambon rule" (*7*) of :GU and AG: (colons indicate cleavage sites). The significance of the lariat structure of the intervening sequence is unclear. Since intervening sequences are rapidly degraded in vivo, whereas exon sequences are comparatively stable, it is possible that the branched structure in the released intervening sequence represents a signal for degradation. However, other features such as nonpolyadenylated 3′ termini or release from an RNA-protein complex could also trigger degradation.

The finding that the 5′ end of the intervening sequence is linked to an oligonucleotide near its 3′ end indicates that the sequences specifying a functional 3′ splice site may be larger than the stretch of 18 pyrimidines followed by CAG, determined by sequence comparisons (*17, 18*). In splicing of yeast mRNA precursors, a required sequence, UACUAACA, is found upstream from the 3′ splice site (*19*). Pikielny *et al.* (*20*) have shown that yeast nuclear RNA is modified at this site as assayed by blockage of reverse transcriptase. It is possible that this sequence is the site of branch formation in yeast precursors. Recently homologs of the yeast sequence have been recognized in mRNA precursors from other eukaryotes (*21*). Similar sequences are found in the branch-containing oligonucleotide identified in this study (Fig. 1A). The sequence specificity of this site in mammalian cells must be rather loose. Wieringa *et al.* (*22*) have generated β-globin constructions with intron deletions up to 16 nucleotides from the 3′ splice site; these constructions remain active for splicing. All sequences upstream from the pyrimidine-rich tract are

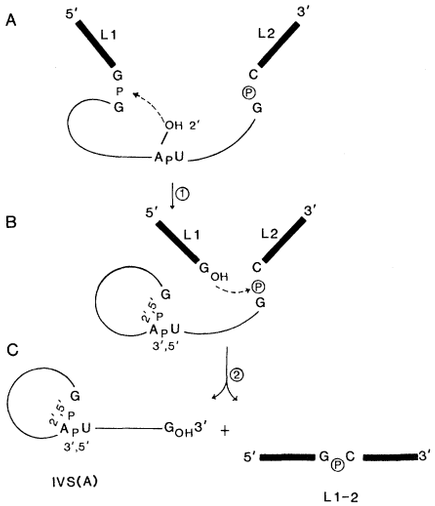

Fig. 6. Proposed scheme for mRNA splicing. L1 and L2 are the first and second leader exons, respectively. L1-2 is the spliced form of the two leader exons. The reactions and structures of the RNA's are described in the text.

deleted and an Xba I linker sequence is placed at the point of deletion. It is possible that an alternative site for branch formation is used in these mutants or that branch formation is not a strict requirement for splicing. However, the quantitative production of the branched intervening sequence in the in vitro reaction suggests that branch formation is required for splicing. If the site of branch formation is not sequence specific then it must be structure specific. The structure of the complex formed by the precursor RNA and other components in the system may play a critical role in determining the specificity of splicing. This hypothetical complex could also specify the site of branch formation.

References and Notes

1. S. F. Hardy, P. J. Grabowski, R. A. Padgett, P. A. Sharp, *Nature (London)* **308**, 375 (1984).
2. N. Hernandez and W. Keller, *Cell* **35**, 89 (1983).
3. A. R. Krainer, T. Maniatis, B. Ruskin, M. R. Green, *ibid.* **36**, 993 (1984).
4. R. A. Padgett, S. M. Mount, J. A. Steitz, P. A. Sharp, *ibid.* **35**, 101 (1983).
5. M. M. Konarska, R. A. Padgett, P. A. Sharp, *ibid.*, in press.
6. P. J. Grabowski, R. A. Padgett, P. A. Sharp, *ibid.* **37**, 415 (1984).
7. R. Breathnach, L. Benoist, K. O'Hare, F. Gannon, P. Chambon, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4853 (1978).
8. J. C. Wallace and M. Edmonds, *ibid.* **80**, 950 (1983).
9. J. G. Moffatt and H. G. Khorana, *J. Am. Chem. Soc.* **83**, 663 (1961).
10. M. M. Konarska, W. Filipowicz, H. J. Gross, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 1474 (1982).
11. S. M. Mount, I. Pettersson, M. Hinterberger, A. Karmar, J. A. Steitz, *Cell* **33**, 509 (1983).
12. T. Kühne, B. Wieringa, C. Weissmann, *EMBO J.* **2**, 727 (1983).
13. C. L. Greer, C. L. Peebles, P. Gegenheimer, J. Abelson, *Cell* **32**, 537 (1983).
14. F. A. Laski, A. Z. Fire, U. L. RajBhandary, P. A. Sharp, *J. Biol. Chem.* **258**, 11974 (1983).
15. W. Filipowicz and A. J. Shatkin, *Cell* **32**, 547 (1983).
16. K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, T. R. Cech, *ibid.* **31**, 147 (1982).
17. P. A. Sharp, *ibid.* **23**, 643 (1981).
18. S. M. Mount, *Nucleic Acids Res.* **10**, 459 (1982).
19. C. J. Langford and D. Gallwitz, *Cell* **33**, 519 (1983).
20. C. W. Pikielny, J. L. Teem, M. Rosbash, *ibid.* **34**, 395 (1983).
21. E. Keller and W. Noon, personal communication.
22. B. Wieringa, E. Hofer, C. Weissmann, in preparation.
23. R. A. Padgett, S. F. Hardy, P. A. Sharp, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 5230 (1983).
24. F. A. Laski, B. Alzner-DeWeerd, U. L. RajBhandary, P. A. Sharp, *Nucleic Acids Res.* **10**, 4609 (1982).
25. M. Silberklang, A. M. Gillum, U. L. RajBhandary, *Methods Enzymol.* **59**, 58 (1979).
26. M. M. Konarska, W. Filipowicz, H. Domdey, H. J. Gross, *Nature (London)* **293**, 112 (1981).
27. P. R. Whitfeld and R. Markham, *ibid.* **171**, 1151 (1953).

28. We thank M. Samuels for providing the XH11 M13 recombinant and U. L. RajBhandary for providing unlabeled nucleotide markers. Supported by a Myron Bantrell postdoctoral fellowship (to R.A.P), a Jane Coffin Childs Memorial Fund postdoctoral fellowship (to M.M.K.), a Helen Hay Whitney postdoctoral fellowship (to P.J.G.), by NSF grant PCM-8200309, by NIH grants R01-GM32467 and PO1-CA26717 (to P.A.S.), and partially supported by NIH Core grant PO1-CA14051.

5 July 1984; accepted 24 July 1984