A functional association between the 5' and 3' splice sites is established in the earliest prespliceosome complex (E) in mammals

Susan Michaud and Robin Reed

Cellular and Molecular Physiology and Program in Cell and Developmental Biology, Harvard Medical School, Boston, Massachusetts 02115 USA

The earliest detectable mammalian prespliceosome complex (E) contains the non-snRNP splicing factor U2AF, U1 snRNP, and several spliceosome-associated proteins (SAPs). We show that specific complexes, designated E3' and E5', assemble independently on RNAs containing only a 3' or 5' splice site, respectively. U2AF is enriched in E3', whereas U1 snRNP is enriched in E5'. Using a highly sensitive substrate-competition assay, we show that both the 5' splice site and the pyrimidine tract at the 3' splice site are required for efficient E complex assembly on intact pre-mRNA. We conclude that the 5' and 3' splice sites are associated functionally as early as E complex by either direct or indirect interactions between U1 snRNP and U2AF. Our observations predict that E complex assembly is a major control point for establishing splice site selection in both constitutively and alternatively spliced pre-mRNAs.

[**Key Words:** E complex; 5' and 3' splice sites; splice site selection; splicing factors; snRNP; substrate-competition assay]

Received March 3, 1993; revised version accepted April 7, 1993.

In both yeast and mammals, the fundamental pre-mRNA splicing reaction accomplishes intron excision and precise ligation of the two flanking exons (for review, see Green 1991; Guthrie 1991; Steitz 1992; Moore et al. 1993). The first insight into the basic mechanism for juxtaposing a given pair of exons came from studies showing that the 3' splice site was required for the first catalytic step of the splicing reaction, cleavage at the 5' splice site, and lariat formation (Reed and Maniatis 1985; Ruskin and Green 1985). This observation indicated that the 5' and 3' exons are initially brought together by interactions between the 5' and 3' ends of the intron. However, little is understood about precisely when and how these interactions between the 5' and 3' splice sites are first established. This problem is especially interesting in the case of metazoan pre-mRNAs, which often have the added complexities of cryptic splice sites, multiple introns, alternative splicing, and remarkably long introns (>100,000 nucleotides), in which the 3' splice site is not even transcribed until several hours after synthesis of the 5' splice site (for review, see Green 1991; Moore et al. 1993).

In yeast, a functional association between the 5' and 3' ends of the intron has been detected in the earliest functional intermediate in spliceosome assembly, designated the "commitment complex" (Ruby and Abelson 1988; Seraphin and Rosbash 1989, 1991; for review, see Ros-

bash and Seraphin 1991; Ruby and Abelson 1991). The commitment complex is an ATP-independent complex that contains U1 small nuclear ribonucleoprotein (snRNP), and competition studies showed that both the branchpoint sequence and the 5' splice site are required for commitment complex activity (Legrain et al. 1988; Seraphin and Rosbash 1989, 1991). These observations led to the proposal that in the commitment complex, U1 snRNP is bound to the 5' splice site and contacts either directly or by another factor, the branchpoint sequence (Seraphin and Rosbash 1989). In a previous study, we showed that a commitment activity similar to that in yeast exists in mammals (Michaud and Reed 1991). This activity was attributed to an ATP-independent complex, designated E complex (early complex), which is detected as a discrete peak by gel filtration (Reed 1990; Michaud and Reed 1991). Using in vitro complementation assays, we showed that the pre-mRNA in E complex isolated by gel filtration was efficiently chased into the first ATP-dependent prespliceosome (A complex), the spliceosome (B complex), and spliced products (Michaud and Reed 1991). E complex assembles on adenovirus major late (AdML), β -globin (Reed 1990; Michaud and Reed 1991), and all other pre-mRNAs tested (S. Michaud and R. Reed, unpubl.), and the efficiency of E complex assembly on each pre-mRNA correlates with the efficiency of spliceosome assembly on that pre-mRNA (Michaud and

Reed 1991). In addition to U1 snRNP, analysis of the protein composition of affinity-purified E complex revealed the presence of the non-snRNP splicing factor U2AF and several spliceosome-associated proteins (SAPs) (Bennett et al. 1992b).

Although there are obvious similarities between the yeast commitment complex and the mammalian E complex, the pre-mRNA sequences required for this early step in mammalian spliceosome assembly have not been determined. Moreover, several ATP-independent complexes have been identified in mammals, and the relationships between these complexes, E complex, and commitment activity are not understood (Michaud and Reed 1991; Jamison et al. 1992; Jamison and Garcia-Blanco 1992). In this report we show that the functional activity that commits pre-mRNA to the spliceosome assembly pathway resides in E complex and not in other ATP-independent complexes. We also show that specific complexes, designated E3' and E5', assemble on RNAs lacking the 5' or 3' splice site, respectively. E3' is a specific precursor to the first ATP-dependent prespliceosome that assembles on an RNA lacking the 5' splice site (A3' complex). However, using a sensitive substrate-competition assay, we demonstrate that E complex assembles more efficiently on a pre-mRNA containing both 5' and 3' splice sites than on an RNA containing either splice site alone. We conclude that in mammals, as in yeast, the 5' and 3' splice sites are recognized and functionally associated with one another in E complex, prior to assembly of A complex.

Results

In initial studies we carried out experiments to determine the relationship between E complex and three other ATP-independent complexes (H, $-ATP$, and commitment) identified in mammals (Fredericewey and Keller 1985; Grabowski et al. 1985; Konarska and Sharp 1987; Reed 1990; Michaud and Reed 1991; Jamison and Garcia-Blanco 1992; Jamison et al. 1992). H complex, which assembles before any of the functional prespliceosome complexes (Michaud and Reed 1991), consists primarily of heterogeneous nuclear ribonucleoproteins (Bennett et al. 1992a) and does not require 5' or 3' splice sites for assembly (Grabowski et al. 1985; Konarska and Sharp 1987; Reed 1990). The $-ATP$ complex contains U2 snRNP and is detected as a large particle on native gels and on density gradients (Jamison and Garcia-Blanco 1992; Jamison et al. 1992). Finally, the commitment complex was identified on density gradients as an activity that commits pre-mRNA to the mammalian spliceosome assembly pathway (Jamison et al. 1992). This activity was not separated as a discrete peak from the H complex or the $-ATP$ complex (Jamison et al. 1992).

To identify each of the discrete ATP-independent complexes ($-ATP$, E, and H), a splicing reaction lacking ATP was fractionated by gel filtration (Fig. 1A), and fractions throughout the column were analyzed by native gel electrophoresis (Fig. 1B). This analysis revealed that the vast majority of pre-mRNA fractionated in the region of the native gel shown previously to contain E and H com-

Figure 1. The ATP-independent activity that commits pre-mRNA to spliceosome assembly resides in E complex. (A) ^{32}P -Labeled AdML pre-mRNA was incubated in a splicing reaction in the absence of ATP and then fractionated by gel filtration. The peaks corresponding to E and H complexes are indicated. The unlabeled peak that elutes before E complex is the void volume of the column, and the unlabeled peak that elutes after H complex is degraded RNA. (B) Fractions throughout the gel filtration column were analyzed by native gel electrophoresis. The fraction numbers are indicated at the top of each lane. The bands corresponding to E and H complexes (E+H) and the $-ATP$ complex are indicated. (C) An aliquot of each of the fractions shown in B was incubated in a complementation reaction under splicing conditions (see Materials and methods; Michaud and Reed 1991) and then fractionated by native gel electrophoresis. The fraction numbers are indicated at the top of each lane. The bands corresponding to E, H, A, and B complexes are indicated.

plexes (Michaud and Reed 1991), whereas a minor fraction of the pre-mRNA was present in the $-$ ATP complex (Fig. 1B). This complex was identified as the $-$ ATP complex (data not shown) on the basis that it assembles in the absence of ATP and fractionates slightly above A complex, as reported previously (Jamison and Garcia-Blanco 1992; Jamison et al. 1992).

To determine which of the ATP-independent complexes cofractionates with commitment activity, an aliquot of each of the fractions shown in Figure 1B was incubated in a complementation reaction under splicing conditions (Fig. 1C; Michaud and Reed 1991). The complementing extract, which does not assemble spliceosomes, or splice naked pre-mRNA, was prepared by preincubating nuclear extract with competitor RNA (see Materials and methods; Michaud and Reed 1991). Significantly, we observed that the pre-mRNA in the gel filtration fractions containing E complex was very efficiently chased into A and B complexes (Fig. 1C, fractions 44–50). The peak of commitment activity corresponds precisely with the peak of E complex (Fig. 1A). Moreover, the data show that commitment activity cannot possibly be due to the $-$ ATP complex, as the levels of this complex are >10 -fold lower than those of A and B complexes generated in the complementation reaction (Fig. 1, cf. B, lanes 46–50, with C, lanes 46–50). We conclude that the ATP-independent activity that commits pre-mRNA to spliceosome assembly resides in E complex, not in H or $-$ ATP complexes. Moreover, because the only commitment activity that we detect cofractionates with E complex, we conclude that the commitment activity, designated commitment complex (Jamison et al. 1992), corresponds to E complex. In contrast to our studies in which E complex is detected as a discrete peak by gel filtration, E (or commitment) complex is not detected as a discrete peak by density gradient sedimentation (Jamison et al. 1992). This probably accounts for the failure to detect E complex in the early studies of spliceosome assembly that also employed density gradients (Fredericewey and Keller 1985; Grabowski et al. 1985; Bindereif and Green 1986; Perkins et al. 1986).

The $-$ ATP complex was detected with a specific derivative of AdML in a previous study (Jamison and Garcia-Blanco 1992) but was not detected as an abundant complex with other pre-mRNAs tested (this study; Jamison and Garcia-Blanco 1992). It is possible that the $-$ ATP complex represents an intermediate step between E and A complexes or that it is an aberrant complex (Jamison and Garcia-Blanco 1992; Jamison et al. 1992).

Sequence requirements for E complex assembly

To determine the pre-mRNA sequence requirements for E complex assembly, we analyzed complex formation with several mutant derivatives of AdML and β -globin pre-mRNAs (Figs. 2 and 3). Equimolar amounts of each RNA were used for E complex assembly reactions. Because the specific activities of the different RNAs vary, only the ratios of the peaks (e.g., E/H ratio) between each gel filtration column profile should be compared. We

also note that the same complex can fractionate in different fractions because of differences between the gel filtration columns. However, the identity of the complexes was confirmed by determining their protein or snRNA compositions, or both (see below).

To determine whether 3' splice-site sequences are essential for E complex assembly, we examined two different 5' half-substrate RNAs, AdML5' and β 5', which contain exon 1 and the 5' portion of the intron but lack the 3' half of the intron and exon 2 (Fig. 2). E complex elutes between the void volume of the column and the H complex (see Fig. 1A). We observed a discrete peak in the position of E complex with both AdML5' and β 5' RNAs (Fig. 2). In contrast, this complex was not detected or was barely detected with three other RNAs, including a non-specific RNA lacking splice sites (NS2), a 5' half substrate with a 5' splice-site deletion ($\beta\Delta$), and a non-specific RNA (NS1) containing a sequence (CAGGTGAGC) that is an 8/9 match to the 5' splice-site consensus sequence (CAGGTGAGT) (Fig. 2). On the basis of these data, we have come to the following conclusions. (1) Specific ATP-independent complexes assemble on RNAs containing only the 5' portion of the pre-mRNA. We have designated these complexes E5', and they are further characterized below. (2) The 5' splice site is required for E5' assembly. The smaller shoulder that elutes before H complex with $\beta\Delta$ RNA is most likely the result of the presence of a cryptic 5' splice site in this RNA. This cryptic 5' splice site is activated to low levels when the normal 5' splice site is deleted (Krainer et al. 1984, 1990b). (3) The 5' splice site alone does not appear to be sufficient for E5' assembly. The observation that E5' was not detected on a non-specific RNA containing a 5' splice site (NS1) suggests that the sequence context of the 5' splice site may play an important role in E5' complex assembly.

The studies presented above show that specific ATP-independent complexes assemble on the 5' portion of the intron. To determine whether a specific complex could be detected on the 3' splice site under these conditions, we examined 3' half-substrates, which contain only the 3' portion of the intron (including the branchpoint sequence and the 3' splice site) and exon 2 (Fig. 3A). Interestingly, although efficient complex assembly was detected with AdML3', only very low levels of complex assembly were observed with β 3'. AdML3' does not contain any sequences likely to serve as cryptic 5' splice sites, suggesting that efficient complex assembly on AdML3' is attributable to the 3' splice site sequences. We have designated this complex E3'.

To test the possibility that differences in 3' splice-site sequences are responsible for the higher levels of E3' complex observed with AdML3' than with β 3', we examined complex assembly on Ad/ β 3'. This RNA contains the 3' intron portion of AdML3' fused precisely to exon 2 of β 3'. As shown in Fig. 3A, the levels of E3' were significantly increased with Ad/ β 3' relative to β 3', although not to the levels observed with AdML3'. These data indicate that either exon sequence or length (AdML3' exon 2 is 45 nucleotides; β 3' exon 2 is 208

Figure 2. A specific ATP-independent complex (E5') assembles on the 5' splice site. The RNAs indicated above each column profile were incubated for 20 min in splicing reactions (100 μ l) lacking ATP and magnesium and then fractionated by gel filtration. Representative column profiles are shown. The peaks containing E5' and H complexes are indicated for each column profile. The peak in fractions 25–35 in each profile is the void volume, whereas the peak in fractions 65–75 is degraded RNA. Note that the large void volumes seen with AdML5' and β 5' are characteristic of the particular columns used, not of these RNAs.

nucleotides; see Materials and methods) is important for efficient complex assembly.

To identify the specific intron sequences in AdML3' that could account for the efficient E3' assembly, we compared the 3' splice-site sequences of AdML3' (CCC-UUUUUUUUCCACAG) and β 3' (UAUUUUCCCAC-CCUAG). One notable difference between these RNAs is the presence of eight contiguous uridine residues in AdML3'. To determine directly whether these nucleotides are important for efficient complex assembly, we constructed a derivative of AdML3' in which these eight uridines were substituted with random sequence (CCCGUGAUCACCCACAG). A dramatic decrease in the levels of E3' was observed with this mutant (AdML Δ ; Fig. 3A). We conclude that efficient E3' complex assembly requires the pyrimidine tract. Further studies are required to determine whether pyrimidine content, or uridine content in particular, is important for assembly.

Although E3' assembly is specific and efficient on AdML3', we questioned the functional significance of E3' because assembly is so inefficient on β 3'. If the interactions that occur during E3' assembly are essential for subsequent steps in the splicing reaction, then β 3' RNA should be significantly less efficient in these steps than AdML3'. Previous studies using native gels have shown that an ATP-dependent complex, designated A complex, assembles on 3' half substrates (Konarska and Sharp 1986). To distinguish between the A complex that

is assembled on intact pre-mRNA and that assembled on 3' half-substrates, we refer to the latter complex as A3' (Bennett et al. 1992b). To determine whether the efficiency of E3' assembly correlates with A3' assembly, we used the same RNAs shown in Figure 3A to examine A3' assembly. This analysis revealed dramatic differences in the efficiency of A3' assembly with the different RNAs (Fig. 3B). As was observed with E3' assembly, β 3' RNA assembles A3' complexes poorly, whereas AdML3' RNA assembles A3' very efficiently (Fig. 3B, lanes 2,3). Similarly, dramatically decreased levels of A3' assembly were observed with AdML Δ (Fig. 3B, lane 4), as expected from previous studies (Konarska and Sharp 1986). Finally, A3' assembly, as with E3' assembly, was intermediate in efficiency with Ad/ β 3' RNA (Fig. 3B, lane 1).

The data presented above show that there is a correlation between the efficiency of E3' and A3' assembly, suggesting that E3' may be a precursor to A3'. To determine this directly, we carried out *in vitro* complementation studies (Fig. 3C). S-100 extracts were used for the complementation reaction (Krainer and Maniatis 1985). These extracts lack the essential splicing factor SF2/ASF and, therefore, do not splice naked pre-mRNA (Krainer and Maniatis 1985; Krainer et al. 1990a,b). To determine whether E3' complex could be chased into A3' complex, AdML3' RNA (containing the branchpoint sequence and the 3' splice site) was incubated in splicing extracts in the absence of ATP, and a small aliquot of this reaction

Figure 3. An ATP-independent complex, E3', assembles specifically on the 3' splice site and is a precursor to the ATP-dependent complex, A3'. (A) Fractionation of E3' complexes by gel filtration. The RNAs indicated above each column profile were incubated for 20 min in splicing reactions (100 μ l) lacking ATP and magnesium and then fractionated by gel filtration. Representative column profiles are shown. The peaks containing E3' and H complexes are indicated for each column profile. The peak in fractions 25–35 in each profile is the void volume, and the peak in fractions 65–75 is degraded RNA. (B) Fractionation of A3' complexes by native gel electrophoresis. RNAs were incubated for 5 min in splicing reactions containing ATP and magnesium and then fractionated by native gel electrophoresis (lane 1, Ad/β; Ad/β3'; lane 2, β; β3'; lane 3, Ad; AdML3'; lane 4, AdΔ; AdMLΔ). The bands corresponding to A3', E3', and H complexes are indicated. (C) In vitro complementation of E3' complexes. (Lane 1) A mock (mck) reaction (see text and Materials and methods), in which nuclear extract and AdML3' RNA were not incubated together before an 87.5-fold dilution into a splicing reaction containing S-100. (Lanes 2–4) E complex, assembled by incubating AdML3' RNA with nuclear extract, before an 87.5-fold dilution into a splicing reaction containing nuclear extract (NE) (lane 2), S-100 (lane 3), or buffer (Buf) (lane 4). Complementation reactions were carried out for 20 min and then fractionated by native gel electrophoresis. The bands corresponding to A3', H, and E3' complexes are indicated (left of B).

was diluted into the complementation reaction containing S-100 (see Materials and methods). As a control, a mock reaction was carried out in which AdML3' RNA and the extract were not incubated together before their dilution into the complementation reaction. The same strategy was used to show that E complex assembled on intact pre-mRNA could be chased into A and B complexes (Michaud and Reed 1991). As was observed previously with E complex (Michaud and Reed 1991), E3' and H complexes cofractionate by native gel electrophoresis (data not shown).

When a reaction containing E3' complex was diluted into a complementation reaction containing normal nuclear extract, A3' complex was assembled efficiently (Fig. 3C, lane 2). In contrast, little or no A3' complex was detected when the control mock complex was diluted into a complementation reaction containing S-100 (lane 1), or when the reaction containing E3' complex was incubated in buffer alone (lane 4). However, A3' complex was readily detected when the reaction containing E3' complex was incubated in a complementation reaction containing S-100 (lane 3). These data indicate that the ATP-independent E3' complex is a precursor to the ATP-

dependent A3' complex. Complex assembly is considerably less efficient in S-100 than in nuclear extract, possibly because S-100 has lower concentrations of splicing factors (Krainer et al. 1990a).

SnRNA and protein compositions of E, E5', and E3' complexes

Our data show that specific ATP-independent complexes, E3' and E5', assemble on RNAs lacking either a 5' or a 3' splice site, respectively. These complexes were not detected in previous studies because, unlike A3' complex, E3' and E5' are not stable in the native gel systems used (Konarska and Sharp 1986, 1987; S. Michaud and R. Reed, unpubl.). To determine the relationships between E complex assembled on intact pre-mRNA and E5' and E3' complexes, we compared the snRNA and protein compositions of highly purified E, E5', and E3' complexes (Figs. 4 and 5). For purification, complexes were assembled on biotinylated pre-mRNA, isolated by gel filtration and affinity-purified in 250 mM salt by binding to avidin-agarose (see Materials and methods; Michaud and Reed 1991; Bennett et al. 1992b).

Figure 4. SnRNA compositions of purified complexes. Splicing complexes were assembled on ^{32}P -labeled AdML pre-mRNA and derivatives of AdML containing only a 5' (5'Ad, AdML5') or a 3' (3'Ad, AdML3') splice site (see Materials and methods). Complexes were fractionated by gel filtration, affinity-purified by binding to avidin-agarose, and washed in 250 mM salt (see Materials and methods). RNA was isolated from equivalent amounts of each of the purified complexes, end-labeled with [^{32}P]pCp and RNA ligase, fractionated on an 8% polyacrylamide gel, and then visualized by PhosphorImage analysis. As a control, the purification was carried out with complexes assembled on nonbiotinylated RNAs. The positions of the snRNAs and the RNA substrates are indicated at left. The band labeled X is present in some preparations of nuclear extract and may correspond to a breakdown product of U1 snRNA (see Michaud and Reed 1991). (Lane 1) B complex + biotin; (lanes 2,6) E complex + biotin; (lane 3) A5' complex + biotin; (lane 4) A3' complex + biotin; (lane 5) A3' complex - biotin; (lanes 7,9) E3' complex + biotin; (lane 8) E5' complex + biotin; (lane 10) E3' complex - biotin; (lanes 11-13) markers for the RNA substrates: (pre-mRNA) AdML; (Ad3') AdML3'; (Ad5') AdML5'.

Analysis of the snRNA compositions of equimolar amounts of purified E and E5' complexes revealed similar levels of U1 snRNA (Fig. 4, lane 2 or 6 and 8; Michaud and Reed 1991). In contrast, much lower levels of U1 snRNA are present in E3' complex (cf. lanes 6-8). Note that the levels of U1 snRNA in E3' complex are above the background levels observed when E3' purification is carried out with a nonbiotinylated RNA as a control (cf. lanes 9 and 10). The levels of U1 snRNA detected in E, E5', and E3' complexes purified in 250 mM salt are similar to the levels detected in 150 mM salt (data not shown). As expected, U2 snRNA is present at high levels in A3' complex (Fig. 4, lane 4) and U2, U4, and U5 snRNAs are enriched in B complex [lane 1; note that U6 snRNA is also enriched in B complex (see below) but does not label efficiently with RNA ligase and pCp]. The low levels of U2, U4, and U5 snRNAs observed in A5' complex suggest the possibility that A5' may be a pseudospliceosome-like complex similar to that shown to assemble on an RNA oligonucleotide containing a consen-

sus 5' splice site (Hall and Konarska 1992). The levels of U1 snRNA are much lower in A3', A5', and B complexes relative to the corresponding ATP-independent complexes, E3', E5', and E (Fig. 4, cf. lanes 1-8). This observation is addressed further below.

The finding that U1 snRNA is the only snRNA specifically enriched in E, E5', and E3' complexes compared with the ATP-dependent complexes is consistent with the idea that E5' and E3' are related to E complex. To further investigate the relationships between these three complexes, we carried out two-dimensional gel analysis of the protein compositions of the purified complexes (Fig. 5). As reported previously (Bennett et al. 1992b), E complex contains the U1 snRNP-specific proteins A and 70K, the snRNP core proteins B and B', the non-snRNP splicing factor U2AF, and five SAPs (42, 72, 88, 92, and 115 kD; the U1 70K does not stain well with silver). The unlabeled proteins in each complex were shown not to be associated specifically with the purified complexes in our previous study (Bennett et al. 1992b). Three of the SAPs (42, 72, and 88 kD) are present in similar levels in both E5' and E3' complexes (Fig. 5). In contrast, the 115-kD SAP and the U1 snRNP A protein are enriched in E5', whereas U2AF is enriched in E3' (Fig. 5). Further studies are needed to understand the significance of the 115-kD SAP in these complexes, as this protein is absent or present in low levels in complexes assembled on other pre-mRNAs (Bennett et al. 1992b; D. Staknis and R. Reed, unpubl.). On the basis of the protein and snRNA compositions of E5' and E3', we conclude that these two complexes are related to E complex assembled on intact pre-mRNA.

Release of U1 snRNP during spliceosome assembly

In our comparison of the snRNA compositions of the ATP-dependent and ATP-independent complexes, we found that U1 snRNA is present at much lower levels in the ATP-dependent complexes (A5', A3', and B; Fig. 4, cf. lanes 1-8). A similar decrease in the levels of U1 snRNA between E and B complexes was observed previously (Michaud and Reed 1991). However, in the present study and that reported previously (Michaud and Reed 1991), the snRNAs were detected by 3' end-labeling with [^{32}P]pCp, which labels different RNAs with very different efficiencies. Thus, we could not distinguish between the possibilities that there was >1 mole of U1 snRNA per mole of E complex, or <1 mole of U1 snRNA per mole of A or B complex. To address this question, we used ethidium bromide staining to compare the snRNA levels in E, A, and B complexes (Fig. 6). Interpretation of these data is complicated because H complex contaminates both E complex and A complex assembled for short periods of time (2 min; Michaud and Reed 1991). As a consequence, the levels of snRNAs in these complexes appear lower than expected when compared with the pre-mRNA. In any case, analysis of E complex clearly shows that there is no more than 1 mole of U1 snRNA per mole of E complex (Fig. 6, lane 1). Moreover, U1 snRNA is substoichiometric relative to the other

Figure 5. Protein compositions of E, E5', and E3' complexes. Affinity-purified E, E5', and E3' complexes were fractionated by two-dimensional gel electrophoresis and silver stained (see Materials and methods; Bennett et al. 1992b). Sizes of molecular mass markers are shown (in kD). The SAPs (42, 72, 88, 92, and 115 kD), the snRNP core proteins B and B', the U1 snRNP-specific protein A, and U2AF are indicated. (Δ) The heat shock proteins, Hsp70/ Hsc70, which are highly abundant in the nuclear extract and may be present in the affinity-purified splicing complexes as contaminants (Bennett et al. 1992b).

snRNAs in B complex (Fig. 6, lane 3). On the basis of these data, we conclude that U1 snRNP is present in E complex but less tightly associated with the pre-mRNA in B complex. Immunoprecipitation studies detect U1 snRNP in B complex (Abmayr et al. 1988), indicating that this snRNP remains loosely bound to the complex.

Figure 6. SnRNA compositions of E, A, and B complexes. RNA was isolated from affinity-purified E, A, and B complexes, fractionated on an 8% denaturing gel, and stained with ethidium bromide. Each lane contains 200 ng of AdML pre-mRNA assembled into the indicated complex. The pre-mRNA and the snRNAs (U1, U2, U4, U5, and U6) are indicated at left. (Lane 1) E complex; (lane 2) A complex; (lane 3) B complex; (lane 4) 200 ng of AdML pre-mRNA.

Whether this is a functional association or not remains to be determined.

We believe that the destabilization of U1 snRNP occurs after A complex assembly because equimolar amounts of U1 and U2 snRNAs are present in A complex assembled for only 2 min (Michaud and Reed 1991). A similar result was obtained when A complex was assembled in U4/U6 (Barabino et al. 1990) or U5 snRNP-depleted extracts (Lamm et al. 1991). We have found that A complex is very short lived when assembled in normal extracts on intact pre-mRNA. Thus, although there is less U1 snRNA than U2 snRNA in the A complex shown in Figure 6 (lane 2), this complex was assembled for 5 min and contains some U4, U5, and U6 snRNAs. This indicates that a portion of the pre-mRNA has assembled into B complex, providing a possible explanation for the lower levels of U1 snRNA. The observation that the levels of U1 snRNA are very low in A5' and A3' complexes (Fig. 4, lanes 3,4) suggests that both the 5' and the 3' splice sites stabilize U1 snRNP interactions in A complex assembled on intact pre-mRNA.

The 5' and 3' splice sites interact in E complex

The observation that E complex does not assemble in its entirety on either the 5' or the 3' splice site alone suggests that both splice sites play a role in E complex assembly. To investigate the functional roles of the 5' and 3' splice sites in E complex assembly, we developed a sensitive substrate-competition assay (Fig. 7). Although a previous study concluded that the 5' and 3' splice sites were required for efficient E (or commitment) complex

Figure 7. The 5' and the 3' ends of the intron interact in E complex. (A) Schematic of *HindIII* and *BamHI* sites used to generate 254-nucleotide (*BamHI*) or 284-nucleotide (*HindIII*) wild type or mutant AdML pre-mRNAs. (B, C, lanes 1–4) Pre-mRNAs used in the competition experiments: (wt) wild-type pre-mRNA; (5'ssΔ) pre-mRNA containing a 5' splice-site mutation; (pyrΔ) pre-mRNA containing a mutation in the pyrimidine tract at the 3' splice site (see Materials and methods). These pre-mRNAs were transcribed from DNA templates digested with either *BamHI* or *HindIII* as indicated. (B, lanes 5–16) wt *Bam* pre-mRNA (2 ng/ml, lanes 5–10; 0.2 ng/ml, lanes 11–16) was mixed with an equal amount of either wt *Hind* pre-mRNA (lanes 5,6,11,12), 5'ssΔ *Hind* pre-mRNA (lanes 7,8,13,14), or pyrΔ *Hind* pre-mRNA (lanes 9,10,15,16). E complex was assembled on the mixed RNAs and fractionated by gel filtration. RNA isolated from the E and H complex peaks was then analyzed by denaturing gel electrophoresis. (C) Same as B, except that wt *Hind* was mixed with wt *Bam* (lanes 5,6,11,12), 5'ssΔ *Bam* (lanes 7,8,13,14) or pyrΔ *Bam* (lanes 9,10,15,16) pre-mRNA.

assembly, this study assayed only for A complex formation (Jamison et al. 1992), without distinguishing between effects on E versus A complex assembly. In our assay, we examined E complex assembly directly (Fig. 7). This was achieved by mixing wild-type pre-mRNA with an equal amount of a pre-mRNA containing a specific mutation. The mixed RNAs were assembled into E complex, fractionated by gel filtration, and the distribution of mutant and wild-type RNAs in H or E complexes was determined by denaturing gel electrophoresis. The mixed pre-mRNAs were distinguished by size; 30 additional nucleotides were present at the 3' end of transcripts when the DNA transcription templates were digested with *HindIII* instead of *BamHI* (Fig. 7A–C, lanes 1–4). As a control for the effects of size differences, wild-type AdML pre-mRNAs truncated at the *BamHI* or *HindIII* sites (wt-Bam and wt-Hind, respectively) were mixed together and the distribution of these RNAs in E and H complexes was determined (Fig. 7B,C, lanes 5,6). These data show that there is slightly more wt-Bam than wt-Hind pre-mRNA in E complex, thus indicating a small deleterious effect of the additional exon sequences. Significantly, however, pre-mRNAs containing substitutions in the 5' splice site (5'ssΔ) or in the pyrimidine tract at the 3' splice site (pyrΔ) were present in much lower levels than the wild-type RNA in E complex (Fig. 7B,C, lanes 7,9). In contrast, the mutant RNAs were overrepresented, or at least present at similar levels as wild type, in H complex (Fig. 7B,C, lanes 8,10). We conclude that the 5' and 3' splice sites interact functionally in E complex.

To determine whether there was competition between

the wild-type and mutant RNAs for E complex formation, or whether the mutant RNAs were incapable of forming E complex, we repeated the experiments (Fig. 7B,C, lanes 5–10) using 10-fold less of each RNA in the reactions (Fig. 7B,C, lanes 11–16). Under these conditions, the levels of the mutant and wild-type pre-mRNAs were more similar in E complex (Fig. 7B,C, lanes 13,15). This observation is consistent with our data showing that E5' and E3' complexes can assemble on RNAs lacking the 3' or the 5' portion of the intron, respectively (see Figs. 2 and 3). Thus, although the 5' and 3' splice sites are recognized independently in E5' and E3' complexes, in the intact pre-mRNA, the 5' and 3' ends of the intron cooperate to form E complex.

Discussion

Understanding the early steps in mammalian spliceosome assembly is critical for elucidating the mechanisms of splice site selection in both constitutively and alternatively spliced pre-mRNAs. E complex, which is the earliest detectable functional intermediate in mammalian spliceosome assembly, accumulates only in reactions lacking ATP. Under these conditions, subsequent complexes are unable to assemble (Michaud and Reed 1991). Thus, E complex appears to be converted very rapidly to the first ATP-dependent prespliceosome complex (A) under normal splicing conditions (i.e., in the presence of ATP). In this study we show that specific ATP-independent complexes, designated E5' and E3', assemble on RNAs lacking a 3' or a 5' splice site, respectively. However, using a sensitive substrate-competi-

tion assay, we demonstrate that a pre-mRNA containing both a 5' and a 3' splice site is assembled into E complex more efficiently than RNAs lacking one or the other splice site. On the basis of this observation, we conclude that the 5' and 3' splice sites are recognized and are functionally associated with one another in E complex. Moreover, this observation suggests that in pre-mRNAs containing multiple splice sites, pairing of the appropriate 5' and 3' splice sites occurs at this very early step of spliceosome assembly. Thus, we predict that E complex assembly may be the major control point for splice site selection in both constitutively and alternatively spliced pre-mRNAs.

5' and 3' splice site interactions

Our studies, together with previous work, suggest a possible mechanism for the recognition and functional association of the 5' and 3' splice sites in E complex. We find that two components of E complex, U1 snRNP and U2AF, are enriched specifically in E5' and E3' complexes, respectively. These observations are consistent with previous studies showing that U1 snRNP binds to the 5' splice site and that U2AF binds to the 3' splice site (Mount et al. 1983; Zamore and Green 1989). Thus, a possible mechanism for bringing together the 5' and 3' ends of the intron in E complex is an association between U1 snRNP and U2AF. This association could be either direct or possibly mediated by one or more of the SAPs present in E complex (Bennett et al. 1992b). In support of this proposal, we do not detect any other proteins in E complex that are specific to the 5' or 3' splice sites. Similarly, in UV cross-linking studies, U2AF is the only protein splicing factor that cross-links to the pre-mRNA in E and E3' complexes (D. Staknis and R. Reed, unpubl.). We cannot rule out the possibility that factors other than U1 snRNP and U2AF mediate the interactions between the 5' and 3' splice sites but are not detectable by UV cross-linking or analysis of our silver-stained two-dimensional gels.

In addition to U2AF and U1 snRNP (Kramer and Utans 1991), at least one member of the SR family of splicing factors is required for A complex assembly (Fu and Maniatis 1990; Krainer et al. 1990a; Fu et al. 1992). The SR family includes SF2/ASF (Ge et al. 1991; Krainer et al. 1991) and SC35 (Fu and Maniatis 1992), and other family members can substitute functionally for either of these proteins in *in vitro* complementation studies (Fu et al. 1992; Mayeda et al. 1992; Zahler et al. 1992). As the SR proteins appear to be involved in 5' (Ge and Manley 1990; Krainer et al. 1990b) and 3' splice-site selection (Fu et al. 1992), it is possible that these proteins function in E complex. However, using antibodies, we are unable to detect the SR proteins on blots of two-dimensional gels of E, A, or B complexes. Thus, further studies are required to determine precisely when during the early steps of spliceosome assembly the SR proteins play a role.

Studies in *Schizosaccharomyces pombe* suggest that U1 snRNA may interact directly with the 3' splice site

by base-pairing interactions with the AG dinucleotide (Reich et al. 1992). In mammals, the AG dinucleotide is required for lariat formation in introns containing short pyrimidine tracts (AG-dependent introns) but not in introns containing long pyrimidine tracts (AG-independent introns; Reed 1989). Using our substrate-competition assay, we observed a reproducible one- to twofold decrease in the efficiency of E complex assembly in the absence of the AG dinucleotide in an AG-dependent intron (S. Michaud and R. Reed, unpubl.). This small effect could be caused by a slight decrease in U2AF binding that may have a small requirement for the AG dinucleotide (Ruskin et al. 1988; Hoffman and Grabowski 1992) or to U1 snRNA-AG interactions, or both (Reich et al. 1992).

In this and other studies, low levels of U2 snRNP have been detected in purified E complex (Michaud and Reed 1991) or specifically bound to pre-mRNA in reactions lacking exogenously added ATP (Jamison and Garcia-Blanco 1992; Wassarman and Steitz 1992). However, extensive efforts have failed to show that U2 snRNP is functionally associated with E complex (S. Michaud and R. Reed, unpubl.). Moreover, using our substrate-competition assay, we obtained no evidence that the branch-point sequence is involved in E complex assembly (O. Gozani and R. Reed, unpubl.). On the basis of these observations, we believe that a functional association between the pre-mRNA and U2 snRNP does not occur until A complex.

Both the initial recognition of the 5' and 3' splice sites by U1 snRNP and U2AF and the proposed association of the splice sites by these factors apparently occur in a very rapid step under normal splicing conditions. Subsequent to E complex, U1 and U2 snRNPs are stably bound to the pre-mRNA in A complex. It is likely that the 5' and 3' splice sites remain associated in A complex, maintaining the interactions that were established previously in E complex. However, in A complex, U1 snRNP bound to the 5' splice site may interact with U2 snRNP bound to the 3' end of the intron, rather than with U2AF. This possibility is suggested by the observation that in A3' complex U2AF is actually displaced from the pre-mRNA when U2 snRNP binds (Bennett et al. 1992b; D. Staknis and R. Reed, unpubl.). However, further studies are required to determine whether U2AF is also displaced in A complex assembled on intact pre-mRNA. Whether the interactions between U1 and U2 snRNPs are direct or indirect remains to be established.

Our data suggest that the interactions between U1 snRNP and the 5' splice site are destabilized around the time of B complex assembly. This conclusion is based on the observation that U1 snRNA and U1 snRNP components are present in significantly lower levels in B than in E complex (this study; Michaud and Reed 1991; Bennett et al. 1992b). In addition, U1 snRNP is present in lower levels in B complex than are U2, U5, and U6 snRNPs. Our suggestion that U1 snRNP is destabilized in B complex is consistent with psoralen cross-linking studies (Wassarman and Steitz 1992), which showed that U1 snRNP cross-links to the 5' splice site during the

same time period that we detect A complex, but then disappears. U5 snRNA then cross-links to the 5' splice site at a time when we detect B complex. Thus, it appears that the U1 snRNA–5' splice site interaction that occurs in E and A complexes is destabilized and replaced in B complex by the U5 snRNA–5' splice site interaction (this study; Newman and Norman 1992; Steitz 1992; Wassarman and Steitz 1992).

What is the significance of complexes assembled on RNAs lacking 5' or 3' splice sites?

In the earliest studies of mammalian spliceosome assembly, it was found that A complex assembled on RNAs lacking a 5' splice site (Konarska and Sharp 1986). This complex, which we have designated A3', is thought to be the same as A complex assembled on intact pre-mRNA because it fractionates in the same position as A complex on native gels and contains U2 snRNP (Konarska and Sharp 1986; Moore et al. 1993). These observations led to the notion that the 5' splice site was dispensable for A complex assembly. However, this notion was difficult to reconcile with the subsequent discovery of E complex, which contains U1 snRNP and, therefore, implicates a role for the 5' splice site before A complex assembly (Reed 1990; Michaud and Reed 1991). Bearing on this paradox, previous studies have shown that U1 snRNP was required for efficient A3' assembly, but the region of U1 snRNA that base-pairs with the 5' splice site was dispensable (Zillman et al. 1988; Barabino et al. 1990). Putting all of these observations together, it was suggested that base-pairing between the 5' splice site and U1 snRNA does not occur in E complex, or if it does occur it is not the U1 snRNP interaction required for subsequent assembly of A complex (for review, see Rosbash and Seraphin 1991).

On the basis of our study, we propose an alternative explanation of the relationships among U1 snRNP, the 5' splice site, and E and A complexes. We identified a specific ATP-independent complex, designated E3', that assembles on RNAs containing a functional 3' splice site but lacking a 5' splice site. Significantly, we demonstrated that E3' complex is a functional precursor to A3' complex. Moreover, consistent with the notion that U1 snRNP is required for efficient A3' assembly (Zillman et al. 1988; Barabino et al. 1990), analysis of the composition of E3' revealed low levels of U1 snRNA and U1 snRNP proteins. Thus, our data indicate that E3'-containing U1 snRNP is a precursor to A3'. However, our substrate-competition experiments showed that RNAs containing both the 5' and the 3' splice site assemble E complex more efficiently than RNAs containing only one splice site. On the basis of these observations, we conclude that although E3' and A3' complexes can assemble, the presence of a 5' splice site bound to U1 snRNP results in the more efficient and stable assembly observed with E and A complexes. Consistent with this proposal, Lamond and co-workers (1987) showed that the presence of a 5' splice site increases the kinetics and efficiency of A complex assembly.

Understanding the relationships among E3', A3', and E and A complexes provides insight into the relationship between E complex in mammals and the commitment complex in yeast. The mammalian and yeast complexes are similar to one another in that the 5' and 3' splice sites communicate before U2 snRNP binding (this study; Legrain et al. 1988; Seraphin and Rosbash 1989, 1991). The notable difference is that the 5' splice site appears to be critical for interactions at the 3' splice site in yeast, whereas in mammals the 5' splice site is important only for the efficiency and stability of these interactions at the 3' splice site. The other distinction between the yeast and mammalian commitment complexes is that the pyrimidine tract in mammals substitutes functionally for the branch site in yeast (this study; Seraphin and Rosbash 1989). As U2AF binds to the pyrimidine tract in E complex in mammals, it will be interesting to identify the yeast counterpart that presumably interacts with the branch site in the yeast commitment complex.

The role of E complex in splice site selection

Splice site selection requires mechanisms for distinguishing between normal and cryptic splice sites and for bringing together the appropriate pairs of 5' and 3' splice sites in pre-mRNAs containing multiple splice sites. Our observation that the 5' and 3' splice sites are first brought together in E complex suggests that the factors in this complex, such as U2AF and U1 snRNP, have a fundamental role in establishing splice site selection. Moreover, these factors are likely targets for regulation in alternatively spliced pre-mRNAs. Consistent with the latter proposal, alternative splicing of the *Drosophila* transformer pre-mRNA was recently found to be negatively regulated by specific competition between sex-lethal protein and U2AF for binding to the 3' splice site (Valcarcel et al. 1993). In addition, regulated alternative splicing of *Drosophila* doublesex pre-mRNA may involve stimulation of U2AF binding to a weak pyrimidine tract by a downstream positive regulator (Tian and Maniatis 1992). Other examples in which the patterns of alternative splicing appear to be determined at the level of E complex include selection of a weak 3' splice site in preprotachykinin pre-mRNA by targeting U2AF to this site by interactions with U1 snRNP bound to a downstream 5' splice site (Hoffman and Grabowski 1992), and somatic inhibition of a *Drosophila* P-element 5' splice site by control of U1 snRNP binding (Siebel et al. 1992). Finally, our data show that the efficiency of E complex assembly can be affected by the context of both the 5' and 3' splice sites. This observation is consistent with the possibility that at least the initial mechanism for distinguishing between cryptic and normal splice sites operates in E complex. Subsequent fine-tuning, such as selection of the precise bond to be cleaved at the 5' or 3' splice sites, may occur at later stages and involve other factors such as U5 snRNP (Newman and Norman 1991, 1992).

Materials and methods

Plasmids

Plasmid pAdML was constructed by subcloning an *EcoRI*–*Sau3A* fragment from pBSAd10 (gift from M. Konarska, The Rockefeller Institute, NY) into the *EcoRI* and *Bam*HI sites of SP72 (Promega Biotech). This fragment contains exon 1 (71 nucleotides), intron 1 (97 nucleotides), and exon 2 (45 nucleotides) derived from the adenovirus 2 major late (AdML) transcription unit. DNA was linearized with *Bam*HI (+213 in pAdML) for transcription of the whole pre-mRNA and with *Fsp*I (+107 in pAdML) for transcription of AdML5' RNA. pAdML3', which contains the 3' portion of the AdML intron and exon 2, was constructed by ligating the *Fsp*I–*Bam*HI fragment from pAdML into the *Pvu*II–*Bam*HI sites of SP73 (Promega Biotech). pAdMLΔ is identical to pAdML3' except that the eight consecutive T residues in the 3' splice site (+155–162 in AdML) are changed to GTGATCAC. pAd3' and pAdΔ were linearized for transcription with *Bam*HI. The wild-type AdML pre-mRNA used in the competition experiment (Fig. 7) was transcribed from pAdML + *Xho*I, which is the same as pAdML except that an *Xho*I linker was inserted into the *Fsp*I site. pAdMLΔ5'ss is the same as pAdML + *Xho*I except that the region around the 5' splice site (GTTGGGGTGAG) has been changed to ATTG–GAGCCAC. pAdMLΔpyr was made by replacing the *Xho*I–*Bam*HI fragment of pAdML + *Xho*I with the *Xho*I–*Bam*HI fragment of pAdMLΔ. pAdML + *Xho*I, pAdMLΔ5'ss, and pAdMLΔpyr were linearized for transcription with either *Bam*HI or *Hind*III, as indicated in the legend to Figure 7. Plasmid pβ5' was constructed from plasmid HβΔ*Bam*-X (Reed et al. 1988), which contains an *Xho*I linker in the *Mbo*II site (+232). pβ5' was obtained by subcloning a *Hind*III–*Xho*I fragment from HβΔ*Bam*-X (+1–232) into SP64 (Promega Biotech); subsequently, a *Hind*III–*Bam*HI fragment containing this *Hind*III–*Xho*I fragment was subcloned into SP73 (Promega Biotech). DNA was linearized with *Eco*RI for transcription. pβ3' was constructed by inserting an *Xho*I–*Bam*HI fragment from HβΔ*Bam*-X (+232–500) into SP64 (Promega Biotech), and then subcloning a *Hind*III–*Bam*HI fragment containing the *Xho*I–*Bam*HI fragment into SP73 (Promega Biotech). DNA was linearized with *Bam*HI for transcription. Plasmid pβΔ was a gift from Michael Green (University of Massachusetts Medical School, Worcester) and is the same as pβ5' except that nucleotides 113–154 are deleted, which include the 5' splice site and one cryptic 5' splice site. One cryptic 5' splice site remains in this RNA. pAd/β was constructed by ligating the *Xho*I–*Alu*I fragment from pAdML3' (+107–168), the *Sma*I–*Bam*HI fragment (containing exon 2 sequences from +11–205) from a derivative of T7Hβ (Reed et al. 1988), and *Xho*I–*Bam*HI vector sequences from SP73 (Promega Biotech). DNA was linearized with *Bam*HI for transcription. The nonspecific RNA lacking functional splice sites (NS2) is the antisense transcript of B3P3 (Mullen et al. 1991; gift from B. Nadal-Ginard, Harvard Medical School, Boston, MA). The nonspecific RNA containing a 5' splice site (NS1) was synthesized by antisense transcription of a human β-globin derivative (HβΔAG; Reed 1990).

Pre-mRNA synthesis and in vitro splicing reactions

All plasmids were transcribed with T7 polymerase except NS1, which was transcribed with SP6 polymerase (Melton et al. 1984). Biotinylated pre-mRNAs (Grabowski and Sharp 1986) were synthesized in standard transcription reactions (Melton et al. 1984) that included 15 μM biotinylated UTP (Enzo Biochemicals) and 100 μM UTP. RNAs were capped during transcription

as described (Konarska et al. 1984). In vitro splicing reactions were carried out according to Krainer and colleagues (1984), except that polyvinyl alcohol was omitted. Reaction volumes ranged from 100 to 125 μl for analytical purposes to 2.4 ml for preparative purification and contained 0.2–4 ng/μl of ³²P-labeled RNA. Reactions contained 30% nuclear extract and were incubated at 30°C for 30 min to form B, E, E5', E3', A5', and A3' complexes, and at 30°C for 5 min to form A complex, unless otherwise indicated in the figure legends. For assembly of E, E5' and E3' complexes, the nuclear extract was depleted of ATP as described (Michaud and Reed 1991) and the reactions lacked added ATP, MgCl₂, and creatine phosphate.

Purification and analysis of splicing complexes

For gel filtration purification of splicing complexes, in vitro splicing reactions were loaded directly onto 1.5 × 50-cm Sephacryl S-500 columns equilibrated in FSP buffer [20 mM Tris (pH 7.8), 0.1% Triton X-100, 60 mM KCl, 2.5 mM EDTA] (Abmayr et al. 1988; Reed et al. 1988). For affinity purification of splicing complexes, the peak fractions were pooled and the salt concentration was adjusted to a final concentration of 250 mM. Avidin–agarose (15 μl/ml) (Vector Labs) was added to the pooled fractions and mixed overnight at 4°C. Avidin beads were then washed four times using a volume of buffer equal to that of the pooled fractions (~3–5 ml for each wash). The washing buffer for the affinity selections was 20 mM Tris (pH 7.9), 250 mM NaCl. The snRNA compositions of the affinity-purified complexes were determined by two methods. Total RNA was prepared from equivalent amounts of each affinity-purified complex and was then either end-labeled with [³²P]pCp and RNA ligase, as described (Reed 1990), or fractionated directly on an 8% denaturing polyacrylamide gel and visualized by ethidium bromide staining. The ³²P-labeled RNAs were visualized by PhosphorImage analysis (Molecular Dynamics). Protein compositions of E, E5', and E3' complexes were determined by two-dimensional gel electrophoresis as described (Bennett et al. 1992b). Proteins were visualized by silver staining (Morrissey 1981). Native gel electrophoresis of splicing complexes was carried out as described (Konarska and Sharp 1987), except that heparin was added to the reactions to a final concentration of 0.5 mg/ml before fractionation on the native gel.

In vitro complementation assays

Extracts used for the in vitro complementation assay in Figure 1C were prepared by incubating nuclear extract for 10 min at 30°C with 9 μg of cold competitor RNA (β5') per milliliter of extract. We showed previously that with this amount of competitor E complex is chased into A and B complexes and into spliced products, whereas H complex is not (Michaud and Reed 1991). For the complementation reaction, 20-μl aliquots of each gel filtration fraction was incubated for 10 min at 30°C in a 50-μl splicing reaction containing 10 μl of the competitor-titrated extract. Heparin was added to the reactions to 0.5 mg/ml, and they were fractionated by native gel electrophoresis as described above. Complexes were visualized by PhosphorImage analysis (Molecular Dynamics). S-100 extracts used in the complementation assay in Figure 3C were a gift from Adrian Krainer (Cold Spring Harbor Laboratory, NY) and can be complemented by the purified splicing factor SF2 (Krainer et al. 1990a). In vitro complementation reactions containing S-100 employed the same dilution assay as was described previously (Michaud and Reed 1991), except that E complex and mock reactions were diluted 87.5-fold before incubation in the complementation re-

action. Reactions were fractionated by native gel electrophoresis as described above.

Acknowledgments

We are grateful to Keiko Kumatori and Rebecca Feld for excellent technical assistance. We thank the members of our laboratory for valuable discussions and comments on the manuscript. HeLa cells for nuclear extracts were provided by the National Institutes of Health (NIH) cell culture facility at Endotronics. S.M. is supported by fellowships from the Markey Program in Developmental Biology and the Albert J. Ryan Foundation. R.R. is a Lucille P. Markey Scholar. This work was supported by a grant from the Lucille P. Markey Charitable Trust and a grant from NIH.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Abmayr, S.M., R. Reed, and T. Maniatis. 1988. Identification of a functional mammalian spliceosome containing unspliced pre-mRNA. *Proc. Natl. Acad. Sci.* **85**: 7216–7220.
- Barabino, S.M.L., B.J. Blencowe, U. Ryder, B.S. Sproat, and A.I. Lamond. 1990. Targeted snRNP depletion reveals an additional role for mammalian U1 snRNP in spliceosome assembly. *Cell* **63**: 293–302.
- Bennett, M., S. Pinol-Roma, D. Staknis, G. Dreyfuss, and R. Reed. 1992a. Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly in vitro. *Mol. Cell. Biol.* **12**: 3165–3175.
- Bennett, M., S. Michaud, J. Kingston, and R. Reed. 1992b. Protein components specifically associated with prespliceosome and spliceosome complexes. *Genes & Dev.* **6**: 1986–2000.
- Bindereif, A. and M.R. Green. 1986. Ribonucleoprotein complex formation during pre-mRNA splicing in vitro. *Mol. Cell. Biol.* **6**: 2582–2592.
- Frederewey, D. and W. Keller. 1985. Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. *Cell* **42**: 355–367.
- Fu, X.-D. and T. Maniatis. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* **343**: 437–441.
- . 1992. The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. *Proc. Natl. Acad. Sci.* **89**: 1725–1729.
- Fu, X.-D., A. Mayeda, T. Maniatis, and A.R. Krainer. 1992. General splicing factors SF2 and SC35 have equivalent activities in vitro, both affect alternative 5' and 3' splice site selection. *Proc. Natl. Acad. Sci.* **89**: 11224–11228.
- Ge, H. and J.L. Manley. 1990. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell* **62**: 25–34.
- Ge, H., P. Zuo, and J.L. Manley. 1991. Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. *Cell* **66**: 373–382.
- Grabowski, P.J. and P.A. Sharp. 1986. Affinity chromatography of splicing complexes: U2, U5, and U4+U6 small nuclear ribonucleoprotein particles in the spliceosome. *Science* **233**: 1294–1299.
- Grabowski, P.J., S.R. Seiler, and P.A. Sharp. 1985. A multicomponent complex is involved in the splicing of messenger RNA precursors. *Cell* **42**: 345–353.
- Green, M.R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu. Rev. Cell Biol.* **7**: 559–599.
- Guthrie, C. 1991. Messenger RNA splicing in yeast: Clues to why the spliceosome is a ribonucleoprotein. *Science* **253**: 157–163.
- Hall, K.B. and M.M. Konarska. 1992. The 5' splice site consensus RNA oligonucleotide induces assembly of U2/U4/U5/U6 small nuclear ribonucleoprotein complexes. *Proc. Natl. Acad. Sci.* **89**: 10969–10973.
- Hoffman, B. and P. Grabowski. 1992. U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes & Dev.* **6**: 2554–2568.
- Jamison, S.F. and M.A. Garcia-Blanco. 1992. An ATP-independent U2 small nuclear ribonucleoprotein particle/precursor mRNA complex requires both splice sites and the polypyrimidine tract. *Proc. Natl. Acad. Sci.* **89**: 5482–5486.
- Jamison, S.F., A. Crow, and M.A. Garcia-Blanco. 1992. The spliceosome assembly pathway in mammalian extracts. *Mol. Cell. Biol.* **12**: 4279–4287.
- Konarska, M.M. and P.A. Sharp. 1986. Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* **46**: 845–855.
- . 1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell* **49**: 763–774.
- Konarska, M.M., R.A. Padgett, and P.A. Sharp. 1984. Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* **38**: 731–736.
- Krainer, A.R. and T. Maniatis. 1985. Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for pre-mRNA splicing in vitro. *Cell* **42**: 725–736.
- Krainer, A.R., T. Maniatis, B. Ruskin, and M.R. Green. 1984. Normal and mutant human β -globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell* **36**: 993–1005.
- Krainer, A.R., G.C. Conway, and D. Kozak. 1990a. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* **62**: 35–42.
- . 1990b. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes & Dev.* **4**: 1158–1171.
- Kramer, A. and U. Utans. 1991. Three protein factors (SF1, SF3, and U2AF) function in pre-splicing complex formation in addition to snRNPs. *EMBO J.* **10**: 1503–1509.
- Lamm, G.M., B.J. Blencowe, B.S. Sproat, A.M. Iribarren, U. Ryder, and A.I. Lamond. 1991. Antisense probes containing 2-aminoadenosine allow efficient depletion of U5 snRNP from HeLa splicing extracts. *Nucleic Acids Res.* **19**: 3193–3198.
- Lamond, A.I., M.M. Konarska, and P.A. Sharp. 1987. A mutational analysis of spliceosome assembly: Evidence for splice site collaboration during spliceosome formation. *Genes & Dev.* **1**: 532–543.
- Legrain, P., B. Seraphin, and M. Rosbash. 1988. Early commitment of yeast pre-mRNA to the spliceosome assembly pathway. *Mol. Cell. Biol.* **8**: 3755–3760.
- Mayeda, A., A.M. Zahler, A.R. Krainer, and M.B. Roth. 1992. Two members of a conserved family of nuclear phosphoproteins are involved in pre-mRNA splicing. *Proc. Natl. Acad. Sci.* **89**: 1301–1304.
- Melton, D.A., P.A. Krieg, M.R. Rebagliati, T. Maniatis, K. Zinn, and M.R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plas-

- mids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**: 7035–7056.
- Michaud, S. and R. Reed. 1991. An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes & Dev.* **5**: 2534–2546.
- Moore, M.J., C.C. Query, and P.A. Sharp. 1993. Splicing of precursors to messenger RNAs by the spliceosome. In *The RNA world*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. (In press.)
- Morrissey, J.H. 1981. Silver stain for proteins in polyacrylamide gels: A modified procedure with enhanced uniform sensitivity. *Anal. Biochem.* **117**: 307–310.
- Mount, S.M., I. Pettersson, M. Hinterberger, A. Karmas, and J.A. Steitz. 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. *Cell* **33**: 509–518.
- Mullen, M.P., C.W.J. Smith, J.G. Patton, and B. Nadal-Ginard. 1991. α -Tropomyosin mutually exclusive exon selection: Competition between branchpoint/polypyrimidine tracts determines default exon choice. *Genes & Dev.* **5**: 642–655.
- Newman, A. and C. Norman. 1991. Mutations in yeast U5 snRNA alter the specificity of 5' splice-site cleavage. *Cell* **65**: 115–123.
- . 1992. U5 snRNA interacts with exon sequences at the 5' and 3' splice sites. *Cell* **68**: 743–754.
- Perkins, K.K., H.M. Furneaux, and J. Hurwitz. 1986. RNA splicing products formed with isolated fractions from HeLa cells are associated with fast-sedimenting complexes. *Proc. Natl. Acad. Sci.* **83**: 887–891.
- Reed, R. 1989. The organization of 3' splice-site sequences in mammalian introns. *Genes & Dev.* **3**: 2113–2123.
- . 1990. Protein composition of mammalian spliceosomes assembled in vitro. *Proc. Natl. Acad. Sci.* **87**: 8031–8035.
- Reed, R. and T. Maniatis. 1985. Intron sequences involved in lariat formation during pre-mRNA splicing. *Cell* **41**: 95–105.
- Reed, R., J. Griffith, and T. Maniatis. 1988. Purification and visualization of native spliceosomes. *Cell* **53**: 949–961.
- Reich, C.I., R.W. VanHoy, G.L. Porter, and J.A. Wise. 1992. Mutations at the 3' splice site can be suppressed by compensatory base changes in U1 snRNA in fission yeast. *Cell* **69**: 1159–1169.
- Rosbash, M. and B. Seraphin. 1991. Who's on first? The U1 snRNP-5' splice site interaction and splicing. *Trends Biochem. Sci.* **16**: 187–190.
- Ruby, S.R. and J. Abelson. 1988. An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* **242**: 1028–1035.
- . 1991. Pre-mRNA splicing in yeast. *Trends Genet.* **7**: 79–85.
- Ruskin, B. and M. Green. 1985. Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing. *Nature* **317**: 732–734.
- Ruskin, B., P.D. Zamore, and M.R. Green. 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* **52**: 207–219.
- Seraphin, B. and M. Rosbash. 1989. Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* **59**: 349–358.
- . 1991. The yeast branchpoint sequence is not required for the formation of a stable U1 snRNA-pre-mRNA complex and is recognized in the absence of U2 snRNA. *EMBO J.* **10**: 1209–1216.
- Siebel, C.W., L.D. Fresco, and D.C. Rio. 1992. The mechanism of somatic inhibition of *Drosophila* P-element pre-mRNA splicing: Multiprotein complexes at an exon pseudo-5' splice site control U1 binding. *Genes & Dev.* **6**: 1386–1401.
- Steitz, J.A. 1992. Splicing takes a Holliday. *Science* **257**: 888–889.
- Tian, M. and T. Maniatis. 1992. Positive control of pre-mRNA splicing in vitro. *Science* **256**: 237–240.
- Valcarcel, J., R. Singh, P.D. Zamore, and M.R. Green. 1993. Sex-lethal antagonizes the splicing factor U2AF to regulate tra alternative splicing in vitro. *Nature* (in press).
- Wassarman, D.A. and J.A. Steitz. 1992. Interactions of small nuclear RNAs with precursor messenger RNA during in vitro splicing. *Science* **257**: 1918–1925.
- Zahler, A.M., W.S. Lane, J.A. Stolk, and M.B. Roth. 1992. SR proteins: A conserved family of pre-mRNA splicing factors. *Genes & Dev.* **6**: 837–847.
- Zamore, P.D. and M.R. Green. 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc. Natl. Acad. Sci.* **86**: 9243–9247.
- Zillmann, M., M.L. Zapp, and S.M. Berget. 1988. Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. *Mol. Cell. Biol.* **8**: 814–821.

A functional association between the 5' and 3' splice site is established in the earliest prespliceosome complex (E) in mammals.

S Michaud and R Reed

Genes Dev. 1993, **7**:

Access the most recent version at doi:[10.1101/gad.7.6.1008](https://doi.org/10.1101/gad.7.6.1008)

References

This article cites 58 articles, 29 of which can be accessed free at:
<http://genesdev.cshlp.org/content/7/6/1008.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

Focused on your science.

USC
SCIENTIFIC
SIDE BY SCIENCE