Mammalian U2 snRNP has a sequence-specific RNA-binding activity

Kristin K. Nelson and Michael R. Green

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138 USA

The RNA branch formed during pre-mRNA splicing occurs at a wide variety of sequences (branch sites) in different mammalian pre-mRNAs. U2 small nuclear ribonucleoprotein (snRNP) binds to the pre-mRNA branch site following the interaction of a protein, U2AF, with the 3' splice site/polypyrimidine tract. Here we show that despite the variability of mammalian branch sites, U2 snRNP has a sequence-specific RNA-binding activity. Thus, RNA branch formation is regulated by two sequence-specific interactions: U2AF with the 3' splice site/polypyrimidine tract, and U2 snRNP with the branch site. The affinity of the branch site for U2 snRNP affects the efficiency of spliceosome assembly and splicing.

[Key Words: U2 snRNP; U2AF; branch site; 3' splice site]

Received June 9, 1989; revised version accepted July 19, 1989.

Assembly of the mammalian spliceosome involves two pre-mRNA/small nuclear ribonucleoprotein (snRNP) interactions; U1 snRNP binds to the 5' splice site, and U2 snRNP binds to a region encompassing the site of RNA branch formation (the branch site) (for review, see Green 1986; Padgett et al. 1986; Maniatis and Reed 1987; Sharp 1987). The specificity of U1 snRNP binding apparently is dictated solely by RNA-RNA base pairing between U1 snRNP and the 5' splice site (Zhuang and Weiner 1986). The determinants of U2 snRNP-binding specificity are more complex. This ATP-dependent binding reaction requires at least one protein, U2AF (Ruskin et al. 1988), and, perhaps, other factors (Kramer 1988), in addition to U2 snRNP. Mammalian branch sites are highly variable, and efficient binding of U2 snRNP requires an additional sequence element, the 3' splice site/polypyrimidine tract (Ruskin and Green 1985a; Chabot and Steitz 1987; Ruskin et al. 1988).

The RNA branch normally forms at an adenosine within a weak consensus located 18–38 nucleotides upstream of the 3' splice site (for review, see Green 1986). Mutational studies have attempted to establish the importance of the specific sequences of mammalian branch sites. In general, the authentic branch site can be deleted or mutated without abolishing accurate splicing, due to activation of new (cryptic) branch sites (Padgett et al. 1985; Ruskin et al. 1985; Hornig et al. 1986; Freyer et al. 1987; Zhuang et al. 1989). These cryptic branch sites, which usually include an adenosine as the branch nucleotide, are located 18–38 nucleotides upstream of the 3' splice site and often do not resemble the authentic branch site.

The mechanism of U2 snRNP binding in *Saccharomyces cerevisiae* differs from that in mammalian cells (for review, see Green 1986; Padgett et al. 1986). In *S.*

cerevisiae the RNA branch always forms at the third adenosine in the highly conserved sequence UA-CUAAC. When the UACUAAC element is deleted, splicing is abolished. Furthermore, the 3' splice site/polypyrimidine tract is not required for either U2 snRNP binding or for subsequent cleavage at the 5' splice site and formation of the lariat intermediate (Rymond and Rosbash 1985). The specificity of U2 snRNP binding in *S. cerevisiae* may be provided solely by RNA-RNA base pairing between the UACUAAC sequence and a complementary region of U2 snRNA (Parker et al. 1987). Thus, branch site selection is primarily sequence dependent in yeast and position dependent in mammalian cells.

We suggested previously that the distance constraint in mammalian branch site selection is due to the requirement for prior binding of U2AF to the 3' splice site/polypyrimidine tract (Ruskin et al. 1988). However, within 18–38 nucleotides upstream of the 3' splice site, the RNA branch usually forms at only one of several adenosines. Thus, there must be an additional component that contributes to the specificity of mammalian branch site selection.

The additional specificity could be imposed either at the level of U2 snRNP binding or at some subsequent step during the process of RNA branch formation. In this report we show that this additional specificity is provided by the sequence-specific binding of U2 snRNP to the branch site.

Results

U2 snRNP binds to the branch site in the absence of the 3' splice site/polypyrimidine tract

Previous studies have shown that the 3' splice site/polypyrimidine tract is required for efficient binding of U2

snRNP to the branch site (Ruskin and Green 1985a; Chabot and Steitz 1987; Ruskin et al. 1988). However, to avoid any specificity imposed by the 3' splice site/polypyrimidine tract (U2AF-binding site), we measured U2 snRNP binding in the absence of this sequence element. Figure 1 shows that U2 snRNP binds specifically to the

Figure 1. U2 snRNP accurately binds to the branch site of a substrate lacking the 3' splice site/polypyrimidine tract. (A) RNase A protection assay. $\Delta E1$ wild-type (WT) and $\Delta E1$ $\Delta PyAG$ β -globin ^{32}P -labeled RNAs were incubated in a HeLa cell nuclear extract under splicing conditions at 23°C and treated with RNase A, and the RNase A-resistant fragments were selected by immunoprecipitation. The antibodies used are anti-Sm, anti-U1/U2 antisera, and an anti-U1 monoclonal antibody directed against the 70-kD U1-specific protein. The RNase A-resistant fragments were fractionated on a 10% denaturing polyacrylamide gel and visualized by autoradiography. The structures of the substrates are diagramed below. Exons are indicated by boxes; introns are indicated by lines; deleted sequences are represented by dotted line. The adenosine at which the RNA branch forms is shown. (B) RNase T1 digestion analysis. The RNase A-resistant fragments were eluted from the gel and digested with RNase T1, and the RNase T1 fragments fractionated on a 20% denaturing polyacrylamide gel and visualized by autoradiography. (Left) Sizes of fragments. (C) RNase A secondary analysis. The 10-nucleotide RNase T1 fragments in B were eluted from the gel and digested to completion with RNase A. The RNase A digestion products were fractionated by two-dimensional thin-layer chromatography and visualized by autoradiography (Ruskin et al. 1984). The composition of the products is indicated.

branch site of an RNA substrate deleted of the 3' splice site/polypyrimidine tract. In these experiments we measure stable U2 snRNP binding, using an RNase A protection/immunoprecipitation assay (Black et al. 1985; Ruskin et al. 1985). U2 snRNP protects the branch site from RNase A digestion, resulting in a 'core' RNase A-resistant fragment, which varies between 28 and 36 nucleotides, depending on the particular RNA substrate (Ruskin et al. 1988; see below). The mutant $\Delta PyAG$, which lacks the 3' splice site/polypyrimidine tract, gives rise to a low level of an RNase A-resistant fragment whose size is identical to that generated from the wild-type substrate (Fig. 1A, lanes 1 and 4).

Immunoprecipitation experiments confirmed that the RNase A-resistant fragment derived from $\Delta PyAG$ resulted from U2 snRNP binding (Fig. 1A). The RNase A-resistant fragments were immunoprecipitated with one of three different antisera: anti-Sm, which recognizes U1, U2, U5, and U4/6 snRNPs; anti-U1/U2, which recognizes U1 and U2 snRNPs; and anti-70 kD, which recognizes U1 snRNP. The RNase A-resistant fragment generated from $\Delta PyAG$ was immunoprecipitated efficiently with the anti-Sm and the anti-U1, U2 sera but not the anti-U1 specific antibody. Thus, protection of both the wild-type and $\Delta PyAG$ branch sites results from U2 snRNP binding.

To determine whether the protected fragments from the $\Delta PyAG$ and wild-type substrates were identical, these fragments were purified and digested to completion with RNase T1. The two RNase T1 digestion patterns are identical (Fig. 1B). The largest RNase T1 fragment, a 10-mer, was isolated and digested to completion with RNase A, and the RNase A digestion products fractionated by two-dimensional thin-layer chromatography (Fig. 1C). The RNase A digestion pattern is diagnostic for the 10-nucleotide RNase T1 fragment that spans the human β -globin branch site (Ruskin et al. 1984). Thus, the RNase A-resistant fragments of wild-type human β -globin and $\Delta PyAG$ include the branch site and are identical to one another.

Accurate lariat formation in the absence of the 3' splice site/polypyrimidine tract

The experiments in Figure 1 demonstrate that in the absence of the 3' splice site/polypyrimidine tract, U2 snRNP bound specifically to the branch site. In light of this result, we tested the mutant substrate to determine whether it could support RNA branch formation. Following incubation of this substrate in nuclear extract, RNA branch formation was assayed by primer-extension analysis (Ruskin et al. 1984). Figure 2 reveals an 85-nucleotide primer-extension product that maps precisely to the adenosine of the authentic branch site. This primer-extension product was eliminated by prior enzymatic debranching (Ruskin and Green 1985b) of the RNA sample, confirming that it resulted from a 2' to 5' phosphodiester bond. Thus, an RNA branch can form accurately on a substrate following deletion of the 3' splice site/polypyrimidine tract. (We view the possibility that a

Figure 2. Primer-extension analysis of Δ PyAG processing products. Transcripts of wild-type (WT) and Δ PyAG substrates were incubated in nuclear extract under splicing conditions for the times indicated above the autoradiogram; a 2-hr time point was further treated with debranching enzyme (2D). Analysis of the processing products by primer extension was as described in Ruskin et al. (1984). The 5' ^{32} P-end-labeled primer is complementary to positions +318 to +337 within exon 2. The primer-extension products were fractionated on a 5% denaturing polyacrylamide gel and visualized by autoradiography. (Right) The sizes of the primer-extension products and the RNA substrates from which they were derived. The RNA substrates and identities of the primer-extension products are shown below the autoradiogram. The position of the 5' ^{32} P-end label is indicated by a star.

cryptic 3' splice site/polypyrimidine tract is responsible for RNA branch formation as unlikely because (1) there is no sequence resembling a 3' splice site/polypyrimidine tract 18–38 nucleotides downstream from the Δ PyAG branch site, and (2) the final products of the splicing reaction are not detected.)

Binding of U2 snRNP to Δ PyAG requires additional factors

Previous studies have shown that stable binding of U2 snRNP to the branch site of a wild-type pre-mRNA requires additional factors (Kramer 1988; Ruskin et al. 1988). Because at least one of these factors, U2AF, in-

teracts with the 3' splice site (Ruskin et al. 1988), we asked whether U2AF (or other) factor(s) was required for U2 snRNP binding to Δ PyAG, which lacks the normal U2AF-binding site. U2 snRNP was separated from U2AF and many other proteins by centrifugation at high ionic strength. Under these conditions, U2 snRNP pellets, whereas U2AF remains in the supernatant (Ruskin et al. 1988). When the wild-type or Δ PyAG pre-mRNAs are incubated in either the pellet fraction or the supernatant fraction, there is no significant protection of the branch site from RNase A digestion (Fig. 3). However, incubation with both the pellet and supernatant fractions supported the binding of U2 snRNP to both substrates. Thus, even though Δ PyAG lacks the 3' splice site, the U2AF-binding site, stable U2 snRNP binding still requires auxiliary factors, presumably including U2AF (discussed below).

A 7-nucleotide branch site sequence is sufficient to direct U2 snRNP binding

The experiments presented above demonstrate that U2 snRNP can bind accurately to the branch site in the absence of a 3' splice site/polypyrimidine tract. It remained possible, however, that sequences surrounding the branch site or in exon 2 were also necessary to direct U2 snRNP binding. To address this issue we asked whether a minimal 7-nucleotide branch site sequence was sufficient for U2 snRNP binding. Because the UA-CUAAC sequence, the *S. cerevisiae* branch site, is a particularly efficient mammalian branch site (Zhuang et al. 1989), a double-stranded oligonucleotide (GGTTTAC-TAACTTCG) containing this minimal branch site was synthesized and inserted into the polylinker of the plasmid pSP73 (Promega Biotec). As a control, a DNA fragment containing the human β -globin 3' splice site/polypyrimidine tract and branch site was inserted into

Figure 3. Protection of the Δ PyAG branch site requires U2AF. Δ E1 wild-type (WT) and Δ PyAG ^{32} P-labeled RNAs were incubated with a U2 snRNP-containing fraction, a U2AF-containing fraction, or a mixture of the two fractions under splicing conditions at 23°C. Separation of U2AF and U2 snRNP was by centrifugation at high ionic strength, as described previously (Ruskin et al. 1988). All reactions were in 25 μ l of total volume containing 7.5 μ l of one fraction and 7.5 μ l of buffer D or 7.5 μ l of each fraction.

the same polylinker. The results in Figure 4 demonstrate that when incubated in a nuclear extract, an RNA containing the UACUAAC sequence can give rise to an RNase A-resistant fragment. In contrast, the same RNA lacking the UACUAAC sequence is not detectably protected from RNase A digestion. Additional control experiments similar to those shown in Figure 1 confirmed that the RNase A-resistant fragment contained the UACUAAC sequence and that the factor conferring RNase A resistance was U2 snRNP (data not shown). We conclude that a 7-nucleotide RNA sequence (UACUAAC) is sufficient to direct stable U2 snRNP binding.

The 3' splice site/polypyrimidine tract can affect the choice of potential U2 snRNP-binding sites

Although the above experiments indicate that the branch site can direct U2 snRNP binding, previous studies implicate the 3' splice site/polypyrimidine tract as the major determinant of U2 snRNP binding (Ruskin et al. 1985, 1988; Hartmuth and Barta 1988). The yeast RP51A pre-mRNA provides an ideal system to evaluate the relative importance of these two elements. Although RP51A pre-mRNA is spliced accurately in both yeast whole-cell and HeLa cell nuclear extracts, the RNA branch forms at different positions in the two systems (Ruskin et al. 1986). In yeast, the RNA branch forms at the third adenosine in the UACUAAC element, located 59 nucleotides upstream from the 3' splice site. In HeLa cell nuclear extracts, an adenosine located within the sequence UACAAAC, 37 nucleotides upstream from the 3' splice site, is used.

We analyzed binding of U2 snRNP to the wild-type RP51A pre-mRNA and to RP51A pre-mRNA substrates

Figure 4. A 7-nucleotide sequence is sufficient to direct U2 snRNP binding. Substrates labeled with ^{32}P were subjected to RNase A protection analysis, as in Fig. 1A. The structures of the substrate RNAs are diagrammed below. (PL) Polylinker; (thin black line) polylinker sequences; (black lines) β -globin intron sequences; (hatched box) β -globin exon 2.

deleted of the 3' splice site/polypyrimidine tract. The identities of the RNase A-resistant fragments resulting from U2 snRNP binding were determined by RNase T1 digestion analysis (Fig. 5). The *S. cerevisiae* branch site, UACUAAC, is contained within a unique 12-nucleotide ^{32}P -A-labeled RNase T1 fragment, whereas the branch site used in the HeLa cell extract, UACAAAC, is contained within a unique 22-nucleotide ^{32}P -A-labeled RNase T1 fragment. Figure 5 shows that using the wild-type RP51A pre-mRNA, the RNase A-resistant fragment contains a 21-nucleotide RNase T1 fragment and no detectable 12-nucleotide RNase T1 fragment. (The 21-nucleotide fragment is derived from the 22-nucleotide RNase T1 fragment; the protection from RNase A does not extend to the final nucleotide of the 22-nucleotide RNase T1 fragment. The 21' fragment results from RNase H-directed cleavage of the 22-nucleotide frag-

Figure 5. Role of the 3' splice site/polypyrimidine tract in branch site selection. The ^{32}P -labeled RNAs were cleaved by incubation in ATP-depleted nuclear extracts in the presence of the indicated oligonucleotide. The ^{32}P -labeled transcripts and an excess of appropriate oligonucleotide (C or D, Rymond et al. 1987) were added to reaction mixtures containing 40% ATP-depleted nuclear extract and 3 mM MgCl_2 and incubated at 30°C for 30 min. The cleaved RNAs were then purified by gel electrophoresis on a 5% denaturing polyacrylamide gel. RNase A-resistant, anti-Sm immunoprecipitable fragments of the RP51 and 3'-III substrates were generated as in Fig. 1A; fragments were purified and digested with RNase T1 as in Fig. 1B. The 21-nucleotide RNase T1 fragment is lacking the final base of the 22-nucleotide RNase T1 fragment due to cleavage by RNase A. 21' is generated by the oligonucleotide-directed cleavage with oligonucleotide C, which cleaves within the original 22-nucleotide RNase T1 fragment. The structures of the substrate RNAs and the identity of the oligonucleotides used for RNase H-directed cleavage are diagrammed below. The relative position and size of the RNase T1-generated fragments is also noted. (M) Complete RNase T1 digest of full-length RP51 RNA.

ment using oligonucleotide C.) Thus, in the presence of the U2AF-binding site, U2 snRNP binds only to the site at which the RNA branch forms in the HeLa cell extract. The RNase A-resistant fragments produced from the substrate deleted of the 3' splice site/polypyrimidine tract give rise to both the 12- and 21-nucleotide RNase T1 fragments, demonstrating that both potential branch sites are bound by U2 snRNP, in the absence of the U2AF-binding site. These results show that the 3' splice site/polypyrimidine tract plays a dominant role in selection of the U2 snRNP-binding site.

Next, we analyzed a mutant RP51A pre-mRNA (3'-III; Jacquier and Rosbash 1986) that contains an A → C transversion at the adenosine used for RNA branch formation in yeast (UACUACC). U2 snRNP binds to the downstream site (21-nucleotide RNase T1 fragment) but not to the mutated upstream branch site (12-nucleotide RNase T1 fragment) (Fig. 5), confirming the sequence-specific nature of this interaction.

The strength of the U2 snRNP/branch site interaction determines the efficiency of spliceosome formation and splicing

To examine the importance of the branch site sequence in the presence of a U2AF-binding site, we analyzed two previously characterized human β -globin branch site mutants. One contains an A → G transition at the adenosine normally used for RNA branch formation (A → G); the other is a substitution of the branch site with a restriction enzyme linker sequence (XR1). In both instances, the 3' splice site/polypyrimidine tract (U2AF-binding site) is normal and the mutant pre-mRNAs are accurately spliced due to activation of cryptic branch points (Ruskin et al. 1985).

The authentic branch site is a better match to the consensus (see Table 2) than either of the cryptic branch sites (see Fig. 6A). If the branch site sequence affects the efficiency of U2 snRNP binding in the presence of a U2AF-binding site, we expect decreased binding of U2 snRNP to the cryptic branch site of these mutants. Figure 6A shows that U2 snRNP indeed binds less efficiently to the branch sites of the two mutant pre-mRNAs than it does to that of the wild-type pre-mRNA.

We also measured U2 snRNP binding and spliceosome assembly, using nondenaturing gels. This gel system resolves at least two major spliceosomal complexes, one containing only U2 snRNP, and the other containing U2, U4/6, and U5 snRNPs (Konarska and Sharp 1986). Compared to wild-type pre-mRNA, both mutant pre-mRNAs are assembled more slowly and to lower levels into both spliceosomal complexes (Fig. 6B). At later times in the reaction, the levels of spliceosomal complexes formed with wild-type and mutant pre-mRNAs appear more comparable, presumably due to turnover of the spliceosome following splicing of the wild-type substrate (Konarska and Sharp 1987).

Finally, we measured the splicing efficiency of these mutants directly (Fig. 6C). Compared to wild-type pre-

mRNA, splicing of both mutant pre-mRNAs is reduced proportionately to their decrease in U2 snRNP binding and spliceosome assembly.

These results demonstrate that in the presence of a U2AF-binding site, the branch site sequence affects the efficiency of U2 snRNP binding. The strength of the U2 snRNP-branch site interaction, in turn, directly affects the efficiency of spliceosome assembly and splicing.

Discussion

In this paper we show that U2 snRNP interacts with its binding site, the branch site, in a sequence-specific manner. Below we discuss these results in conjunction with previous studies and propose a model for selection of mammalian branch sites.

Role of the 3' splice site/polypyrimidine tract in branch site selection

Three observations support the view that the primary constraint on mammalian branch site selection is relative position within the intron. First, with one notable exception (discussed below), the RNA branch forms 18–38 nucleotides upstream of the 3' splice site, regardless of intron size (Green 1986). Second, upon mutation of the authentic branch site, RNA branches form at cryptic sites, which again are located within the 18- to 38-nucleotide distance. Third, an authentic branch site can be inactivated by moving it farther upstream from the 3' splice site than 38 nucleotides (Ruskin et al. 1985). Likewise, an exceptionally efficient branch site, UACUAAAC (Zhuang et al. 1989), located 59 nucleotides upstream of the 3' splice site, is inactive in a HeLa cell extract (Ruskin et al. 1986). In fact, the distance constraint is so strong that in the absence of an adenosine residue within the 18- to 38-nucleotide range, the branch will form at a cytosine rather than at an adenosine farther upstream (Hartmuth and Barta 1988).

The data presented here are also consistent with the notion that the 3' splice site/polypyrimidine tract is the dominant factor in branch site selection: Deletion of the RP51A 3' splice site/polypyrimidine tract unmasks a new upstream branch site, which is a perfect match to the consensus. That is, in the presence of the 3' splice site/polypyrimidine tract, a consensus branch site located upstream is inactive.

This distance constraint is likely mediated by U2AF, which binds to the 3' splice site/polypyrimidine tract and is assumed to contact U2 snRNP directly. Here we show that in the absence of the 3' splice site/polypyrimidine tract, a U2AF-containing fraction is still required for stable binding of U2 snRNP. There are several possible explanations for this apparent inconsistency. For example, U2AF and U2 snRNP may initially contact one another followed by binding of this putative U2AF-U2 snRNP complex to the branch site. Alternatively, U2AF may bind nonspecifically to the pre-mRNA, followed by specific binding of U2 snRNP to the branch site. We

Figure 6. The branch site sequence affects spliceosome assembly and splicing efficiency. (A) Time course of U2 snRNP binding. The RNase A-resistant fragments of the wild-type (WT) and mutant substrate RNAs were generated as in Fig. 1A, except that the incubations with nuclear extract were at 30°C for the times indicated. The structure of the substrates is diagrammed below and the sequence of the branch region is included. The adenosine at which the RNA branch forms is underlined. (B) Spliceosome assembly is affected by branch site mutations. Electrophoretic separation of splicing complexes was carried out as described by Konarska and Sharp (1986), with the modifications described in Nelson and Green (1988). Substrate RNAs labeled with ^{32} P were incubated with nuclear extract under splicing conditions. At the indicated times (in minutes), an aliquot was removed and heparin was added to 1 μ g/ml. The complexes were fractionated on a native 3.5% polyacrylamide–0.5% agarose gel and visualized by autoradiography. (NS) Nonspecific complex;

(A) U2 snRNP-containing complex; (B) U2, U4/6, U5 snRNP-containing complex. (C) Splicing efficiency is affected by branch site mutations. Substrate RNAs labeled with ^{32} P were incubated with nuclear extract under splicing conditions for the times indicated. The RNA species were purified and fractionated on a 5% denaturing polyacrylamide gel. Because intron-containing species from each substrate migrate at different positions, these species have not been labeled.

favor this latter possibility because (1) U2AF can bind weakly to RNAs lacking a 3' splice site/polypyrimidine tract (Ruskin et al. 1988; P.D. Zamore and M.R. Green, in prep.), and (2) removal of all sequences downstream from the branch site prevents U2 snRNP binding (data not shown).

The single well-characterized example of an RNA branch forming farther than 38 nucleotides upstream from the 3' splice site is an alternatively spliced intron of the α -tropomyosin pre-mRNA (Smith and Nadal-Ginard 1989). In this case, the RNA branch forms immediately upstream of a highly pyrimidine-rich region, whereas multiple purines interrupt the actual 3' splice site/polypyrimidine tract. Previous studies have shown that the polypyrimidine tract is more important than the AG dinucleotide of the 3' splice site for RNA branch formation (Ruskin and Green 1985) and that the AG dinucleotide is not absolutely required for binding of U2AF (Ruskin et al. 1988). Thus, it is likely that even in this apparent exception, the position of the RNA branch is determined by nearby binding of U2AF.

Role of the branch site sequence in RNA branch formation

The dominant role of the 3' splice site/polypyrimidine tract in U2 snRNP binding has made it difficult to ascertain whether U2 snRNP has an intrinsic binding specificity. Although we demonstrated the sequence-specific binding of U2 snRNP, by necessity, in the absence of a 3' splice site/polypyrimidine tract, we also provide evidence for the importance of this interaction when the 3' splice site/polypyrimidine tract is present. A single-base substitution in the authentic branch site significantly decreases U2 snRNP binding. The significance of this interaction is also supported by analysis of a compilation of mammalian branch sites. Table 1 lists the mapped branch sites of 31 wild-type and mutant pre-mRNAs. There are distinct sequence preferences at multiple positions, based upon which a consensus, UN-CURAC, can be derived.

The sequence specificity of U2 snRNP binding could involve base pairing of U2 snRNA to the branch site,

Table 1. Compilation of mammalian branch site sequences

<u>Intron</u>	<u>Sequence</u>
Human β -globin IVS1	CACUGAC UCUCUCUGCCUAUUGGUCUAUUUUUCCACCCUUAG
A→G	CACUGGCUCUCUC UGCCUAU UGGUCUAUUUUUCCACCCUUAG
Δ 86	UGGUAUCAAGGU AGCCUAU UGGUCUAUUUUUCCACCCUUAG
XI-1	CUGCCUCUAGAGC UGCCUAU UGGUCUAUUUUUCCACCCUUAG
3' Δ X24	UUGGC UCUAGAGC UGCCUAU UGGUCUAUUUUUCCACCCUUAG
3' Δ X34	GAGAAGACUCUUGGC UCUAGAG GUCUAUUUUUCCACCCUUAG
XR-1	CACUC UCUAGAGC UGCCUAU UGGUCUAUUUUUCCACCCUUAG
Δ 22	AGACUCUUGGGU UUCUGAU AGGGUCUAUUUUUCCACCCUUAG
Δ 56	AGACCAA UAGAAAC UGGGCAUGGUCUAUUUUUCCACCCUUAG
Human γ -globin IVS1	UGGCACCC UUCUGAC UGUCAAAACUGUUCUUGUCAUUCACAG
Human ϵ -globin IVS1	UUGCAU CUCUAU UUUUGUAUCUGAUAGGUGUCAUUUCAUAG
Mouse β -globin IVS1	ACACUAAC UUUUCAGUGUCCCCUGUCUAUGUUUCCUUUUUAG
Rabbit β -globin IVS1	AGG UGGUGA CUUCUCUCCCCUGGGCUGUUUUAUUUUUCUCAG
Rabbit β -globin IVS2	CCCUC UGCUAAC CAUGUUCUAGCCUUCUUCUUUUUCCUACAG
3'ss LIVS-24	UAAACUUUAGC UCUAGAG CAUGCCUUCUUCUUUUUCCUACAG
mini LIVS 38/129	GC UUCUCAU AGCUAGAG CAUGCCUUCUUCUUUUUCCUACAG
mini LIVS 38/102	GU UGGGAAC CG GAGAGAG CAUGCCUUCUUCUUUUUCCUACAG
Human α -globin IVS1	CCCGGACCCAAACCCAC CCCUCAC CUCUGCUUCUCCCCGCAG
Human α -globin IVS2	GGCGGCGCGGCUUGGGCCG CACUGAC CCCUCUUCUCUGCACAG
H. Growth hormone IVS1	UUGCUCUCC GGCUC CCUCUGUUGCCUCUGGUUUCUCCCCAG
H. Growth hormone IVS4	ACCCAAG CGCUUGGCCUCUCCUUCUCCUUCACUUUGCAG
hCS-3	CUUCCUCUCCGGCUC CCUCCA UUGCCUCCGGUUUCUCCCCAG
H. Calcitonin/CGRP 1 IVS3	AUUCUGGUGCAUGGU UACUGUC UGGUUAUGUGUUUCCUGCAG
H. Calcitonin/CGRP 1 IVS4	U CACUCAC CAGAUCUUCUCUUCUUCUCCAUCCUGCAAUUCAG
Rat insulin	UACAUGUACCUUUUGCUAG CCCUCA CCCUGACUAUCUCCAG
Adenovirus 5 E1a	UUUUGUG GUUUAA AGAAUUUUGUAUUGUGAUUUUUUUAAAAG
Adenovirus major late	CUUGAUGAUGUCA UACUUA UCCUGUCCUUUUUUUCCACAG
Adenovirus E2a	UCCUCCUUCUC GACUGAC UCCAUGAUCUUUUUCUGCCUAUAG
SV40 T/t	UAAUGUGUAAACUACUG AUUCUAU UGUUUGUGUAUUUUAG
Drosophila ftz	CUCAUUG AGCUAAC CCAUUUUUUCUUUUGCUUAUGCUUACAG
Yeast RP51A	UACAAAC UUUUUAUUUUGUAUUGCUUUUCGUCAUUUUAAUAG

The sequence of the 3' end of the intron from 31 normal and mutant pre-mRNAs of which the branch site has been mapped are listed. The nucleotide at which the RNA branch forms is underlined and the branch site sequences are in boldface type. The boxed sequences represent those that are a better match to the consensus branch site than the one used. Introns indented are mutants of the gene listed directly above. (hCS-3) Human chorionic somatomammotropin. The primary references for these sequences are available on request.

recognition of the branch site by a U2 snRNP polypeptide, or interaction of an as yet unidentified branch site-binding factor with U2 snRNP. The U2 snRNA base-pairing model is attractive for several reasons. First, base pairing between U2 snRNA and the branch site has been demonstrated in *S. cerevisiae* (Parker et al. 1987) and, more recently, in mammalian cells (J. Wu and J.L. Manley, in prep.; Y. Zhuang and A.M. Weiner, in prep.). Second, the potential to form both AU and GU base pairs increases the number of sites with which an RNA can interact. Examination of Table 2 reveals that at positions +1, -1, and -3, the second most favored nucleotide would preserve base pairing. Third, RNA-RNA base pairing interactions can be tolerant to mismatches. For example, mammalian U1 snRNA base-pairs with the 5' splice site, and the sequences of mammalian 5' splice sites are quite diverse. These latter two points would help explain how U2 snRNP can bind in a sequence-specific fashion to a wide variety of sites.

According to the current model for base pairing be-

tween U2 snRNA and the branch site, the adenosine at which the RNA branch forms is unpaired (Parker et al. 1987). Thus, if RNA-RNA base pairing is the sole determinant of specificity in U2 snRNP binding, the identity of the bulged nucleotide should not affect the specificity or efficiency of this interaction. However, we find that the identity of the nucleotide at this position does affect U2 snRNP-binding efficiently (Fig. 5), suggesting that another factor, such as a U2 snRNP polypeptide, also contributes to sequence-specific binding.

A model for branch site selection

On the basis of this and previous studies, we propose a model for selection of mammalian branch sites (Fig. 7): (1) U2AF binds to the 3' splice site/polypyrimidine tract; (2) U2 snRNP is recruited to the U2AF/pre-mRNA complex, presumably through interaction with bound U2AF; (3) U2 snRNP positioned near the 3' splice site/polypyrimidine tract binds stably to the highest affinity site

Table 2. A mammalian branch site consensus

Nucleotide frequency							
	-5	-4	-3	-2	-1	BN	+1
U	15	5	8	17	4	1	8
C	8	9	20	6	4	1	15
A	3	10	0	8	10	29	1
G	5	7	3	0	13	0	7
Consensus							
5'UNCURAC3'							
3'AUGAU—G5'							
U2 snRNA sequence							

The frequency with which a nucleotide appears at each position within the branch sites in Table 1 has been tabulated. A consensus has been derived, based on this compilation. The consensus is aligned with the region of U2 snRNA that has been shown to base-pair to the UACUAAC sequence in yeast introns, listed below the consensus. (BN) Branch nucleotide; (—) no nucleotide; the A from the branch sequence is presumably bulged out. (R) purine; (N) any nucleotide.

within the 18- to 38-nucleotide range. This model predicts that the branch site is the best match to the consensus within 18–38 nucleotides upstream of the 3' splice site. In 27 of 31 cases, this enables the branch site to be predicted correctly.

In contrast to mammalian cells, formation of the RNA branch in yeast does not require the 3' splice site/polypyrimidine tract (Rymond and Rosbash 1985; Cellini et al. 1986; Fouser and Friesen 1987). Thus, in yeast, apparently only one sequence element, the branch site, directs RNA branch formation. Accordingly, single-base substitutions in branch sites are generally more deleterious in yeast than in mammalian cells.

Branch site sequence and spliceosome assembly

Our results indicate that the strength of the U2 snRNP/branch site interaction is related to the efficiency of spliceosome formation and splicing. This reinforces the view that U2 snRNP is an early and, perhaps, the rate-limiting (Bindereif and Green 1987) step in spliceosome assembly. We note that another study (Reed and Maniatis 1988) did not observe an effect of branch site sequence on spliceosome assembly. This discrepancy may be due to differences in the assays used for spliceosome assembly, the reaction times when spliceosome assembly was monitored, and other aspects of the experimental design.

It is conceivable that the branch site sequence has a function(s) in addition to that of a U2 snRNP-binding site. For example, a point mutant in the yeast branch site (UACUAAC > UACUACC) has a more severe effect on splicing than it does on formation of the U2 snRNP/pre-mRNA complex (Pikielny et al. 1986). Fur-

Figure 7. A model for the selection of mammalian branch sites. U2AF binds to the 3' splice site/polypyrimidine tract of the pre-mRNA and recruits U2 snRNP. U2 snRNP then selects the best site available within the U2AF-imposed distance constraint and stably binds to the pre-mRNA. The large A is the adenosine used for RNA branch formation; the small As represent other potential branch sites.

thermore, some mammalian branch site mutants are blocked following 5' splice site cleavage and lariat formation (Hornig et al. 1986; Freyer et al. 1987). Whether these effects are all a consequence of U2 snRNP binding or are due to interactions of other splicing components with the branch site remains to be determined.

Methods

Materials

SP6 polymerase, RNasin, DNase I, AMV reverse transcriptase, DNA ligase, and restriction enzymes were from Promega Biotec or New England BioLabs. GpppG and ribonucleotides and deoxynucleotides were from Pharmacia. RNase A was from Boehringer–Mannheim Biochemicals. RNase T1 was from Calbiochem. Heparin was from Sigma. [α - ^{32} P]UTP (410 Ci/mmol) was purchased from Amersham or New England Nuclear. Anti-Sm serum was purchased from Vitrotec Laboratories, Inc. The anti-70-kD antibodies were a generous gift of S. Hoch (Billings et al. 1982), and the anti-U1/U2 snRNP antibody was a gift of W. van Venrooij (Habets et al. 1985). The oligonucleotides used for oligonucleotide-directed RNase H cleavage of the RP51 substrates were a generous gift of Brian Rymond (Rymond et al. 1987).

RNA substrates

The wild-type β -globin (pSP64H β Δ 6, Krainer et al. 1984) and Δ PyAG (Ruskin and Green 1985c) have been described previously. The Δ E1 versions include *Bst*NI–*Bam*HI (+136 to +477; Lawn et al. 1980) in pSP64. The RP51A wild-type substrate was transcribed from 5'–20 (Pikielny and Rosbash 1985), and the mutant from 3'–III (Jacquier and Rosbash 1986). PL-YBP was constructed by inserting a blunt-ended, double-stranded oligonucleotide (GGTTTACTAACTTCG) into the *Sma*I site of the polylinker of pSP73 (Promega Biotec). PL- β XS was constructed by inserting a blunt-ended *Xba*I–*Sin*I (+222 to +293; Lawn et al. 1980) fragment from NL-X (Ruskin et al. 1985) into the *Sma*I site of the pSP73 polylinker. For in vitro transcription with SP6 polymerase, these templates were linearized with *Pvu*II.

RNase protection assays

A modified RNase A protection assay (Ruskin and Green 1985a) was used. The incubations of the RNA with nuclear extract were carried out at 23°C, unless noted otherwise. The RNase A treatment and immunoprecipitation were as described previously (Ruskin et al. 1988). The antibody used for immunoprecipitation is a polyclonal α -Sm, unless noted otherwise.

Acknowledgments

We thank W. van Venrooij and S. Hock for valuable immunological reagents, C. Pikielny for clones, and B. Rymond for oligonucleotides. We gratefully acknowledge J. Lillie, C. Pikielny, and other members of the laboratory for providing critical comments on the manuscript. K.K.N. was supported by a National Science Foundation predoctoral training grant. This work was supported by grants from the National Institutes of Health and the Chicago Community Trust/Searle Scholars program to M.R.G.

References

- Bindereif, A. and M.R. Green. 1987. An ordered pathway of snRNP binding during mammalian pre-mRNA splicing complex assembly. *EMBO J.* **6**: 2415–2424.
- Billings, P.B., R.W. Allen, F.C. Jensen, and S.O. Hoch. 1982. Anti-RNP monoclonal antibody derived from a mouse strain with lupus-like autoimmunity. *J. Immunol.* **128**: 1176–1180.
- Black, D.L., B. Chabot, and J.A. Steitz. 1985. U2 as well as U1 small ribonucleoprotein particles are involved in pre-mRNA splicing. *Cell* **42**: 737–750.
- Cellini, A., E. Felder, and J.J. Rossi. 1986. Yeast pre-messenger RNA splicing efficiency depends on critical spacing requirements between branch point and 3' splice site. *EMBO J.* **5**: 1023–1030.
- Chabot, B. and J.A. Steitz. 1987. Multiple interactions between the splicing substrate and small nuclear ribonucleoproteins in spliceosomes. *Mol. Cell Biol.* **7**: 281–293.
- Fouser, L.A. and J.D. Friesen. 1987. Mutations in a yeast intron demonstrate the importance of specific conserved nucleotides for the two stages of nuclear mRNA splicing. *Cell* **45**: 81–93.
- Freyer, G.A., J. Arenas, K. Perkins, H.M. Furneaux, L. Pick, B. Young, R.J. Roberts, and J. Hurwitz. 1987. In vitro formation of a lariat structure containing a G^{2',5'}G linkage. *J. Biol. Chem.* **262**: 4267–4273.
- Green, M.R. 1986. Pre-mRNA splicing. *Annu. Rev. Genet.* **20**: 671–708.
- Guthrie, C. and B. Patterson. 1988. Spliceosomal snRNAs. *Annu. Rev. Genet.* **22**: 387–419.
- Habets, W., M. Hoet, P. Bringmann, R. Lurhmann, and W. van Venrooij. 1985. Autoantibodies to ribonucleoprotein particles containing U2 small nuclear RNA. *EMBO J.* **4**: 1545–1550.
- Hartmuth, K. and A. Barta. 1988. Unusual branch point selection in processing of human growth hormone pre-mRNA. *Mol. Cell Biol.* **8**: 2011–2020.
- Hornig, H., M. Aebi, and C. Weissmann. 1986. Effect of mutations at the lariat branch acceptor site on Δ -globin pre-mRNA splicing in vitro. *Nature* **324**: 589–592.
- Jacquier, A. and M. Rosbash. 1986. Dramatic effect of a mutant yeast branch point on splicing and intron turnover. *Proc. Natl. Acad. Sci.* **83**: 5835–5839.
- 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 nucleoprotein particles in formation of spliceosomes. *Cell* **49**: 763–774.
- Krainer, A.R., T. Maniatis, B. Ruskin, and M.R. Green. 1984. Normal and mutant human Δ -globin pre-mRNAs are faithfully and efficiently spliced in vitro. *Cell* **36**: 993–1005.
- Kramer, A. 1988. Presplicing complex formation requires two proteins and U2 snRNP. *Genes Dev.* **2**: 1155–1167.
- Lawn, R.M., A. Efstratiadis, C. O'Connell, and T. Maniatis. 1980. The nucleotide sequence of the human β -globin gene. *Cell* **21**: 647–651.
- Maniatis, T. and R. Reed. 1987. The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing. *Nature* **325**: 673–678.
- Nelson, K.K. and M.R. Green. 1988. Splice site selection and ribonucleoprotein complex assembly during in vitro pre-mRNA splicing. *Genes Dev.* **2**: 319–329.
- Padgett, R.A., P.J. Grabowski, M.M. Konarska, and P.A. Sharp. 1987. Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* **55**: 1119–1150.
- Padgett, R.A., M.M. Konarska, M. Aebi, H. Hornig, C. Weissmann, and P.A. Sharp. 1986. Nonconsensus branch-site sequences in the in vitro splicing of transcripts of mutant rabbit β -globin genes. *Proc. Natl. Acad. Sci.* **82**: 8349–8353.
- Parker, R., P.G. Siliciano, and C. Guthrie. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNP. *Cell* **49**: 229–239.
- Pikielny, C.W. and M. Rosbash. 1985. mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell* **41**: 119–126.
- Pikielny, C.W., B.C. Rymond, and M. Rosbash. 1986. Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes. *Nature* **324**: 341–345.
- Reed, R. and T. Maniatis. 1988. Role of the mammalian branch point sequences in pre-mRNA splicing. *Genes Dev.* **2**: 1268–1276.
- Ruskin, B. and M.R. Green. 1985a. Specific and stable intron-factor interactions are established early during in vitro pre-mRNA splicing. *Cell* **43**: 131–142.
- . 1985b. Detection and characterization of an RNA processing activity that debranches RNA lariats. *Science* **229**: 135–140.
- . 1985c. Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing. *Nature* **317**: 732–734.
- Ruskin, B., J.M. Green, and M.R. Green. 1985. Cryptic activa-

- tion allows accurate in vitro splicing of human β -globin intron mutants. *Cell* **41**: 833–844.
- Ruskin, B., P.D. Zamore, and M.R. Green. 1988. A factor, U2AF is required for U2 binding and splicing complex assembly. *Cell* **52**: 207–219.
- Ruskin, B., A.R. Krainer, T. Maniatis, and M.R. Green. 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing. *Cell* **38**: 317–331.
- Ruskin, B., C.W. Pikielny, M. Rosbash, and M.R. Green. 1986. Alternative branchpoints are selected during splicing of a yeast pre-mRNA in mammalian and yeast extracts. *Proc. Natl. Acad. Sci.* **83**: 2022–2026.
- Rymond, B.C. and M. Rosbash. 1985. Cleavage of 5' splice site and lariat formation are independent of 3' splice site in yeast mRNA splicing. *Nature* **317**: 735–737.
- Rymond, B.C., D.D. Torrey, and M. Rosbash. 1987. A novel role for the 3' region of introns in pre-mRNA splicing of *Saccharomyces cerevisiae*. *Genes Dev.* **1**: 238–246.
- Sharp, P.A. 1987. Splicing of messenger RNA precursors. *Science* **235**: 766–771.
- Smith, C.W.J. and B. Nadal-Ginard. 1989. Mutually exclusive splicing of α -tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. *Cell* **56**: 749–758.
- Zhuang, Y. and A.M. Weiner. 1989. A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. *Cell* **46**: 827–835.
- Zhuang, Y., A.M. Goldstein, and A.M. Weiner. 1989. UA-CUAAC is the preferred branch site for mammalian mRNA splicing. *Proc. Natl. Acad. Sci.* **86**: 2752–2756.

Mammalian U2 snRNP has a sequence-specific RNA-binding activity.

K K Nelson and M R Green

Genes Dev. 1989, **3**:

Access the most recent version at doi:[10.1101/gad.3.10.1562](https://doi.org/10.1101/gad.3.10.1562)

References

This article cites 39 articles, 13 of which can be accessed free at:
<http://genesdev.cshlp.org/content/3/10/1562.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](#).