REPORT

Branchpoint selection in the splicing of U12-dependent introns in vitro

TIMOTHY S. MCCONNELL,^{1,2} SOO-JIN CHO,^{1,3} MIKKO J. FRILANDER,^{1,4}
and JOAN A. STEITZ¹

¹Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute,
Yale University, New Haven, Connecticut 06536, USA

ABSTRACT

In metazoans, splicing of introns from pre-mRNAs can occur by two pathways: the major U2-dependent or the minor U12-dependent pathways. Whereas the U2-dependent pathway has been well characterized, much about the U12-dependent pathway remains to be discovered. Most of the information regarding U12-type introns has come from in vitro studies of a very few known introns of this class. To expand our understanding of U12-type splicing, especially to test the hypothesis that the simple base-pairing mechanism between the intron and U12 snRNA defines the branchpoint of U12-dependent introns, additional in vitro splicing substrates were created from three putative U12-type introns: the third intron of the *Xenopus* RPL1a gene (XRP), the sixth intron of the *Xenopus* TFIIS.oA gene (XTF), and the first intron of the human Sm E gene (SME). In vitro splicing in HeLa nuclear extract confirmed U12-dependent splicing of each of these introns. Surprisingly, branchpoint mapping of the XRP splicing intermediate shows use of the upstream rather than the downstream of two consecutive adenosines within the branchpoint sequence (BPS), contrary to the prediction based on alignment with the sixth intron of human P120, a U12-dependent intron whose branch site was previously determined. Also, in the SME intron, the position of the branchpoint A residue within the region base paired with U12 differs from that in P120 and XTF. Analysis of these three additional introns therefore rules out simple models for branchpoint selection by the U12-type spliceosome.

Keywords: ATAC spliceosome; intron branch site; minor class introns; snRNPs; splicing mechanism

INTRODUCTION

In higher eukaryotes, a minority of introns are excised by a distinct low-abundance spliceosome that involves the U12 snRNP. Although not as well characterized as its classical U2-dependent counterpart, the U12-dependent spliceosome has provided general insights into the mechanism of pre-mRNA splicing. The assembly of the U12-dependent spliceosome follows a pathway comparable to that of the major spliceosome, with

the U11, U12, U4atac, and U6atac snRNPs replacing the U1, U2, U4, and U6 snRNPs (Tarn & Steitz, 1997; Burge et al., 1999). U5 is the only snRNA shared by the two spliceosomes, but the list of common protein factors is growing (Luo et al., 1999; Will et al., 1999, 2001). Although the corresponding snRNAs are quite different in sequence, their secondary structures as well as the snRNA–snRNA and snRNA–substrate interactions they undergo during the course of splicing are analogous in the two systems (Tarn & Steitz, 1996a; Frilander & Steitz, 2001). Recently, it has been shown that the *Drosophila melanogaster* U12 and U6atac genes are required and play essential roles in development (Otake et al., 2002).

During the splicing of U12-dependent introns, the branchpoint sequence is recognized by base pairing with the U12 snRNP in one of the earliest steps committing an intron to the minor splicing pathway (Hall & Padgett, 1996; Tarn & Steitz, 1996b; Frilander & Steitz,

Reprint requests to: Joan Steitz, Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06536, USA; e-mail: joan.steitz@yale.edu.

²Present address: Rib-X Pharmaceuticals, 25 Science Park, Suite 611, New Haven Connecticut 06511, USA.

³Present address: Medical Scientist Training Program, Washington University School of Medicine, St. Louis, Missouri 63110, USA.

⁴Present address: Institute of Biotechnology, PL56 (Viikinkaari 9), 00014 University of Helsinki, Finland.

1999). U2-dependent introns have loosely conserved sequences at the branchpoint as well as at the 5' splice site and 3' splice site (Burge et al., 1999; Moore, 2000; Reed, 2000), with the most conserved features being the terminal dinucleotides GT and AG. In contrast, U12-dependent introns contain more highly conserved sequences at the 5' splice site and branchpoint that are different from the U2-type consensus sequences (Jackson, 1991; Hall & Padgett, 1994; Sharp & Burge, 1997; Tarn & Steitz, 1997). Their preferred terminal dinucleotides are either AT-AC or GT-AG (Dietrich et al., 1997; Sharp & Burge, 1997; Levine & Durbin, 2001), with only an 8 to 21 nt spacing between the branch site and the 3' splice site (Dietrich et al., 2001a; Levine & Durbin, 2001). An early bioinformatic study identified at least 60 nonredundant genes containing U12-dependent introns from a variety of organisms (Burge et al., 1998). More recently, the completed human genome sequence was scanned, yielding over 400 candidates for U12-type introns (Levine & Durbin, 2001). However, only a few of these introns—from the human nucleolar protein (P120) gene (Hall & Padgett, 1996; Tarn & Steitz, 1996b), the human gene encoding the voltage-gated skeletal muscle sodium channel α subunit (SCN4A; Wu & Krainer, 1996), human ADPRP (Dietrich et al., 1997), and three *Drosophila* genes (*prospero*, involved in neural development, CT23545, and CT36969; Otake et al., 2002)—have actually been demonstrated to splice in a U12-dependent manner.

In U2-dependent splicing, there is strong preference for adenosine as the branchpoint nucleophile. When the A within the branchpoint consensus sequence, YURAC, is mutated to another nucleotide, a nearby A is used instead of the mutated residue (Query et al., 1994, 1996). In U12-dependent introns, the branchpoint sequence (BPS), UUUCCUUAACYCY, is longer and more highly conserved (Burge et al., 1998) and exhibits extensive complementarity to U12 snRNA (Hall & Padgett, 1996; Tarn & Steitz, 1996b). This suggests that U12 is

more stringent than U2 in its selection of the BPS and that the selection may be decided at the level of U12•BPS base pairing. However, to this point, the identity of the branchpoint nucleotide has been determined only for P120 (Tarn & Steitz, 1996b) of all the U12-type introns so far identified.

Here, the *in vitro* splicing of three U12-dependent introns was investigated using HeLa nuclear extracts. The substrates were derived from the genes for *Xenopus* RPL1a (XRP), a ribosomal protein, *Xenopus* TFIIIS.oA (XTF), a transcription factor, and human Sm E (SME), a core snRNP protein, each representing a different machinery in the gene expression pathway. We demonstrate that the three substrates all assemble into U12-type spliceosomes *in vitro* and identify the branchpoint nucleotide used in their excision. The mapping of these branchpoint adenosines challenges the simplicity of current models for the U12•BPS interaction.

RESULTS

U12-dependent introns

To expand the repertoire of introns demonstrated to splice via the U12-dependent pathway, we selected candidates for their variety of BPS sequence, original intron position, and gene function. We constructed three new splicing substrates containing at least 41 nt of 5' and 3' exon sequences flanking the U12-type intron (Table 1). To enhance *in vitro* splicing, the lengths of the introns were shortened by truncation upstream of the putative branchpoint region for the XTF and SME substrates (see Materials and Methods). All three introns exhibit strong U12-type consensus sequences at the 5', 3', and branchpoint regions. XTF follows the P120 hallmark of AT-AC terminal nucleotides, whereas XRP and SME have GT-AG termini, like most U2-dependent introns. Each intron does, however, present some challenges for selection of the branchpoint aden-

TABLE 1. Information on the genes containing U12-type introns used in this study.^a

	P120	XRP	XTF	SME
GenBank accession #	M33132	X06552	X97658	M21258
Organism	<i>Homo sapiens</i>	<i>Xenopus laevis</i>	<i>X. laevis</i>	<i>H. sapiens</i>
Intron	6 (14)	3 (9)	6 (9)	1 (4)
Intron size	99	226 (241)	117 (593)	124 (481)
5' exon size	158	108	59	111
3' exon size	100	144	155	41
Substrate size	357	478	331	285
5' splice site	agg at atccttgcagg	aac gt atccttttggg	aag at atcctttttat	atc gt atcctacgcag
Branch site	cagttcctta ac agg	ttttcctta at att	atattcctta aa tcc	ttttccttagcc ac
3' splice site	tccttaacaggccc ac atg	ttaatatctcatg ag atg	ttaaatacccttgc ac cg-	agccactgtggtg ag aa

^aThe intron number is that intron which is U12 dependent and the number in parentheses is the total number in the gene. The total lengths of some introns were reduced to enhance *in vitro* splicing; shown is the actual length of the intron used and the natural intron size in parentheses. The bold letters of the splice sites indicate the terminal dinucleotides of the introns and the bold a in the branch site is the nucleophile for the first step of slicing.

osine. Whereas XRP has an apparently generic U12 consensus branchpoint sequence, XTF has three consecutive adenosines, and SME has a G at the predicted branchpoint position with an A immediately upstream (see Table 1).

The U12-dependent spliceosome assembles on the XRP, XTF, and SME substrates

During assembly of the U12 spliceosome on the P120 intron, RNA–RNA interactions that can be detected by psoralen crosslinking form early between U12 and the BPS, followed by a second set of psoralen-detectable interactions corresponding to the formation of helices Ia and Ib between U12 and U6_{atac} in the active site of the spliceosome (Tarn & Steitz, 1996a, 1996b; Frilander & Steitz, 2001). We first asked whether similar crosslinks could be detected for the XRP, XTF, and SME splicing substrates and whether the formation of these crosslinks would require the expected U12/BPS

interactions. Even though XRP and XTF originate from *Xenopus*, we used HeLa nuclear extract because previous attempts in our laboratory failed to produce observable U12-dependent splicing in *Xenopus* oocytes (Y.-T. Yu, M.-D. Shu, & J.A. Steitz, unpubl. results). As seen in Figure 1, in the presence of a 2'-O-methyl oligoribonucleotide complementary to the U2 BPS binding region (U2b), which prevents the nonproductive interaction of the U2 snRNP with the substrate (Tarn & Steitz, 1996b), we observed the formation of psoralen crosslinks between U12 and other RNA constituents by probing a northern blot for U12. Subsequent probing for U6_{atac} and for each of the three new substrates plus the P120 control (data not shown) identified the cross-linking partner of U12 in each band. Figure 1, lanes 3 and 4, show that both a U12•BPS and two U12•U6_{atac} crosslinks form with the control P120 substrate and do not appear when U12 is blocked by a 2'-O-methyloligoribonucleotide complementary to the BPS-pairing region of U12 (U12a, Fig. 1, lane 5). The SME substrate

FIGURE 1. U12- and U6_{atac}-specific crosslinks form with the SME, XTF, and XRP substrates. Splicing reactions containing 2'-O-methyl oligonucleotides complementary to U12 (U12a, or a above the lane) or U2 (U2b, or b above the lane) were exposed to 365 nm light (indicated by + above the lane) in the presence of psoralen. Purified RNA was resolved on a 5% denaturing gel, transferred to nylon membrane, and the resulting northern blot probed for U12 snRNA. A cross-reacting band (*) detected even in the absence of UV irradiation either ran below or comigrated with the lower U12•U6_{atac} crosslink (helix Ib), depending on the experiment. Lane 1 contains a molecular weight marker, a ³²P-labeled pBR322 *Msp*I digest. Lane 2 is a control reaction without any added substrate RNA.

(Fig. 1, lanes 7 and 8) forms a crosslink with U12 that has a faster mobility, as predicted from its smaller size. Again, this crosslink and the other two crosslinks diagnostic of the active site helices Ia and Ib between U12 and U6_{atac} appear only in the absence of the U12-blocking oligonucleotide (Fig. 1, compare lanes 8 and 9). XRP (Fig. 1, lanes 12 and 13) and XTF (Fig. 1, lanes 17 and 18) form similar U12•BPS and U12•U6_{atac} crosslinks in a U12-dependent fashion (Fig. 1, lanes 14 and 19), although the lower U12•U6_{atac} crosslink is partially obscured by a cross-reacting band in this particular U12-probed northern blot (Fig. 1, lanes 10–19). In earlier experiments, we found that constructs containing the full-length introns of XRP and XTF similarly formed U12•BPS and U12•U6_{atac} crosslinks (data not shown), consistent with excision via the U12-dependent pathway. The truncated XRP and XTF splicing substrates were used in this and later experiments because of their enhanced splicing activity and because their lariat intermediates were readily resolved in 10% gels, thus allowing branch site mapping by primer extension (see below).

In other respects, the three introns behave like other U12-dependent introns studied. Stimulation of splicing by the presence of flanking introns, as described for the sodium channel intron (Wu & Krainer, 1998, 1999), was confirmed by adding an upstream polypyrimidine tract or a downstream U1 binding site; in each case, a two- to fivefold increase in spliceosome assembly as measured by U12-dependent crosslinks was observed (data not shown). Chimeric substrates created from P120 and XRP sequences to test the role of a purine-rich sequence in the 5' exon of P120 exhibited moderately enhanced spliceosome assembly. Exonic enhancers have been demonstrated to modulate U12-type splicing (Dietrich et al., 2001b; Hastings & Krainer, 2001), and this putative exonic splicing enhancer fits the GAR repeat consensus for SR protein binding (Lavigne et al., 1993; Sun et al., 1993; Watakabe et al., 1993; Ramchatesingh et al., 1995).

Identifying lariat intermediates and mapping the branchpoint nucleotides

The truncated forms of the XRP and XTF introns, as well as the full-length SME intron, yielded observable splicing intermediates and products in the HeLa *in vitro* system. On a 10% denaturing gel, lariat intermediate was best visualized by 3'-end labeling of the substrate, as shown for XRP in Figure 2A. A 5-h splicing reaction produced a measurable amount of a slowly migrating band (LI, Fig. 2A, lane 5), not seen when the extract was preincubated with the U12a oligonucleotide (Fig. 2A, lane 1). Subsequent treatment with ribonuclease H and a DNA oligonucleotide complementary to the first exon (E1, Fig. 2A, lane 4), the intron (I, Fig. 2A, lane 3) or the second exon (E2, Fig. 2A, lane 2) identified this slowly

migrating band as lariat intermediate. (Note that a majority of substrate was cleaved to a distinct RNase H product in each case.) Similar characterization identified the lariat intermediates generated during the splicing of the XTF and SME U12-type introns (data not shown).

To obtain adequate amounts of U12-dependent lariat intermediates for analysis, large scale splicing reactions (200 μ L) were resolved (as in Fig. 2A), and the lariat intermediates were isolated and subjected to primer extension with reverse transcriptase to identify their branchpoint nucleotides. In each case, a 5'-³²P-labeled deoxyoligonucleotide complementary to exon 2 was used as primer and the products were run in parallel with a DNA sequencing ladder (Fig. 2B).

For SME, primer extension produced a stop that aligns with the G residue (Fig. 2B, lane 5) immediately downstream of the single A residue present in the BPS (Table 1). The conclusion that this A residue acts as nucleophile for the first step of splicing is supported by the lack of any significant primer extension stop in the mock-treated (Fig. 2B, lane 6) or pre-mRNA alone (data not shown) control reactions.

With the XTF substrate, the predicted branchpoint adenosine is situated in the middle of a three-adenosine stretch—a configuration that would provide a different challenge in branchpoint selection for the U12-dependent spliceosome. Here, mapping of the branchpoint showed a stop that aligns with the downstream A (Fig. 2B, compare lanes 16 and 18), indicating that the predicted middle A is the branchpoint nucleotide.

Surprisingly, the intron with the most generic BPS, XRP, produced the most unusual reverse transcription stop. The band (Fig. 2B, lane 11) aligns with the downstream A (Fig. 2B, lane 10), indicating that the upstream A, instead of the expected nucleotide, was utilized as the branchpoint. Thus, the XRP•U12 duplex resembles the SME•U12 interaction rather than the anticipated duplex of P120•U12 or XTF•U12 (Fig. 3), even though an apparently suitable A residue is present in the downstream (expected) position.

DISCUSSION

Because of the extensive potential base pairing with U12 evident just upstream of the putative branchpoint A residue in virtually all U12-type introns, our expectation was that branchpoint nucleotide selection would be easily predictable. Yet, the four experimentally determined branch sites compiled in Figure 3 demonstrate that this is not the case. Two introns (P120 and XTF) follow one rule, whereas the other two (XRP and SME) follow another with regard to the position of the branch site A relative to the register of base pairing between the intron and U12. For the SME intron, only a single A is present: Its utilization shows that strong preference for an A as the branchpoint nucleotide is

FIGURE 2. Mapping the branchpoints of the SME, XRP, and XTF introns. **A:** A splicing reaction containing 3' ^{32}P -end labeled XRP pre-mRNA was incubated 5 h in the presence of 2'-*O*-methyl oligonucleotide U2b (lanes 2–5) or U12a (lane 1). Subsequently, reactions containing U2b 2'-*O*-methyl oligonucleotide were treated with ribonuclease H and a DNA oligonucleotide complementary to the 5' exon, the intron, or the 3' exon (lanes 2–4). The identified lariat intermediate (LI) was separated from the pre-mRNA on a 10% denaturing polyacrylamide gel. **B:** Reverse transcription of isolated lariat intermediates (LI) like that shown in **A** (lane 5) was performed with ^{32}P -labeled DNA oligonucleotides complementary to exon 2 of each substrate. Sequencing reactions (ACGT) of the corresponding plasmid DNA, and primer extension reactions of mock-treated (M) RNA extracted from another region of the same gel lane and of the pre-mRNA substrate (R) are shown.

conserved between the U2 and U12 spliceosomes. However, the unexpected observation that XRP uses the upstream rather than the downstream A argues against any simple explanation for choice of the nucleophile for the first step of U12-dependent splicing.

We can rule out several trivial explanations for the apparent complexity of the situation. First, the primer extension blockage results have been reproduced over a period of greater than a year by different experimenters and appear unambiguous (see Fig. 2B). If both As in XRP were used with some frequency, an additional stop 1 nt upstream would have been detected. Second, the two *Xenopus* introns fall into two different classes, so a species specificity in branch-site selection is not the explanation. Third, the two artificially shortened introns (XTF and SME) likewise fall into two different

classes, arguing against the possibility that manipulation has altered branch-site selection in the splicing substrate from that occurring in the original intron. Fourth, the distances between the observed branch-point A residue and the 3' splice sites all fall within the 8 to 21 nt spacing recently deduced from bioinformatic analyses of U12-type introns (Dietrich et al., 2001a; Levine & Durbin, 2001). Fifth, neither the identity nor the strength of any BPS•U12 base pairs that might potentially form proximal to the 3' splice site can explain the grouping into two different classes. Note that, even in the U2-dependent spliceosome, the base pairing of the nucleotide 3' to the branch site A residue has not been established (Query et al., 1994).

Although we cannot pinpoint the exact basis for the differential branchpoint usage, we do observe two in-

FIGURE 3. Branchpoint selection by the U12-dependent spliceosome. The schematic shows the terminal dinucleotides and potential contiguous base pairing of U12 with each intron whose branchpoint has been mapped. Lines and dots indicate Watson–Crick and AG or GU pairs, respectively. The nucleophilic branchpoint A is in larger font. The lengths (Σ) are the total number of nucleotides between the nucleophilic A and the 3' splice site. The U12-type consensus sequence is from Burge et al. (1998). The 3' sequence of U12 shown (AUAAC) can alternatively base pair with U6_{atac} to form a putative U6_{atac}/U12 helix III (Sun & Manley, 1995; Tarn & Steitz, 1996a).

dependent correlations from our data set. First, by comparing the potential for U12•BPS snRNA base pairing among the different substrates used in this study (see Fig. 3), it can be seen that the contiguous intermolecular Watson–Crick base pairing for XRP and SME extends upstream of the BPS beyond the consensus defined by Burge et al. (1998), whereas for P120 and XTF, it is shorter than the consensus. This suggests that a stronger U12•BPS interaction may lead to the use of the upstream A whereas a weaker interaction activates the downstream A. A second correlation evident in Figure 3 is between the terminal dinucleotides of each intron and branch site usage: Among the splicing substrates examined in our study, the GU–AG introns (SME and XRP) choose the upstream A whereas AU–AC introns (P120 and XTF) select the downstream A as the branchpoint for the U12-type spliceosome. Because the BPS recognition by U12 does require simultaneous binding of the U11 snRNP at the 5' splice site (Frilander & Steitz, 1999), the sequence at the 5' end of the intron could change the affinity of the U12 snRNP for the branch site—thus providing a possible explanation for the observed correlation.

An alternative possibility is that any BPS recognition step occurring prior to U12 snRNP association might determine branch nucleotide selection. In the U2-dependent spliceosome, two protein interactions are believed to assist in the choice and bulging of the correct A. The structure of splicing factor 1 (SF1) associated with a canonical U2-dependent BPS has recently been solved by NMR (Liu et al., 2001) and reveals a bulged A in the anticipated position (UAUACUAACAA); but it is not known whether SF1 or a similar factor is used by the U12-type spliceosome. A second protein,

p14, can be crosslinked to the branch site adenine in the U2-dependent pre-spliceosome and spliceosome (at least until the first step; MacMillan et al., 1994); p14 is also present in the U12-type spliceosome (Will et al., 2001), but nothing that is known concerning this factor would explain our unanticipated results.

In the U2-dependent splicing system, Query et al. (1994) carried out detailed analyses of branch-site formation and concluded that either of two adjacent purines within a strong branchpoint consensus sequence (UGCURRC) could be used alternatively. In the U12-dependent system, the BPS consensus sequence is much more conserved (UUUCCUUAACYCY), but interestingly, the two positions in the U2-dependent system correspond to those we have documented in the U12-dependent system (see Fig. 3). However, despite the high level of conservation in the U12-dependent system, there is still some flexibility to permit G in either purine position. For example, SME is one of at least four U12-dependent introns that has a G at the downstream purine position of the putative BPS consensus, whereas at least eight have a G upstream of the putative branchpoint A in the original set of U12-type introns compiled by Sharp and Burge (1997). Although the unexpected selection observed for the XRP substrate is still not explained, the parallels between the two splicing systems are underscored.

A mechanistic explanation for our observations may come from the recent determination of the crystal structure of domains 5 and 6 of a group II self-splicing intron. The structural data unexpectedly reveal that not only the nucleophilic adenosine but also its 3' neighbor are bulged from the domain 6 helix (Zhang & Doudna, 2002). Perhaps this is also true of the U12-dependent (and the U2-dependent) spliceosome, allowing flexibility in the choice between two adjacent A residues.

MATERIALS AND METHODS

Plasmids and splicing substrates

The XRP construct was created by amplifying a portion of the *Xenopus* RPL1a gene from *Xenopus* total genomic DNA using nested PCR. A fragment containing nt 2331 to 2817 of the RPL1a gene (according to GenBank accession number X06552) and a promoter for T7 RNA polymerase were subcloned into the *Hind*III and *Eco*RI sites of pSP64 adding 5 nt to the transcript at the 3' end. The sequence agreed with the GenBank sequence with the exception of two deletions in the intron—one of 2 nt 21 residues downstream of the 5' splice site (2459–2460) and another of 13 nt 41 residues upstream of the branch site (2610–2623). The XTF construct was created by likewise amplifying nt 155 to 260 and 747 to 962 of the *Xenopus* TFIIIS.oA gene (GenBank accession number X97662). The 5' and 3' fragments contained heterologous *Spe*I sites at their 3' and 5' ends. The two fragments were digested with *Hind*III, *Spe*I, and *Sal*I and subcloned into the *Hind*III and *Sal*I sites of pSP64. The resulting transcript added

2 Gs at the 5' end for T7 initiation, the 6 nt of the *SpeI* site in the intron and 4 additional nucleotides of the *SalI* site at the 3' end. The SME construct was created similarly from human total genomic DNA and contained nt 730 to 902, a heterologous *SpeI* site, and nt 1260 to 1348 (GenBank accession number M21253). The resulting transcript added 2 Gs at the 5' end for T7 initiation, the 6 nt of the *SpeI* site in the intron and 5 additional nucleotides of an *XbaI* site at the 3' end. All constructs were confirmed by sequence analysis (Keck DNA Sequencing Lab, Yale University).

Plasmids were linearized by cleaving at the 3' restriction site (*EcoRI* for XRP, *SalI* for XTF, and *XbaI* for SME) to create templates for in vitro transcription. Reactions were carried out in 250 μ L containing: 600 μ M ATP, CTP, and UTP; 60 μ M GTP; 1.6 mM GpppG (cap), $\sim$ 200 ng/ μ L template DNA, and approximately 4 U/ μ L T7 RNA polymerase. Substrates were purified on a 5% or 6% denaturing polyacrylamide gel. For splicing and primer extension reactions, the substrates were 3'-end labeled with α - 32 P-cordycepin (Lingner & Keller, 1993) and subsequently purified on a 6% denaturing polyacrylamide gel.

Splicing reactions and ribonuclease H analysis of the lariat intermediate

The reaction conditions for in vitro splicing were essentially as described (Tarn & Steitz, 1994, 1996b): 60% HeLa nuclear extract, 1 U/ μ L Ribonuclease inhibitor (Boehringer-Mannheim), 2 mM dithiothreitol, 20 mM creatine phosphate, 0.5 mM ATP, 2.5 mM $MgCl_2$, and 10 mM KCl. The concentrations of 2'-*O*-methyl oligonucleotides U2b (complementary to nt 27–49 of U2 snRNA) and U12a (complementary to nt 11–28 of U12 snRNA) were 4.8 μ M and 0.5 μ M, respectively. The nuclear extract was incubated at 30°C with the various 2'-*O*-methyl oligonucleotides for 10–15 min prior to substrate addition.

To confirm the identification of lariat intermediates, ribonuclease H digestions were carried out on splicing reactions after 5 or 6 h (the last time point) by adding 20 μ M of DNA oligonucleotide and 1 U (per 12.5 μ L splicing reaction) ribonuclease H (Boehringer-Mannheim), followed by an additional 20 min at 30°C. The 5' exon, intron, and 3' exon oligonucleotides, respectively, were GAATACGAGCAACG, CGATATTC GTCAGT, and TCTTAGTTGGGGCAAACA for XRP, AGCAC CAAGTTCAT, GGTACAAACTAGT, and TTGTATTTTGCA TCTGTGTT for XTF, and ACGCCATGGTGGAA, CGCACCC CGTATTC, and CAACTTACGTACATTTTGTA for SME. The reactions were subsequently treated with proteinase K, extracted with phenol-chloroform-isoamyl alcohol (PCA), and precipitated with ethanol. The products of splicing and of ribonuclease H digestion were separated on a 10% denaturing polyacrylamide gel.

Psoralen crosslinking

Splicing reactions were carried out as described above (except that the concentration of unlabeled splicing substrate was increased to 20 nM and approximately 1% polyvinyl alcohol (PVA) was added) and incubated at 30°C for 40 min. Samples were then transferred to ice, received 4'-amino-methyl-4, 5',8-trimethyl (AMT) psoralen to a final concentration of 20 μ g/mL, and were irradiated with 365 nm UV light

for 10 min as described previously (Sontheimer, 1994; Frilander & Steitz, 1999). After proteinase K treatment, phenol extraction, and ethanol precipitation, the crosslinked RNAs were resolved on a 5% denaturing polyacrylamide gel. The RNAs were subsequently transferred to Zeta-Probe (Biorad) membranes at 15 V in 0.5 $\times$ Tris-borate-EDTA (TBE) buffer overnight. The membranes were prehybridized at 42°C for 2 h in 25 mM sodium phosphate, pH 6.5, 6 $\times$ SSC, 5 $\times$ Denhardt's solution, 0.5% SDS, 50% formamide (with riboprobes), and 9 mg carrier RNA. They were hybridized at 42°C overnight with $\sim$ 5–10 $\times$ 10⁶ cpm U12-specific antisense riboprobes, produced by in vitro transcription as described previously (Wassarman & Steitz, 1992; Tabak, 1993; Tarn & Steitz, 1996b). DNA oligonucleotide probes were used to visualize U6_{atac}- and substrate-specific crosslinks. The membranes were then washed at room temperature for 15 min per step, first in 2 $\times$ SSC and 0.1% SDS, then in 0.5 $\times$ SSC and 0.1% SDS, and in 0.1 $\times$ SSC and 0.1% SDS, then finally at 55°C in 0.1 $\times$ SSC and 0.1% SDS. To reprobe membranes, they were first stripped by boiling in 0.1 $\times$ SSC and 0.5% SDS for 15 min.

Mapping the branchpoint by primer extension using reverse transcriptase

Splicing reactions were carried out as above using 2 to 10 fmol 32 P-labeled splicing substrate in a reaction volume of 200 μ L and incubating for 6 h. The products were separated on a 10% denaturing polyacrylamide gel and the lariat intermediate was excised from the gel, purified, and used as the template for primer extension (AMV RT, Roche) with the same oligonucleotides as used for 3' exon targeting in the ribonuclease H digestions. A 5 $\times$ 10⁵ cpm 5'-phosphorylated primer was prehybridized to each template by heating at 95°C for 1 min, then allowing the mixture to cool at room temperature for $\sim$ 20 min. Extension was carried out at 42°C for 45 min with 0.5 mM dNTPs. The dideoxy sequencing ladders were obtained by adding 0.1 mM ddNTP to four different reverse transcriptase reactions. The reactions were subsequently treated with ribonuclease A at 42°C for 5–10 min, with an equal volume of 10 M urea at 50°C for 20 min, and with proteinase K. After PCA extraction and ethanol precipitation, the extension products were separated on a 10% denaturing polyacrylamide gel and analyzed with a Molecular Dynamics PhosphorImager.

ACKNOWLEDGMENTS

We thank Therese Yario for technical assistance and the members of the Steitz Laboratory for helpful discussions. This work was supported by the National Institutes of Health (Grant GM26154 to J.A.S.), the European Molecular Biology Organization (long-term fellowship ALTF 616-1996 to M.J.F.), and the Academy of Finland (Grants 50116 and 50527 to M.J.F.). J.A.S. is an Investigator of the Howard Hughes Medical Institute.

Received January 15, 2002; returned for revision February 13, 2002; revised manuscript received February 27, 2002

REFERENCES

- Burge CB, Padgett RA, Sharp PA. 1998. Evolutionary fates and origins of U12-type introns. *Mol Cell* 2:773–785.
- Burge CB, Tuschl T, Sharp PA. 1999. Splicing precursors to mRNAs by the spliceosomes. In: Gesteland RF, Cech T, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 525–560.
- Dietrich RC, Incorvaia R, Padgett RA. 1997. Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. *Mol Cell* 1:151–160.
- Dietrich RC, Peris MJ, Seyboldt AS, Padgett RA. 2001a. Role of the 3' splice site in U12-dependent intron splicing. *Mol Cell Biol* 21:1942–1952.
- Dietrich RC, Shukla GC, Fuller JD, Padgett RA. 2001b. Alternative splicing of U12-dependent introns in vivo responds to purine-rich enhancers. *RNA* 7:1378–1388.
- Frilander MJ, Steitz JA. 1999. Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/branchpoint interactions. *Genes & Dev* 13:851–863.
- Frilander MJ, Steitz JA. 2001. Dynamic exchanges of RNA interactions leading to catalytic core formation in the U12-dependent spliceosome. *Mol Cell* 7:217–226.
- Hall SL, Padgett RA. 1994. Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. *J Mol Biol* 239:357–365.
- Hall SL, Padgett RA. 1996. Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science* 271:1716–1718.
- Hastings ML, Krainer AR. 2001. Functions of SR proteins in the U12-dependent AT-AC pre-mRNA splicing pathway. *RNA* 7:471–482.
- Jackson IJ. 1991. A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res* 19:3795–3798.
- Lavigne A, La Blanche H, Kornblihtt AR, Chabot B. 1993. A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes & Dev* 7:2405–2417.
- Levine A, Durbin RA. 2001. A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Res* 29:4006–4013.
- Lingner J, Keller W. 1993. 3'-end labeling of RNA with recombinant yeast poly(A) polymerase. *Nucleic Acids Res* 21:2917–2920.
- Liu Z, Luyten I, Bottomley MJ, Messias AC, Houngrinou-Molango S, Sprangers R, Zanier K, Kramer A, Sattler M. 2001. Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science* 294:1098–1102.
- Luo HR, Moreau GA, Levin N, Moore MJ. 1999. The human Prp8 protein is a component of both U2- and U12-dependent spliceosomes. *RNA* 5:893–908.
- MacMillan AM, Query CC, Allerson CR, Chen S, Verdine GL, Sharp PA. 1994. Dynamic association of proteins with the pre-mRNA branch region. *Genes & Dev* 8:3008–3020.
- Moore MJ. 2000. Intron recognition comes of AGE. *Nat Struct Biol* 7:14–16.
- Otake LR, Scamborova P, Hashimoto C, Steitz JA. 2002. *Drosophila* U6atac and U12 snRNAs are essential for development. *Mol Cell* 9:439–446.
- Query CC, Moore MJ, Sharp PA. 1994. Branch nucleophile selection in pre-mRNA splicing: Evidence for the bulged duplex model. *Genes & Dev* 8:587–597.
- Query CC, Strobel SA, Sharp PA. 1996. Three recognition events at the branch-site adenine. *EMBO J* 15:1392–1402.
- Ramchatesingh J, Zahler AM, Neugebauer KM, Roth MB, Cooper TA. 1995. A subset of SR proteins activates splicing of the cardiac troponin T alternative exon by direct interactions with an exonic enhancer. *Mol Cell Biol* 15:4898–4907.
- Reed R. 2000. Mechanisms of fidelity in pre-mRNA splicing. *Curr Opin Cell Biol* 12:340–345.
- Sharp PA, Burge CB. 1997. Classification of introns: U2-type or U12-type. *Cell* 91:875–879.
- Sonthheimer EJ. 1994. Site-specific RNA crosslinking with 4-thiouridine. *Mol Biol Rep* 20:35–44.
- Sun JS, Manley JL. 1995. A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes & Dev* 9:843–854.
- Sun Q, Mayeda A, Hampson RK, Krainer AR, Rottman FM. 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. *Genes & Dev* 7:2598–2608.
- Tabak HF. 1993. 1993 Nobel Prize for Medicine for the discovery of split genes. *Ned Tijdschr Geneesk* 137:2690–2693.
- Tarn WY, Steitz JA. 1994. SR proteins can compensate for the loss of U1 snRNP functions in vitro. *Genes & Dev* 8:2704–2717.
- Tarn WY, Steitz JA. 1996a. Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science* 273:1824–1832.
- Tarn WY, Steitz JA. 1996b. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell* 84:801–811.
- Tarn WY, Steitz JA. 1997. Pre-mRNA splicing: The discovery of a new spliceosome doubles the challenge. *Trends Biochem Sci* 22:132–137.
- Wassarman KM, Steitz JA. 1992. The low-abundance U11 and U12 small nuclear ribonucleoproteins (snRNPs) interact to form a two-snRNP complex. *Mol Cell Biol* 12:1276–1285.
- Watakabe A, Tanaka K, Shimura Y. 1993. The role of exon sequences in splice site selection. *Genes & Dev* 7:407–418.
- Will CL, Schneider C, MacMillan AM, Katopodis NF, Neubauer G, Wilm M, Luhmann R, Query CC. 2001. A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J* 20:4536–4546.
- Will CL, Schneider C, Reed R, Lührmann R. 1999. Identification of both shared and distinct proteins in the major and minor spliceosomes. *Science* 284:2003–2005.
- Wu Q, Krainer AR. 1996. U1-mediated exon definition interactions between AT-AC and GT-AG introns. *Science* 274:1005–1008.
- Wu Q, Krainer AR. 1998. Purine-rich enhancers function in the AT-AC pre-mRNA splicing pathway and do so independently of intact U1 snRNP. *RNA* 4:1664–1673.
- Wu Q, Krainer AR. 1999. AT-AC pre-mRNA splicing mechanisms and conservation of minor introns in voltage-gated ion channel genes. *Mol Cell Biol* 19:3225–3236.
- Zhang L, Doudna JA. 2002. Structural insights into group II intron catalysis and branch-site selection. *Science* 295:2084–2088.