Cooperative interaction of branch signals in the actin intron of Saccharomyces cerevisiae

Daniela Castanotto and John J. Rossi*

Molecular Biology Department, Beckman Research Institute of the City of Hope, 1450 East Duarte Road, Duarte, CA 91010, USA

Received June 16, 1998; Revised and Accepted July 24, 1998

ABSTRACT

In pre-mRNA splicing, specific spliceosomal components recognize key intron sequences, but the mechanisms by which splice sites are selected are not completely understood. In the Saccharomyces cerevisiae actin intron a silent branch point-like sequence (UAAUAAG) is located 7 nt upstream of the canonical sequence. Mutation of the canonical UACUAAC sequence to UAAUAAC reduces utilization of this signal and activates the cryptic UACUAAG. Splicing-dependent β-galactosidase assays have shown that these two splice signals cooperate to enhance splicing. Analyses of several variants of this double branch point intron demonstrate that the upstream UACUAAG sequence significantly increases usage of the UAAUAAC as a site of lariat formation. This activation is sequence-specific and unidirectional. However the ability of the UACUAAG signal to activate the downstream branch point is dependent on the presence of a short non-conserved sequence located a few nucleotides upstream of the UACUAAG. Mutation of this sequence leads to the disappearance of the cooperative interactions between the two branch signals. Our results show that this non-conserved sequence and the UACUAAG signal must both be present to achieve activation of the downstream branch point and suggest that a specific structure may be necessary to allow efficient recognition of the UAAUAAC.

INTRODUCTION

The strong conservation of the intron-encoded signals in Saccharomyces cerevisiae appears in part to account for intron recognition in this organism. In pre-mRNA splicing, specific spliceosomal components recognize key intron sequences. The yeast branch point sequence UACUAAC is highly conserved and is recognized at least twice within the course of splicing: the first is during commitment complex formation (1,2) and the second prior to step 1 of splicing, when it is base paired with U2 snRNA (3). During commitment complex formation, the U1 snRNP, the proteins Mud2P and BBP (branch bridging protein) as well as other uncharacterized proteins interact with the precursor mRNA prior to recruitment of U2 snRNP to the branch point (1,2,4,5). Interaction of BBP with the UACUAAC is believed to provide the earliest definition of this region (5). Point mutations within but not surrounding the branch point signal reduce the binding affinity of BBP in gel shift assays (5). The pairing of U2 snRNA with the branch sequence is similarly affected by mutations within the UACUAAC signal (3). Thus, in yeast, the fidelity of branch site selection is maintained by both protein binding and RNA–RNA base pairing interactions.

In the S. cerevisiae actin (ACT1) intron a silent branch point-like sequence (UAAUAAG) is located 7 nt upstream of the wild-type sequence and can be activated as a site of branch formation if the canonical UACUAAC sequence is mutated (6,7). We activated the cryptic UACUAAG by mutating the wild-type UACUAAC to UAAUAAC. In this construct branching takes place at both sequences, although not with equivalent efficiencies. When either signal is present alone within the ACT intron, branch formation and, concomitantly, splicing efficiency is reduced to <10% of wild-type. Unexpectedly, when the cryptic UACUAAG and mutant UAAUAAC are present together, these two splice signals cooperatively interact to enhance splicing to >50% of wild-type levels (6,7). In an effort to better understand the molecular basis for this more than additive increase in splicing efficiency, we initiated a molecular genetic analysis of the cryptic and mutant branch point signals and surrounding non-conserved sequences.

In this work, we show that in the absence of a canonical branch point signal, the cryptic and mutant sequence combination increases the overall efficiency of the splicing reaction and that this increase in splicing efficiency is due to a sequence-specific, unidirectional activation of the downstream sequence by the upstream UACUAAG. Furthermore, we have identified three small regions of complementarity located in the vicinity of the 5′, 3′ and branch point sequences. Mutational alteration of these regions indicates that in the UACUAAG/UAAUAAC mutant intron, these sequences are able to influence the first and second steps of the splicing reaction. To our knowledge, this represents the first case of branch point sequence utilization that can be altered both by cooperative interactions of two non-canonical sequences as well as by short, non-conserved intron sequences. These results suggest that although pre-mRNAs seem to lack

*To whom correspondence should be addressed. Tel: +1 626 359 8360; Fax: +1 626 301 8271; Email: jrossi@smtplink.coh.org
phylogenetically conserved structures, some structural signals are necessary for efficient completion of the splicing reaction.

**MATERIALS AND METHODS**

**Materials and strains**

Restriction endonucleases, T4 DNA ligase and T4 kinase were purchased from Bethesda Research Laboratories. Taq DNA polymerase was purchased from Perkin Elmer-Cetus. 32P-Labeled nucleotides were purchased from NEN. Avian myeloblastosis virus (AMV) reverse transcriptase was obtained from Life Sciences. T7 RNA polymerase and RNasin were purchased from Promega Biotech. The oligonucleotides used for PCR and primer extension analyses were synthesized in our laboratory using the ABI 391 PCR-MATE synthesizer.

Escherichia coli strain MC1061 (8) was used for routine manipulations. Escherichia coli strain CJ236 (MATa, leu2, ura3, trp1, his3) was kindly provided by J. Mc Ewen. Saccharomyces cerevisiae strain JM43 (MATa, his7, leu2, ura3-52) was kindly provided by J. Abelson. Escherichia coli strains were transformed by the method of Kushner (9) and his7 leu2 trp1 ura3 strain CJ236 (MATa, his7, leu2, ura3-52) was kindly provided by J. Abelson. Escherichia coli strains were transformed by the method of Kushner (9) and his7 leu2 trp1 ura3 strain CJ236 (MATa, his7, leu2, ura3-52) was kindly provided by J. Abelson.

**Plasmid constructs**

Plasmid pTZ18-ACT (11) was mutagenized as described by Kunzel (12) or by using mutagenic oligonucleotides in a PCR (13). The AC/AC mutation at the 3′ splice site of pTZ18-ACT was created using the BioRad mutagenesis kit to generate mismatched heteroduplex molecules. The resultant plasmid (pWTACT-AC) was then used in PCR reactions to create additional mutations. Specifically, a mutagenic oligonucleotide containing the ClaI restriction site at its 5′-end and one mismatch at the canonical branch point sequence was used to create the double mutant branch point sequence (UACUAAG-UAAUAAC) (p256LACZ). The two mutagenic oligonucleotides complementary to the T7 promoter sequence were used as primers in PCR reactions along with the specific mutagenic oligos containing the ClaI restriction site and various mismatches to create deletions, randomized sequences, inversions of the branch point sequences or mutations in the branch point sequence. The resulting PCR fragments were digested and cloned into pWTACT-AC using the EcoRI and ClaI restriction enzymes.

Plasmids pRCUpbp1, pCoopOut and pCoopRest, which have mutations upstream of the branch points, downstream of the 5′ splice site and in both locations, respectively, were also created in a similar manner. The 5′ primer used to create pRCUpbp1 (SCE-R1) was a mutagenic oligonucleotide complementary to the sequence of exon 1 with the exception of 1 nt which introduced a point mutation to create an EcoRI restriction site. For the construction of pCoopOut and pCoopRest, the 5′ primer (SCE-R1m1) included an additional portion that extended downstream of the 5′ splice site and contained the appropriate mismatches. The 3′ mutagenic oligonucleotide used to construct pRCUpbp1 and pCoopRest contained mismatches in the region upstream of the branch point sequences and included the ClaI restriction site. The template used in these PCRs was p256LACZ. The resulting fragments were digested and cloned into pWTACT-AC using the EcoRI and ClaI restriction enzymes.

Plasmids pRDWUp and pRUdUp were constructed using plasmids pCRDw and pCRUp as templates in mutagenic PCRs. SCE-R1 was the 5′ primer used for these constructions. The 3′ mutagenic oligonucleotide used for these PCRs contained mismatches in the region upstream of the branch point sequences and included the ClaI restriction site. The resulting PCR fragments were digested and cloned into pWTACT-AC using the EcoRI and ClaI restriction enzymes.

Plasmids p256#1, pCDelDw, pCRDw, pCRUp, pMVb, pHVRUp, pHVRDw, pBBPtest, pRCUpbp1, pCoopOut, pCoopRest, pRDWUp, pRUdUp and pBBPac test were digested with the EcoRI and HindIII restriction endonucleases. The resulting fragments were ligated into EcoRI/HindIII-cut Bluescript KS+, digested again with BamHI and SalI and cloned into the same sites of pBM258 (14), which harbors the GAL1-10 promoter used for inducible expression in S. cerevisiae. The vector derived by cloning the actin intron in pBM258 is called pB5ssbpACstr.

Plasmids pB5256ACstr, pBpBACstr, pB5ssACstr and pB5ssbpACstr were constructed using p256LACZ, pRCUpbp1, pCoopOut and pCoopRest, respectively, as templates in mutagenic PCRs. SCE-R1 was the 5′ primer used for these constructions. The 3′ primer mutagenized the 3′ splice site (AC/AC) and the region upstream of the 3′ splice site and created a SalI restriction site at the 3′-end of the fragments. These resulting PCR fragments were digested with the EcoRI and SalI restriction enzymes, ligated into EcoRI/SalI-cut Bluescript KS+, digested again with BamHI and SalI and cloned into the same sites of pBM258. These plasmids also had their 3′ splice sites restored and the mutagenized actin genes were cloned in-frame with the lacZ gene in the pYH7 expression vector (15) to perform in vivo β-galactosidase assays. The vector named ‘LACZ’ was added to the names of these plasmids to distinguish them from those having the AC mutation at the 3′ splice site.

Plasmids pWLTLACZ, p256LACZ, pCDelUpLACZ, pCDelDw-LACZ, pLVNpLACZ, pRCUpbLACZ and pBBPtestLACZ were constructed using pWT18-ACT, p256LACZ, pCDelUp, pCDelDw, pMVb, pRUdUp and pBBPtest as templates in mutagenic PCRs with a 5′ oligonucleotide complementary to exon 1 sequences and a mutagenic 3′ primer which restored the original 3′ splice site and introduced a SalI restriction site at the 3′-end of the fragments. The resulting PCR fragments were digested with the XhoI and SalI restriction enzymes and cloned into the pYH7 lacZ expression vector. The XhoI site is a unique restriction site located in the ACT intron between the 5′ splice site sequence and the branch point sequence.

Plasmid pBbCoopOutACT was created by cloning the actin promoter into pBbCoopOut using the BamHI and EcoRI restriction enzymes. Plasmids pBbCoopOutLACZ and pBbCoopRestLACZ were constructed by digesting pBbCoopOutACT with BamHI and XhoI and cloning the resulting fragment into the same sites of p256LACZ and pRCupbpLACZ, respectively. Plasmid pB256AGstr was constructed by PCR mutagenesis using pB256ACstr as template. The sequence of the 5′ primer is complementary to ACT exon 1 except for the point mutation.
which creates the EcoRI site. The 3′ primer restores the 3′ splice site and creates a SalI site at the 3′-end of the fragment. The PCR fragment was digested with the EcoRI and SalI restriction enzymes and cloned into pBCoopOutACT to yield pB256AGstr.

To construct pB5′ssAGstr, pBbpAGstr and pB5′ssbpAGstr, plasmids pCoopOutLACZ, pRCUbp1LACZ and pCoopRestLACZ were digested with the BamHI and ClaI restriction enzymes and the resulting fragments cloned into BamHI/ClaI-cut pB256AGstr. Plasmids pB256AGstr, pB5′ssAGstr, pBbpAGstr and pB5′ssbpAGstr were digested with the BamHI and SalI restriction endonucleases and cloned into the same sites of the pJYH7 vector to construct pB256AGstrLACZ, pB5′ssAGstrLACZ, pBbpAGstrLACZ and pB5′ssbpAGstrLACZ, respectively.

Mutations were confirmed by either dideoxy sequencing (16) of plasmid templates using the US Biochemical Sequenase kit or by using the City of Hope DNA sequencing core facility. The sequences of the oligonucleotides used for the plasmid constructions are available upon request.

RNA induction and primer extension analyses

Yeast transformants were grown in glucose minimal medium (17) to an A_{600} nm of 0.8. The cells were washed once in T_{20E} (10 mM Tris–HCl, pH 8.0, 1 mM EDTA) and the plasmid-derived RNA was induced by growing the cells in 2% galactose minimal medium for 3 h. For temperature-sensitive transformants, RNA induction was carried out in galactose minimal medium for 5 h at different temperatures. Preparation of the RNA and primer extension analyses were performed according to Donnely et al. (18). The oligonucleotide used in the primer extension analyses to verify branch point utilization for all the constructs with the AG/AC mutation at the 3′ splice site is complementary to 10 nt on either side of the 3′ splice site (22mer 303–305, 5′-GCAGCAACGTGTAAACATATAAA-3′). For the remaining constructs that have the pJYH7 vector background and utilize the ACT promoter, cells were grown in glucose minimal medium and the RNA analyzed by primer extension using a primer complementary to a region of the lacZ gene ~80 nt downstream of the 3′ splice site (β-gal 23mer). To verify the position of 5′ cleavage for all the constructs we used an intron-specific primer (IVS) which is complementary to the 22 nt region downstream of the 5′ splice site spanning the XhoI restriction site. An RNA-dependent product generated by a primer complementary to the 3′-end of U6 snRNA was used as an internal control to verify the amount and integrity of the RNA for each construct. For all primer extension reactions the annealing mixes were divided in half and simultaneously analyzed using either the U6 complementary primer or the actin constructs-specific primer. Reverse transcriptase was used to extend the RNA–oligonucleotide complex in the presence of dNTPs and ddNTPs at 45°C for 30 min. The products of the primer extension reactions were analyzed by gel electrophoresis in a 6% polyacrylamide–8 M urea sequencing gel.

β-Galactosidase assays

β-Galactosidase assays were done in triplicate and repeated a minimum of three times. Cells were grown in liquid minimal medium to an A_{600} nm of 0.8. Aliquots of 50 µl each culture were mixed with 950 µl Z buffer (60 mM Na_{2}HPO_{4}, 40 mM NaH_{2}PO_{4}, 10 mM KCl, 1 mM MgSO_{4}, 50 mM β-mercaptoethanol, pH 7.0) and assayed as described previously (19).

In vitro transcription

In vitro transcription was carried out in 10 µl reactions containing 50–100 µg/ml linearized (HindIII) DNA template mixed with 20 U T7 RNA polymerase in 40 mM Tris–HCl, pH 7.5, 20 mM NaCl, 6 mM MgCl_{2}, 10 mM DTT, 2 mM spermidine, 40 U RNasin, 0.5 mM each CTP, ATP and GTP, 25 µM UTP and 1 µCi/ml [α-32P]UTP. The reactions were incubated at 37°C for 2 h and the RNA transcripts were purified by gel electrophoresis.

RESULTS

A UACUAAG sequence increases lariat formation to a second cryptic branch site

We have previously constructed an intron with a branch point mutation in which the UACUAAC sequence of the S.cerevisiae ACT1 intron was converted to UAUAAAC. This point mutation in the context of the wild-type ACT intron resulted in only a modest (~50%) reduction in splicing efficiency. We reasoned that this change should have resulted in a more dramatic reduction in splicing based upon its effects on U2 base pairing interactions (3). This rather marginal effect on splicing could have been due to the presence of a nearby UACUAAG sequence, which can also serve as a branch point signal (6). To test this possibility, we deleted in sequential fashion either the UACUAAG or UAUAAAC sequences and examined the consequences of this on splicing efficiency. When the upstream sequence is deleted (pCDelUPLACZ), the downstream UAUAAAC allows only 9% of wild-type splicing efficiency (Table 1, d). When the downstream sequence is deleted (pCDelDWLACZ) the upstream sequence behaves similarly, generating 10% splicing efficiency (Table 1, e). However, when both non-canonical sequences are present (p256LACZ), the splicing efficiency is 54% of wild-type (pWTLACZ) and three times higher than the expected 19% of a purely additive interaction (Table 1, c).

In order to analyze the site and relative efficiencies of branch formation, we constructed plasmid pD256#1 (Fig. 1), in which in addition to the cryptic mutant branch point signals contains an AG/AG→AC/AC mutation at the 3′ splice site designed to block the second step of the splicing reaction. This mutation does not alter the first step of splicing and results in the accumulation of intron lariats still covalently attached to exon 2 (7,20). Other relevant constructs were similarly modified to allow analysis of step 1 of splicing by primer extension. Results obtained using total RNA isolated from S.cerevisiae transcripts show that when both the UACUAAG and UAUAAAC sites are present, branching occurs at both signals, with an approximate 40:60 ratio of upstream to downstream signal (Fig. 2, lane 1). In the absence of the upstream UACUAAG sequence (pCDelUp), branch formation at the UAUAAAC site is barely detectable (Fig. 2, compare lanes 1 and 2). This is consistent with the splicing assay results (Table 1, d). Quite surprisingly, branch formation at the UACUAAG site is not affected by the absence of the UAUAAAC sequence (pCDelDW; Fig. 2, lanes 1 and 3). This result is not consistent with the β-galactosidase splicing assays, in which this mutant only generated 10% of wild-type splicing efficiency (Table 1, e), and thus explanations other than efficiency of branching must be sought to explain the reduced level of splicing in the light of the relatively efficient branch formation.

Since deletions alter spacing between critical sequences, we randomized the UACUAAG or the UAUAAAC sequence and repeated the primer extension analyses. The results obtained with
Table 1. *In vivo* splicing efficiencies of introns containing mutant branch point sequences as determined by β-galactosidase measurements

<table>
<thead>
<tr>
<th>Construct</th>
<th>Splicing Efficiency</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) WT-LACZ</td>
<td>5'ss_________UACUAAG_UACUAAC_______AG</td>
<td>100%</td>
</tr>
<tr>
<td>b) pDWTLACZ</td>
<td>5'ss_________XXXXXXX_UACUAAC_______AG</td>
<td>100%</td>
</tr>
<tr>
<td>c) p256LACZ</td>
<td>5'ss_________UACUAAG_UAAUAAC_______AG</td>
<td>54%</td>
</tr>
<tr>
<td>d) pCDeU/pLACZ</td>
<td>5'ss_________XXXXXXX_UAAUAAC_______AG</td>
<td>9%</td>
</tr>
<tr>
<td>e) pCDeDwLACZ</td>
<td>5'ss_________UACUAAG__XXXXXXX_______AG</td>
<td>10%</td>
</tr>
<tr>
<td>f) pNvbpLACZ</td>
<td>5'ss_________UAAUAAC_UACUAAG_______AG</td>
<td>27%</td>
</tr>
<tr>
<td>g) pRCupbp1LACZ</td>
<td>5'ss______AUGAU_ UACUAAG_UAAUAAC_______AG</td>
<td>20%</td>
</tr>
<tr>
<td>h) pBBPAGtstLACZ</td>
<td>5'ss_________UACAAAG_UAAUAAC_______AG</td>
<td>19%</td>
</tr>
</tbody>
</table>

To monitor splicing efficiency exon 1, IVS and part of exon 2 of the *ACT* gene were cloned in the pJYH7 vector under the control of the actin promoter (see text), to obtain a spliced product in-frame with the β-galactosidase gene. The AUG translational start codon for lacZ gene expression is present in exon 1 and is in the correct frame for translation only after splicing of the actin mRNA. Thus the level of splicing directly correlates with the level of β-galactosidase activity. X indicates nucleotide deletion and the randomized nucleotides upstream of the branch point are shown. The effects of these mutations on splicing are reported with respect to the wild-type *ACT* intron (WT-LACZ), which produced 1747 β-galactosidase units. These assays were done in duplicate and repeated a minimum of three times using independent yeast transformants. The SD column indicates the standard deviations calculated for each construct.

Figure 1. Construction of a double cryptic branch point intron. Plasmid pBTz 258 containing the *S.cerevisiae ACT-HIS4* gene fusion is shown. Distances from the native and cryptic branch points to the 3′ splice site are given in nucleotides. The nucleotide changes required to create plasmid p256#1 are shown above the wild-type sequence. The mutated nucleotides are underlined. p256#1 contains two mutations; one converts the branch point sequence UACUAAC to a cryptic UAAUAAC signal and the other changes the AG/AG of the 3′ splice site to AC/AC. This plasmid was used as a substrate to create other constructs containing various mutations (Materials and Methods). The 3′ splice site was restored in all of the plasmids and the resulting genes were cloned in-frame with the lacZ gene under the control of the actin promoter for β-galactosidase determination (Materials and Methods).

the randomized sequences (pCRUp and pCRDw; Fig. 2, lanes 4 and 5) were the same as those obtained with the deletions, demonstrating that spacing alterations were not responsible for the observed effects. Presence of the UACUAAG element is sufficient to activate the UAAUAAC as a branch point, whereas branching at the UACUAAG is not affected by the presence or
Determination of \textit{in vivo} branch point selection by primer extension analyses. An oligonucleotide complementary to 22 nt at the intron/exon 2 junction of the \textit{S.cerevisiae} \textit{ACT} gene was used in primer extension analyses of total RNA prepared from \textit{S.cerevisiae} transformed with pD256#1 and its derivatives. The RT stops at the upstream and downstream branch points are indicated by open and filled arrowheads, respectively. An RNA-dependent product generated by a primer complementary to the 3' end of U6 snRNA (shown above) was used as an internal control to verify the amount and integrity of the RNA for each construct. Below the autoradiographic exposure are the DNA sequences of the constructs analyzed in the primer extension reactions. The lower case letters in parentheses are presented for comparison with Table 1. From lane 1 to 10 the plasmids analyzed are p256#1, pCDelUp, pCDelDw, pCRUp, pCRDw, pINVbp, pINVRU, pINVRDw, pRCUpbp1 and the wild-type \textit{S.cerevisiae} construct (pWTACT-AC). The additional RT stop, indicated with a small square, corresponds to a cytosine nucleotide located in a non-conserved region 7 nt upstream of the UACUAAG sequence. In lane 4, no RT stops corresponding to the expected sizes for the downstream branch point or the non-conserved 5 nt sequence were observed. The products of the primer extension reactions were electrophoresed in a 6% polyacrylamide–8 M urea sequencing gel.

absence of the mutant branch signal. Reversal of the positions of the two elements revealed that the UAAUAAC signal must be located downstream of the UACUAAG element to be activated (Fig. 2, compare lanes 1 and 6). Based on these and on previous results we can conclude that activation of the UAAUAAC sequence by the UACUAAG element is sequence-specific and unidirectional. Despite the fact that branching efficiency within the UACUAAG sequence does not appear to be dependent upon the presence or absence of the UAAUAAC signal, splicing as measured by \(\beta\)-galactosidase activity does not proceed efficiently to step 2 when the branch is in the UACUAAG signal (Table 1).

A small non-conserved sequence affects the ability of the UACUAAG signal to activate the downstream UAAUAAC

Primer extension analyses of the intronic RNAs revealed a reverse transcriptase (RT) stop located 7 nt upstream of the UACUAAG sequence (Fig. 2). The possibility that this could be an additional site of lariat formation was examined by treating isolated RNAs with debranching enzyme. Primer extension analyses of total RNA from pD256#1 transformants either treated or not treated with debranching enzyme were carried out on RNA templates (Fig. 3, lanes 1 and 2). The results obtained show that the RT stop is still present after treatment with debranching enzyme, while the authentic branches were removed. However, the RNA sequence mapped the RT stop to a cytosine located 7 nt upstream of the UACUAAG sequence (Fig. 3) and GC branch points have been shown to be resistant to the activity of debranching enzyme (21). To test if this RT stop was a GC branch point we transformed a temperature-sensitive yeast mutant defective for the first step of splicing, prp5, with our construct pD256#1 and grew the transformants at increasing temperatures (from 23 up to 42°C). It has been shown that PRP5 mediates a U2 snRNA conformational change which is necessary for U2 snRNA binding to the branch sequence (22). Consequently, the propagation of this mutant at increasing temperatures determines a progressive loss of branch point formation. Primer extension analyses of the RNA (Fig. 4) showed that branch formation at the two branch points decreased with increasing temperature and finally disappeared at the non-permissive temperature (37°C). However, the RT stop upstream of the branch point sequences still persisted at the non-permissive temperature (Fig. 4), demonstrating the absence of a branch at that position. In addition, when similar primer extension analyses were carried out using \textit{ACT}intron RNA prepared from T7 \textit{in vitro} transscripts, the same strong stop was observed (Fig. 3). These combined results confirm that this stop is due to a structural feature of the RNA and not to a covalent modification.

The cytosine corresponding to the RT stop is part of a 5 nt sequence (5'-UGCUA) which can potentially base pair with another 5 nt sequence (5'-UAGCG) beginning 8 nt downstream of the 5' cleavage site. To examine whether or not this strong RT stop was due to intramolecular RNA interactions, we randomly mutated the sequence upstream of the UACUAAG element (5'-UGCUA) to 5'-AUGAU. This mutational alteration eliminated the strong RT stop (Fig. 2, lane 9). Functional analyses of this construct (pRCUphpb1LACZ) demonstrated that mutagenizing this region also affected the splicing reaction: overall splicing
efficiency was drastically reduced (Table 2, g). Perhaps most interestingly, this alteration negated the cooperativity between the two crippled branch point sequences (Table 1, compare g and c). The 5 nt sequence upstream of the UACUAAG sequence was reduced to the level observed in the absence of the UACUAAG (pRCUppbp1; Fig. 2, compare lane 9 with lanes 1–3). To test possible intramolecular pairings within this pre-mRNA transcript, we made a series of mutations and compensatory changes and analyzed their effects on the splicing efficiency of the double cryptic branch point construct. Compensatory mutations designed to restore the potential pairing of this 5 nt sequence with the region 8 nt downstream of the 5′ splice site did not restore splicing efficiency (Table 2, m). In addition, no changes in branch point selection were detected with this construct (data not shown). These results argue against the existence of an intramolecular pairing between the short sequence upstream of the branch point signal (5′-UGCUA) and the sequence downstream of the 5′ splice site (5′-UAGCG). We also examined the possibility that the 5 nt mutation upstream of the branch signals was affecting usage of single branch signals as opposed to having an effect on the cooperative interaction of these elements. The mutation was individually combined with either the UACUAAG (pCRDw) or UAAUAAC (pCRUp) constructs. The relative efficiencies of branch formation in these constructs was not reduced by the presence of the 5 nt randomization (Fig. 5, left panel, lanes 5 and 6), confirming that this mutation has its primary effect on cooperativity (Fig. 5, compare lanes 2 and 7 in the left panel or lanes 2 and 3 in the right panel).

Small non-conserved sequences affect the overall splicing efficiency of the cryptic mutant branch point construct

The 5 nt sequence upstream of the UACUAAG element is duplicated 13 nt upstream of the 3′ splice site. Thus, this 3′ splice site proximal sequence also has base pairing potential with the sequence beginning 8 nt downstream of the 5′ cleavage site. To test the existence of this base pairing interaction, each sequence was individually altered (pCoopOutLACZ and pB256 AG-strLACZ). These variants were tested for their potential effects on splicing, either independently or in combination with each other (pB5′ssAGstrLACZ). The results are summarized in Table 2. Each of the mutations affected overall splicing efficiency. The mutation upstream of the 3′ splice site (pB256AGstrLACZ) produced the largest reduction in spliced product, lowering the efficiency of the reaction to 27% of the parental construct (Table 2, l). In contrast, mutations downstream of the 5′ splice site produced the smallest effect on splicing (Table 2, i). Compensatory mutations that restore the potential pairing between these two sequences brought the splicing efficiency to near its original level (Table 2, o).

Compensatory mutations made in combination with the mutation upstream of the branch point signals did not restore splicing to the parental level (pB5′ssbpAGstrLACZ; Table 2, p). We have no explanation as to why the splicing reaction was somewhat more efficient when mutations upstream of the branch point sequence and downstream of the 3′ splice site were combined (pBbpAGstrLACZ; Table 2, n).
Table 2. Effects of mutations in non-conserved intron sequences upon \textit{in vivo} splicing efficiencies

<table>
<thead>
<tr>
<th>Construct</th>
<th>Spliced product sequence</th>
<th>In vivo splicing efficiency</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>c) p256LCZ</td>
<td>5'ss_UAGCG_UACUAAG_UAUAAC_UACUAAG</td>
<td>100%</td>
<td>+/- 2.80%</td>
</tr>
<tr>
<td>g) pRCUpppLCZ</td>
<td>5'ss_UGACU_UACUAAG_UAUAAC_UACUAAG</td>
<td>37%</td>
<td>+/- 2.82%</td>
</tr>
<tr>
<td>h) pBBAGtestLCZ</td>
<td>5'ss_UAGCG_UACUAAG_UAUAAC_UACUAAG</td>
<td>36%</td>
<td>+/- 0.94%</td>
</tr>
<tr>
<td>i) pCoopOutLCZ</td>
<td>5'ss_UACAU_UACUAAG_UAUAAC_UACUAAG</td>
<td>63%</td>
<td>+/- 2.44%</td>
</tr>
<tr>
<td>l) pB256AGstrLCZ</td>
<td>5'ss_UACAU_UACUAAG_UAUAAC_UACUAAG</td>
<td>27%</td>
<td>+/- 3.37%</td>
</tr>
<tr>
<td>m) pCoopRestLCZ</td>
<td>5'ss_UACAU_UACUAAG_UAUAAC_UACUAAG</td>
<td>54%</td>
<td>+/- 4.19%</td>
</tr>
<tr>
<td>n) pBbpAGstrLCZ</td>
<td>5'ss_UACAU_UACUAAG_UAUAAC_UACUAAG</td>
<td>72%</td>
<td>+/- 3.77%</td>
</tr>
<tr>
<td>o) pB5'ssAGstrLCZ</td>
<td>5'ss_UACAU_UACUAAG_UAUAAC_UACUAAG</td>
<td>80%</td>
<td>+/- 4.02%</td>
</tr>
<tr>
<td>p) pB5'ssShpAGstrLCZ</td>
<td>5'ss_UACAU_UACUAAG_UAUAAC_UACUAAG</td>
<td>42%</td>
<td>+/- 1.69%</td>
</tr>
</tbody>
</table>

Spliced products from these constructs are translationally in-frame with the \textit{lacZ} gene. As shown in Table 1, the level of splicing directly correlates with the level of \(\beta\)-galactosidase activity. Nucleotide mutations in the three non-conserved regions of the \textit{ACT} intron are depicted in small bold letters. The wild-type sequences of these regions are presented in a smaller font (c). The effects on splicing are reported with respect to the parental double cryptic branch point intron, which produced 1006 \(\beta\)-galactosidase units. These assays were done in duplicate and repeated a minimum of three times using independent yeast transformants. The SD column indicates the standard deviations calculated for each construct.

To analyze the first step of the splicing reaction for these constructs, we mutated the 3' splice site AG to AC (pCoopOut, pCoopRest, pB256ACstr, pBbpACstr, pB5'ssACstr and pB5'ssShpACstr) and performed primer extension analyses. The results showed no additional changes in branch point selection (data not shown). The fidelity of the 5' splice site cleavages for all the constructs analyzed in this study was also verified by primer extension analyses using an intron-specific primer (IVS). No changes in the 5' cleavage site were detected for any of the constructs (data not shown).

**DISCUSSION**

**In vivo** cooperativity between cryptic splice sites increases the efficiency of the splicing reaction

In this work we have investigated the role of non-conserved sequences in pre-mRNA splicing. For our studies, we have chosen the \textit{S.cerevisiae} actin intron, which has the peculiar feature of a branch point-like sequence (UACUAAG) located immediately upstream of the canonical UACUAAC. The cryptic branch point is only utilized if the canonical signal is weakened by a point mutation (7). We and others (7) have shown that the splicing efficiency of this double cryptic branch point intron is more than the additive contribution of each of these signals. In this work we have demonstrated that this higher splicing efficiency is attained through activation of the downstream branch point by the upstream UACUAAG sequence. This phenomenon is unidirectional and sequence-specific. Interestingly, branch formation at the UACUAAG sequence in the absence of the downstream UACUAAC is more efficient than the corresponding branch formation at the UACUAAG sequence in the absence of the upstream UACUAAG. However, when each of these sequences is present alone, the splicing efficiencies are almost identical (Table 1). Since branching at the upstream UACUAAG is more efficient than branching at the downstream signal and since the mRNA levels are comparable when either signal is used, we conclude that the efficiency of the second step of splicing is reduced when branching occurs at the upstream position in the intron, presumably due to an unfavorable position relative to the 3' splice site. Results obtained using a construct where the UACUAAG sequence had been mutated to a strong UACUAAC signal gave qualitatively similar results. From these results and the comparison of the double UACUAAC intron both sequences were utilized for branch formation (the downstream signal was preferentially used), but the overall splicing efficiency was 80% of wild-type (data not shown). This decrease in splicing efficiency would be expected if branching to the upstream UACUAAC signal resulted in a less efficient second step of splicing.

The marked preference for selection of the UACUAAG sequence versus the UACUAAC as a site of lariat formation (e.g. compare Fig. 2, lanes 4 and 7 and lanes 5 and 8) cannot be explained by more favorable pairing with U2 snRNA, even when surrounding nucleotides are taken into account. Thus, it is possible that some other splicing factor(s) preferentially recognizes the UACUAAG signal. This specific recognition of the UACUAAG sequence could explain why the synergism between the two branch sites is sequence-specific and unidirectional and why the
Likewise, deletions at the corresponding sites of the RP51a intron 4144′3′5′a few nucleotides upstream of these two signals. This sequence, UACUAAG sequence is lost by mutating a short region located upstream of the branch point (the wild-type sequence UGCUA is mutated to AUGAU). The R indicates the 5 nt mutation upstream of the splice site. Interestingly, sequence downstream of the 5′ splice site did not change much, since compensatory mutations designed to restore the potential pairing between these two sequences did not have any effect on splicing nor did they restore cooperative interaction between the two branch sites.

Activation of the downstream branch point is also dependent on the presence of the UACUAAG signal and this effect is sequence-specific and unidirectional. These results demonstrate that the UACUAAG signal and the short sequence upstream of it are both necessary to achieve activation of the downstream branch point. Primer extension analyses of the parental intronic RNA shows a RT stop that can be mapped to the cytosine located within the short sequence (Fig. 3). Remarkably, this strong RT stop, which as expected is eliminated by randomizing the short sequence, is also absent in constructs where the UACUAAG signal has been removed or its position relative to the UAAUAAC has been altered (Fig. 2, lanes 2, 4 and 6–8). In all of these constructs, cooperative interaction between the two branch signals is annulled. These results show that the 5′-UGCUA sequence and the UAAUAAC signal may be components of a structure that is directly or indirectly (through binding of trans-acting factors) responsible for efficient selection of the downstream branch point. Recently, a new protein that possesses sequence specificity for the branch point sequence has been isolated (4,5). Data from these studies indicate that this protein (BBP) contributes to recognition and selection of the branch point prior to U2 snRNA binding (5). It has also been shown that BBP can bind to a UACUAAG sequence (5). Mutation of the UACUAAG sequence to a UACA sequence, located upstream of the UACUAAG, eliminates cooperative interaction between the two branch signals (Fig. 2, lane 9 and Table 1, compare c with g). Activation of the downstream branch point is also dependent on the presence of the UACUAAG and this effect is sequence-specific and unidirectional. These results demonstrate that the UACUAAG signal and the short sequence upstream of it are both necessary to achieve activation of the downstream branch point. Primer extension analyses of the parental intronic RNA shows a RT stop that can be mapped to the cytosine located within the short sequence (Fig. 3).

Mutations of the UAAUAAC signal are activated by a short non-conserved sequence and a cryptic branch point

Figure 5. Effects of randomization of 5 nt on selection of single branch points analyzed by primer extension. Primer extension analyses were performed as described in Figure 2. The product yielded by extension of the U6 snRNA (shown above) was used as an internal standard for these assays. The RT stops at the upstream and downstream branch points are indicated by Bp1 and Bp2 followed by arrows. The additional stop upstream of the UACUAAG sequence is indicated by a dot (•). In the left panel the plasmids analyzed are: lane 1, pWTACT (5′-ss_UA CUAAG_UACUAAC__AC); lane 2, p256#1 (5′-ss___UA CUAAG_UACUAAC__AC); lane 3, pCRDw (5′-ss__UACUAAG random__AC); lane 4, pCRUp (5′-ss__random_UACUAAC__AC); lane 5, pRDwUp (5′-ss_R_UACUAAG random__AC); lane 6, pRUpUp (5′-ss__R_random_UACUAAC__AC); lane 7, pCRUpUp1 (5′-ss__R_UACUAAG_UAAUAAC__AC). The R indicates the 3 nt mutation upstream of the branch point (the wild-type sequence UGCUA is mutated to ACGAU). The randomized sequence replacing the branch point sequence is UCGAUGU. In the right panel the plasmids analyzed are: lane 1, pWTACT (5′-ss___UA CUAAG_UACUAAC__AC); lane 2, p256#1 (5′-ss___UA CUAAG_UACUAAC__AC); lane 3, pCRUpUp1 (5′-ss__R_UA CUAAG_UAAUAAC__AC); lane 4, pBBPest (5′-ss___UA CAAAG_UAAUAAC__AC). The UACUAAG sequence on its own is preferable to UAAUAAC as a site for lariat formation.

Non-conserved sequences within the intron play important roles during splicing

Activation of the downstream branch point in the presence of the UACUAAG sequence is lost by mutating a short region located a few nucleotides upstream of these two signals. This sequence, 5′-UGCUA, and an analogous sequence located upstream of the 3′ splice site can both potentially base pair with a complementary sequence located downstream of the 5′ splice site. Interestingly, these sequences also have base pairing potential with nucleotides A52G6G61C61AG2 of U6 snRNA. Mutation of the 3′ splice site proximal sequence does not affect branch point selection (data not shown), yet causes a drastic reduction in the overall splicing efficiency (Table 2, 1). In an independent study, Kivens and Siliciano (23) have shown that mutations at the same sites (nucleotides C and G in the sequence located upstream of the 3′ splice site) in the actin intron (D-17 and D-18) decrease splicing activity in both mutant (5′-A, 3′-C) and wild-type introns. Likewise, deletions at the corresponding sites of the RP51a intron (D-19 and D-20) also reduce splicing efficiency (23). In the present study, compensatory mutations in a complementary sequence located downstream of the 5′ splice site did not change branch selection, but were able to restore splicing to almost the parental levels (Table 2, o). Interestingly, Libri et al. (24), using an in vivo selection system, have found that intramolecular RNA–RNA interactions can occur between the ends of introns and function as enhancers of splicing efficiency.

In contrast, the sequences upstream of the branch point that affect branch point selection do not seem to interact with the sequence downstream of the 5′ splice site, since compensatory mutations designed to restore the potential pairing between these two sequences did not have any effect on splicing nor did they restore cooperative interaction between the two branch sites.

The UAAUAAC signal is activated by a short non-conserved sequence and a cryptic branch point

In the present study, compensatory mutations in a complementary sequence located downstream of the 5′ splice site did not change branch selection, but were able to restore splicing to almost the parental levels (Table 2, o). Interestingly, Libri et al. (24), using an in vivo selection system, have found that intramolecular RNA–RNA interactions can occur between the ends of introns and function as enhancers of splicing efficiency.

In contrast, the sequences upstream of the branch point that affect branch point selection do not seem to interact with the sequence downstream of the 5′ splice site, since compensatory mutations designed to restore the potential pairing between these two sequences did not have any effect on splicing nor did they restore cooperative interaction between the two branch sites.

The UAAUAAC signal is activated by a short non-conserved sequence and a cryptic branch point

In the present study, compensatory mutations in a complementary sequence located downstream of the 5′ splice site did not change branch selection, but were able to restore splicing to almost the parental levels (Table 2, o). Interestingly, Libri et al. (24), using an in vivo selection system, have found that intramolecular RNA–RNA interactions can occur between the ends of introns and function as enhancers of splicing efficiency.

In contrast, the sequences upstream of the branch point that affect branch point selection do not seem to interact with the sequence downstream of the 5′ splice site, since compensatory mutations designed to restore the potential pairing between these two sequences did not have any effect on splicing nor did they restore cooperative interaction between the two branch sites.
point signal of an essential gene such as actin, cooperative interactions between cryptic sequences could increase mRNA production to a level sufficient for cell survival.

ACKNOWLEDGEMENTS

We thank M. Green for providing the debranching enzyme, members of the Rossi laboratory for helpful discussions, R. J. Lin for suggestions and critical reading of the manuscript and Wanda Fitzgerald for editorial assistance. This work was supported by National Institutes of Health grant AI 29329.

REFERENCES