RNA (2000), *6*:352–368. Cambridge University Press. Printed in the USA.
Copyright © 2000 RNA Society.

Splicing enhancement in the yeast rp51b intron

D. LIBRI,[1] A. LESCURE,[2] and M. ROSBASH[3]

[1]Centre National de la Recherche Scientifique, Centre de Génétique Moléculaire, Gif-sur-Yvette, France
[2]Unité Propre de Recherche 9002 Structure des Macromolécules Biologiques et Mécanismes de Réaction, 67084 Strasbourg Cedex, France
[3]Howard Hughes Medical Institute, Brandeis University, Waltham, Massachusetts 02254, USA

ABSTRACT

Splicing enhancement in higher eukaryotes has been linked to SR proteins, to U1 snRNP, and to communication between splice sites across introns or exons mediated by protein–protein interactions. It has been previously shown that, in yeast, communication mediated by RNA–RNA interactions between the two ends of introns is a basis for splicing enhancement. We designed experiments of randomization-selection to isolate splicing enhancers that would work independently from RNA secondary structures. Surprisingly, one of the two families of sequences selected was essentially composed of 5′ splice site variants. We show that this sequence enhances splicing independently of secondary structure, is exportable to heterologous contexts, and works in multiple copies with additive effects. The data argue in favor of an early role for splicing enhancement, possibly coincident with commitment complex formation. Genetic compensation experiments with U1 snRNA mutants suggest that U1 snRNP binding to noncanonical locations is required for splicing enhancement.

Keywords: enhancer; randomization-selection; splicing; U1 snRNP

INTRODUCTION

Splicing is the process that leads to the production of coding mRNAs from split genes through the precise and regulated removal of introns (for reviews, see Moore et al., 1993; Hodges & Bernstein, 1994; Madhani & Guthrie, 1994). The maturation of precursor RNA to yield one or more mRNAs is a multistep, nuclear process that occurs within the spliceosome, a large ribonucleoprotein particle. Spliceosome assembly precedes the first catalytic reaction (cleavage of the 5′ exon-intron junction) and occurs by multiple steps, some of which can be rate limiting for intron removal.

Early intron recognition depends on the interaction of ribonucleoprotein particles (U snRNPs) and proteins with splicing signals located at the 5′ end (5′ splice site or donor site) or near the 3′ end (branch site, pyrimidine stretch, and 3′ AG) of the intervening sequence. The 5′ splice site is first recognized by U1 snRNP, mainly via base pairing with a short region of the U1 snRNA. The branchpoint is also recognized very early by a protein complex containing at least two factors: Mud2p and branchpoint binding protein (BBP) (Abovich & Rosbash, 1997; Berglund et al., 1997). Because it is inessential, Mud2p is believed to play an auxiliary role in the interaction between BBP and the branchpoint (Berglund et al., 1997). This role is paralleled in higher eukaryotes by the role played by U2AF (a protein factor that recognizes the branchpoint region) in the stabilization of the interaction between the SF1 factor (Arning et al., 1996) (ortholog of BBP) and the less well-defined branch site (Berglund et al., 1998). Mud2p might play additional and more direct roles in early intron recognition (Rain & Legrain, 1997) and/or intron bridging in light of its genetic (Abovich et al., 1994) and physical interactions with the U1 snRNP particle (Fromont-Racine et al., 1997). The interaction of Mud2p with the U1 snRNP-associated protein Prp39p as well as the simultaneous interaction of BBP with the 3′ end of the intron and the U1 snRNP protein Prp40p (Kao & Siliciano, 1996) define a network of protein interactions that bridges the two ends of yeast introns in very early steps of spliceosome assembly (Abovich & Rosbash, 1997; Berglund et al., 1997).

The interaction of U2 snRNP and additional protein factors with the branchpoint is the first ATP-requiring step. This is followed by the addition of the triple snRNP U4:U6:U5, which is most likely responsible for the displacement of U1 snRNP (Konforti et al., 1993; Crispino & Sharp, 1995; Konforti & Konarska, 1995). In fact, during spliceosome assembly, the 5′ splice site sequence is recognized at least two times: first by U1

Reprint requests to: D. Libri, Centre National de la Recherche Scientifique, C.G.M., Gif-sur-Yvette, France; e-mail: Libri@cgm.cnrs-gif.fr.

snRNP and subsequently by U6 snRNP. In the latter case there are Watson–Crick interactions between a highly conserved ACA on U6 and the last three nucleotides (UGU) of the yeast donor site (Kandels-Lewis & Séraphin, 1993; Lesser & Guthrie, 1993b). The first transesterification reaction follows a conformational change, which is believed to activate the spliceosome by enabling the U2:U6 complex to carry out its catalytic functions.

Both in yeast and in higher eukaryotes, splicing is not an all-or-none process. A multiplicity of RNA conformations and RNA-binding activities are likely to determine not only splice-site recognition, but also a wide range in the efficiencies with which introns are excised.

Splicing enhancers have been described in higher eukaryotes and are believed to act at early steps of spliceosome assembly. Enhancers are defined as exonic or intronic sequences that increase splicing efficiency, can promote exon inclusion, and are generally functional in heterologous contexts (Somasekhar & Mertz, 1985; Reed & Maniatis, 1986; Cooper & Ordahl, 1989; Streuli & Saito, 1989; Balvay et al., 1992; Black, 1992; Cooper, 1992; Libri et al., 1992; Huh & Hynes, 1993, 1994; Lavigueur et al., 1993; Tian & Maniatis, 1993, 1994; Watakabe et al., 1993; Caputi et al., 1994; Dirksen et al., 1994; Tanaka et al., 1994; Humphrey et al., 1995; Sirand-Pugnet et al., 1995; Staffa & Cochrane, 1995; Tian & Kole, 1995; Carlo et al., 1996). These sequences have been shown to bind to protein factors, most of which are characterized by the presence of SR (Ser-Arg) repeats (for review, see Fu, 1995; Manley & Tacke, 1996; Valcárcel & Green, 1996). It has been shown that splicing enhancers can promote *trans*-splicing in HeLa cells extracts when present on the downstream RNA molecule (Bruzik & Maniatis, 1995; Chiara & Reed, 1995).

The presence of a donor site downstream of an exon is considered a bona fide enhancer for the splicing of an upstream intron. In fact, it has been shown that inactivation of such a donor site most often leads to exon skipping (Robberson et al., 1990; Talerico & Berget, 1990; Berget, 1995). It is now widely accepted that the presence of a functional donor site favors spliceosome assembly independently of its involvement in a splicing reaction (Grabowski et al., 1991; Kuo et al., 1991; Lou et al., 1995; Hwang & Cohen, 1996a). Strikingly, it has been shown that a 5′ splice site can functionally substitute for a purine-rich splicing enhancer as a requirement for a *trans*-splicing reaction in HeLa cells extracts (Chiara & Reed, 1995). The enhancement effect is believed to be due to the indirect recruitment to the upstream intron of splicing factors and/or SR proteins by the U1 snRNP bound to the downstream donor site (Watakabe et al., 1993; Staknis & Reed, 1994).

Comparable splicing enhancers have not yet been described in yeast. This is perhaps not surprising, considering the absence of known bona fide yeast SR proteins. In previous reports (Newman, 1987; Goguel & Rosbash, 1993; Libri et al., 1995; Charpentier & Rosbash, 1996), it has been shown that the interaction of relatively short and imperfect inverted repeats in yeast introns leads to an RNA secondary structure-mediated form of splicing enhancement. In the absence of such interaction, or when it is significantly weakened by mutations, splicing as well as commitment complex formation are reduced (Charpentier & Rosbash, 1996). This indicates that secondary structure formation affects the very early steps of intron recognition.

In this report we identify a novel form of yeast splicing enhancement, the mechanism of which resembles more closely that existing in higher eukaryotes. We provide evidence suggesting that U1 snRNP is the *trans*-acting mediator of this effect, which likely concerns very early assembly steps.

RESULTS

Identification of two families of sequences that improve splicing efficiency in yeast

We have previously identified sequences that improve splicing efficiency by allowing the formation of intramolecular RNA–RNA interactions between the two ends of the intron (Libri et al., 1995). With the present work we aimed at isolating sequences that would work in yeast introns as more traditional, structure-independent, splicing enhancers. We performed experiments of randomization-selection in vivo using a selection system based on an intron-containing *CUP1* gene (Fig. 1), so that splicing efficiency would be reflected by the levels of copper resistance conferred to yeast. In addition to allowing for a fast selection of better splicers, the reporter provides a convenient and extremely sensitive tool to evaluate splicing efficiency (Lesser & Guthrie, 1993a; Stutz & Rosbash, 1994; Libri et al., 1995).

To favor the isolation of sequences that would specifically improve early stages of intron recognition, we performed most of the selection experiments in a *MUD2* knock-out strain. This should introduce a defect in early splicing steps, because Mud2p is involved in early intron recognition (Rain & Legrain, 1997) and commitment complex formation (Abovich et al., 1994; Abovich & Rosbash, 1997). To confirm this expectation, we assayed the effect of Mud2p deletion on splicing of a battery of introns (Fig. 1C). In many cases, the deletion produces a decrease of about 50% in copper resistance and splicing efficiency of introns that do not have obvious defects in early recognition steps (Acc, rp51a, b/bI/b.st, bi/bi; see Materials and Methods and Fig. 1). A stronger and sometimes dramatic effect is noticed for introns with mutations in the donor site (Don2C), the branchpoint (BpA2C, BpA6C), or the secondary structure (3mUB1 and 5mUB1). These results underscore

the synergistic nature of early splice-site recognition, secondary structure formation, and Mud2p activity.

As a substrate for our randomization experiments, we used a variant of the rp51b intron (SB4/DB4; Libri et al., 1995) that we introduced in the CUP1 based reporter. In this intron a short sequence in the proximity of the 5′ splice site (SB4, Selected Box 4) interacts with a sequence (DB4) located immediately downstream of

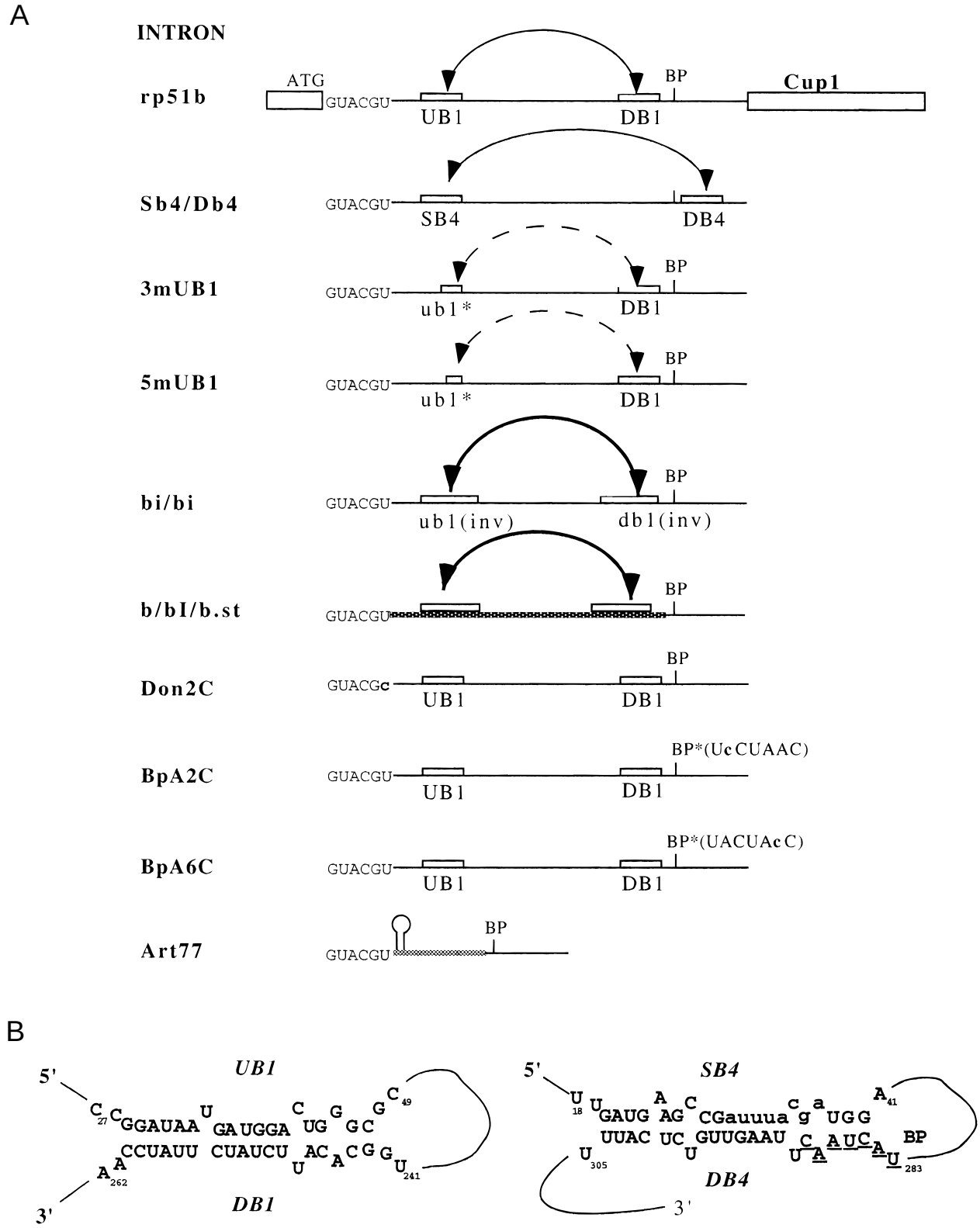


FIGURE 1. (*Figure continues on facing page.*)

the branchpoint (Fig. 1). Splicing of this intron is severely affected by the absence of Mud2p, possibly because the SB4/DBA interaction is not allowed to form or is not productive without Mud2p. Accordingly, we reasoned that selection from a pool of introns containing random sequences in the same location as DB4 (immediately downstream of the branchpoint) would be unlikely to yield sequences capable of base pairing with the 5′ half of the intron in the absence of Mud2p.

We constructed a pool of intron sequences [SB4/X_8, previously called SB4/db4N (Libri et al., 1995)] containing 8 random nt situated 2 nt downstream of the branch site (same location as DB4; Fig. 2). This pool of introns was cloned into the *CUP1*-based reporter gene and introduced into a mud2Δ yeast strain. Because the introduction of random sequences in DB4 (SB4/X_8) prevents the formation of the SB4-DB4 stem (and the absence of Mud2p greatly enhances the effects of secondary structure defects; see Fig. 1C), it was not surprising that growth of the pool on copper was almost undetectable (see below, Fig. 3B). Individual clones were then selected based on copper resistance between 0.3 and 0.6 mM $CuSO_4$.

The intron sequences obtained from this experiment are shown in Figure 2B. We used the CONSENSUS package of programs (Hertz et al., 1990; Hertz & Stormo, 1999) to identify conserved patterns in the selected sequences. Programs WCONSENSUS and PATSER (see Materials and Methods) allowed the identification of a family of sequences containing the consensus pattern UGUAU/CGU, which we will refer to hereafter as the Ψ-donor family. A second family was defined based on the extreme U-richness of its components. Finally, two sequences could not be obviously assigned to either family (#30 and #31, shown as "others" in Fig. 2).

Figure 2 also shows sequences obtained from selections performed in the presence of Mud2p (Fig. 2C). "Winners" could be easily assigned to the same two families (see Materials and Methods), although the percentage of U-rich sequences is higher in these selec-

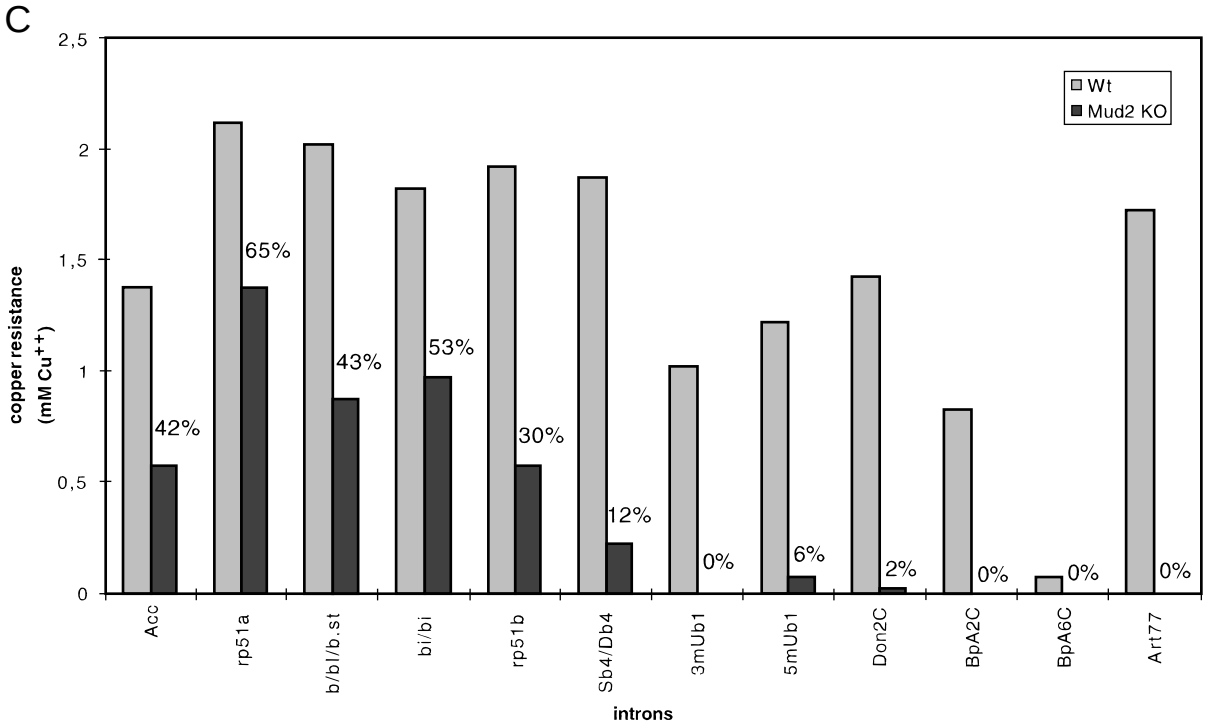


FIGURE 1. A: Scheme of the various introns introduced in the *CUP1* reporter gene. The interaction of UB1 with DB1 enhances splicing efficiency in the wild-type intron. The sequence SB4 was obtained after randomization-selection of the wild-type sequence UB1. The interaction of SB4 with a different region of the intron (DB4) allows efficient splicing of the SB4/DB4 version. Note that this last interaction involves the region immediately downstream of the branch site. The RNA–RNA interactions are shown by double headed arrows; thicker lines in bi/bi and b/bI/b.st indicate the presence of a perfect stem; dashed lines in 3mUb1 and 5mUb1 indicate a weaker interaction due to mutations in the Ub1 region. Mutations are indicated in lowercase letters. Introns b/bI/b.st (319 nt) and Art77 (99 nt) contain a sequence different from rp51b (315 nt) (shown by a different line pattern). Art77 contains a putative binding site for the hU1A protein (AUUGCACUCC) closed by two GC pairs and an imperfect stem afterwards (shown by a stem). **B**: A schematic drawing of the UB1/DB1 and SB4/DB4 interactions. **C**: Effect of *MUD2* deletion on splicing of various introns. The copper resistance (which is linked to splicing efficiency) is shown in the presence or absence of Mud2p after background subtraction. The effect of the Mud2p absence is expressed as the percentage of the splicing efficiency in a wild-type, Mud2p proficient, background. The background level (0.075 mM Cu^{2+}) is defined as the copper concentration that is tolerated by the cup1Δ strain in the absence of any reporter.

A

B

SB4/db4, mud2Δ

Ψ-donor

#	sequence
1,2	UUGUGUAUuc
3	AUGUACGC
4,9	AUGUAUGU
5	GUGUACGU
8	GUGUACGN
10	AUGUACGA
12,16	uUGUA.GUUA
13	uUGUUUGUA
14	uuGUAUAUAG
15	AUGUACUG
17	AUGUAAGU
18	GUGUAUGU
19	UUGUUUGUC

U-rich

#	sequence
6	UUUUUGUU
7	UUUUUGUA
11	UGAGUUUA

others

#	sequence
30	AUAUUGAG
31	GUUGCAAA

C

SB4/db4, MUD2

W1	UAUGUAUGu
W6	UUUGUACGu
W10	GUGUACGC
W11	GUGUACGU
W4	AUGUAUGA
W17,24	uUUUAUGU

W5	UUAUUUUC
W7	AUUUGUUC
W8	GUUUUUUU
W12	UUUGUUUC
W13	UUUUUUUA
W14	UUUUUGUU
W15	GUUUUGUU
W16	UUGUUUUC
W18	UUAGUUUA
W20	AUUUUUAA
W21	UUUUGUUC
W22	UUUGUUAA

W9	GUGUUGUC
W2	AABUGUUC
W19	UUGUUAUA
W3	UGUGAUGA

D

Nucleotides frequency

Pos. / nt	-2	-1	1	2	3	4	5	6
A	0.37	0	0	0	0.89	0.06	0.05	0.11
C	0	0	0	0	0	0.44	0	0.16
G	0.31	0	0.95	0	0	0	0.84	0.06
U	0.31	1	0.05	1	0.11	0.5	0.11	0.66
Inf.	0.21	1.61	1.03	1.61	1.79	0.80	0.62	0.41

Ψ-donor consensus

A	**U**	**G**	**U**	**A**	**U**	**G**	**U**
G					**C**		
U							

5' splice site consensus

A	**G**	**G**	**U**	**A**	**U**	**G**	**U**

FIGURE 2. (*Legend on facing page.*)

tion conditions. Four sequences could not be obviously assigned to either family ("others" in Fig. 2C).

The pattern that emerges from the alignment of the Ψ-donor family is strikingly similar (Fig. 2D) to the intronic portion of a 5′ splice site (GUAUGU). In some cases (#4,9, #18, #W1) the sequence is identical. Most of the variations are compatible with the consensus sequence for yeast donor sites (see, for instance sequence #5, GUACGU, which is identical to the donor site of the rp51b intron) and occur at position 4 (a C for the consensus U) and position 6. A major difference is observed, however, in the nucleotide that precedes the sequence. This is almost always a U, although a G is the preferred nucleotide (G_{-1}) in the case of a bona fide donor site.

The donor site-like sequence is a secondary structure-independent splicing enhancer

We decided to investigate further the properties of the donor site-like putative splicing enhancer, hereafter referred to as Ψ-donor enhancer. We inserted a "model" selected sequence (AUGUAUGU) into a set of different introns, in a location identical to that of the selected sequences. If added to the wild-type rp51b intron, the enhancer significantly increases splicing efficiency in the absence of Mud2p, as measured by copper resistance and primer extension (Fig. 3 and data not shown). In its presence, however, we did not detect any difference in splicing efficiency, possibly because the enhancer affects a step that is not limiting in this case.

Consistent with this hypothesis, coupling the enhancer to a mutation in the donor site of the rp51b intron (Don2C) resulted in a positive effect, both in the presence and absence of Mud2p. (Note that the 5′ splice site of the rp51b intron contains a nonconsensus C at position 4; the Don2C 5′ splice site contains two nonconsensus C at positions 4 and 6, GUAcGc.) This is particularly conspicuous in the absence of Mud2p, as splicing of the same intron without the enhancer is almost undetectable under these conditions. Similarly, versions of the rp51b intron bearing mutations that disrupt the required secondary structure (3mUB1 and $SB4/X_8$) are positively affected by the introduction of the Ψ-donor enhancer, again both in the presence and absence of Mud2p. (Note that we compared SB4/Ψ-donor to $SB4/X_8$ that contains random sequences in DB4, as in both cases the SB4-DB4 pairing is disrupted.) Finally, introns BpA2C and BpA6C (Fig. 1) contain A-to-C transversions in positions 2 and 6 of the branchpoint sequence respectively. These mutations have been shown to affect BBP binding and early intron recognition (Berglund et al., 1997; Rain & Legrain, 1997) but are also expected to affect later splicing steps (e.g., U2snRNP binding) that might be rate limiting (Fouser & Friesen, 1986; Parker et al., 1987; Rain & Legrain, 1997). In both cases, in the presence as well as in the absence of Mud2p, the splicing enhancer did not improve splicing efficiency, consistent with a specific early role for the enhancer.

In the case of the Ψ-donor enhancer, it was necessary to insure that splicing enhancement was not due to base pairing of this sequence with other intron sequences (Libri et al., 1995) leading to reconstitution of an RNA–RNA enhancer. To this end, we inserted the UGUAUGU sequence downstream of the branchpoint of an artificial intron, which verifies that the Ψ-donor enhancer is also functional when exported to a sequence environment with completely unrelated pairing potential. Intron Art77 contains only the last 46 nt of the rp51b intron (from the branchpoint to the 3′ end) and a different 5′ sequence. This intron bears an early step splicing defect, which is mainly revealed in the absence of Mud2p (Fig. 1C; see Materials and Methods). As expected, splicing in the absence of Mud2p was detectable for Art77 only in the presence of the enhancer (Fig. 3B). As with the wild-type rp51b intron, we suggest that the lack of detectable splicing enhancement because of the Ψ-donor sequence in the presence of Mud2p (Fig. 3A) is due to the step affected by the enhancer not being rate limiting.

FIGURE 2. Randomization-selection experiment on SB4/DB4 intron. Eight nucleotides have been randomized starting at position +2 after the branch site (**A**). The sequences issued from the experiment performed in a mud2Δ yeast strain and with the SB4/DB4 intron are shown in **B**. A similar experiment performed in the presence of Mud2p with the SB4/DB4 intron gave the sequences shown in **C**. The sequences of the Ψ-donor family (shown in **A** and **B**) have been aligned and analyzed with the CONSENSUS package of programs (Hertz et al., 1990; Hertz & Stormo, 1999) (see Materials and Methods). Flanking sequences have been indicated in lowercase whenever they contribute to the consensus pattern. Sequence #12,16 was manually aligned by introducing a gap (shown by a dot). **D**: Frequency matrix associated with the alignment of all the sequences of the Ψ-donor family. Positions are defined by analogy with the bona fide 5′ splice site (i.e., position +1 is the first intron nucleotide). The last row indicates the positional information content (Schneider et al., 1986; Stormo & Hartzell, 1989) of the consensus (in bits) that is a measure of the constraints imposed to the nucleotide choice at a given position and varies between 0 and 2. A small sample correction was applied (average information background expected from a random alignment of $n = 19$ sequences) and the starting pool bias was included in the calculation. The standard deviation expected from a random alignment is 0.10 bits. "B" indicates G, U, or C; "N" any nucleotide. The sequence used for the retesting experiments is AUGUAUGU (#4,9), which was assigned the highest score (together with sequence #W1) by the PATSER program.

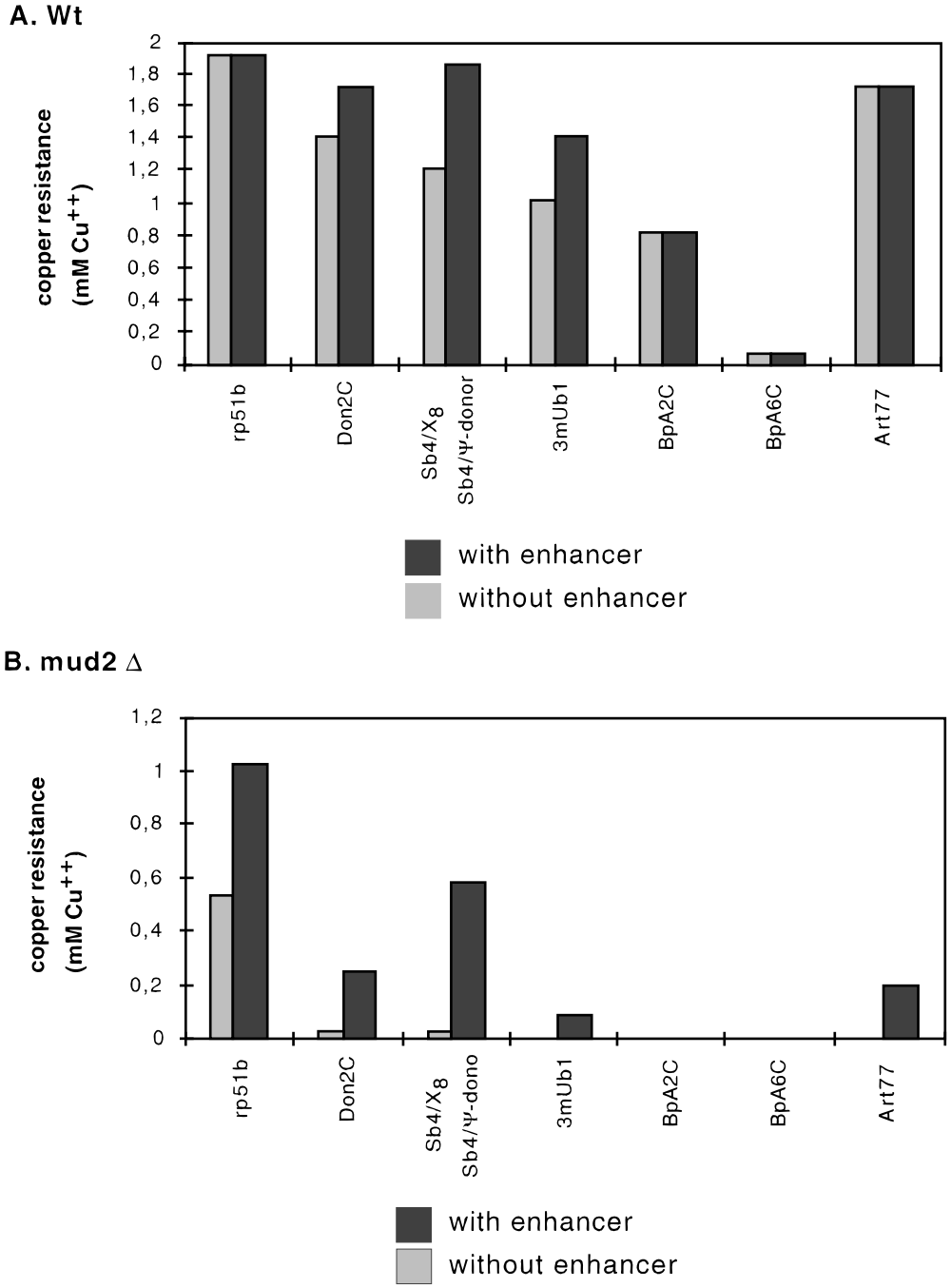
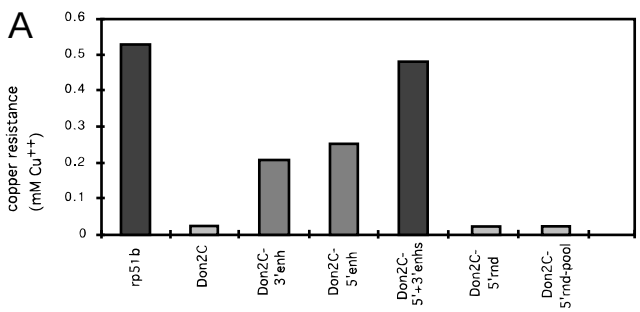


FIGURE 3. Splicing enhancement induced by the selected sequence in different introns. Note that in the case of intron SB4/DB4, the enhancer-containing version was compared to the starting pool (SB4/X_8), as in both cases the SB4-DB4 interaction has been disrupted. **A**: Wild-type background. **B**: mud2Δ background.

The Ψ-donor enhancer can be exported to the 5′ end of the intron

We have previously shown (Libri et al., 1995) that at very early steps of intron recognition the rp51b pre-mRNA is folded with a structural arrangement in which the branchpoint and the donor site are in close proximity. An obvious prediction is that a Ψ-donor enhancer should also be functional in the vicinity of the 5′ splice site.

We have artificially introduced a perfect copy of the Ψ-donor enhancer (UGUAUGU) at intron position +18. (Note that use of the Ψ-donor as a 5′ splice site would yield an out-of-frame mRNA.) We have coupled this to the Don2C mutation, and examined the splicing efficiency of the resulting intron, in the absence (Don2C-5′enh) and in the presence (Don2C-5′+3′enhs) of an additional copy of the enhancer in its original position (downstream of the branchpoint). As can be seen in Figure 4, one copy of the enhancer in its 5′ position



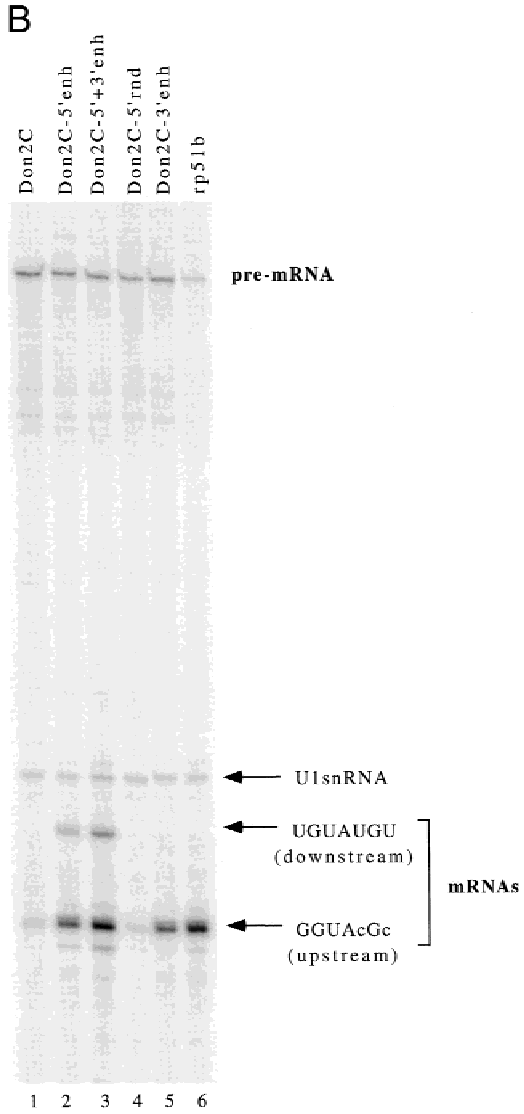


FIGURE 4. **A**: Introduction of one perfect copy of the Ψ-donor enhancer in the 5′ region of intron rp51b, version Don2C, either alone (Don2C-5′ enh) or with an additional copy in the 3′ location (Don2C-5′+3′ enhs). Copper resistance in mud2Δ background, compared to the wild-type rp51b and Don2C-3′ enh. As a control, Don2C-5′rnd-pool is a pool of sequences containing a random region in the same position as the 5′ enhancer in Don2C-5′ enh and Don2C-5′+3′ enhs. Don2C-5′rnd is one of the sequences of this pool. **B**: Primer extension to monitor the use of the Ψ-donor enhancer as a bona fide 5′ splice site in Don2C-5′enh (lane 2) and Don2C-5′+3′enhs (lane 3). Lane 1: Don2C. Lane 4: Don2C-5′rnd containing an unrelated sequence in the same position as the 5′ enhancer. Lane 5: Don2C-3′enh, containing a single copy of the enhancer sequence in the 3′ location. Lane 6: rp51b wt intron. The primer annealing site is in the *CUP1* gene. As a loading control, the endogenous U1 snRNA has been reverse transcribed. Use of the enhancer as a splice site (downstream site) gives rise to a slower migrating mRNA band that codes for a frame-shifted protein. Upstream mRNA/pre-mRNA ratios are as follows: Don2C: 0.17 ± 0.039 (n = 5); Don2C-5′ enh: 1.61 ± 0.25 (n = 5); Don2C-3′enh: 0.86 ± 0.12 (n = 3); Don2C-5′+3′ enhs: 3.95 ± 0.81 (n = 5); rp51b: 3.82 ± 0.73 (n = 3).

(Don2C-5′enh) produced a clear splicing enhancement effect in a mud2Δ background. This was even slightly better than the effect of the same sequence when located in a 3′ position (Don2C-3′enh). More strikingly, two copies of the enhancer further boosted splicing efficiency and nearly suppressed the U_6C splicing defect present in Don2C (i.e., Don2C-5′+3′enhs is spliced almost as efficiently as rp51b). As a control, Don2C-5′rnd-pool and Don2C-5′rnd, which contain, respectively, a randomized segment and a randomly chosen sequence in the same location as the enhancer in the 5′ location, are spliced as the parental Don2C clone.

Primer extension of the RNAs derived from these constructs (Fig. 4B) shows that both the bona fide 5′ splice site and the Ψ-donor sequence are used as donor sites, although the latter is used with an approximately four times lower efficiency. Importantly, the Ψ-donor sequence increased splicing at the original donor site, consistent with the Cu^{2+}-resistance measurements and indicating enhancer activity. An identical experiment performed in the presence of Mud2p gave a similar result, although the increase in splicing efficiency induced by the enhancer is smaller (data not shown). Besides showing that the enhancer can be exported to a different location in the intron, this experiment shows that the addition of a second copy has an additive effect on splicing.

Mutational analysis of the Ψ-donor enhancer

The impressive similarity of a bona fide donor site with the Ψ-donor enhancer prompted us to investigate the possible involvement of the U1 snRNP particle as a mediator of this splicing enhancement. We first undertook a mutational analysis of the enhancer sequence. A mutated enhancer was introduced in the wild-type rp51b intron in the 3′ location, and the resulting en-

hancement effects were examined in the absence of the Mud 2 splicing factor (Fig. 5).

The most striking difference between a bona fide donor site and the Ψ-donor enhancer is the absence of the G that would be the last nucleotide of the exon (G/GUAUGU for a donor site versus UGUAUGU for the enhancer). This G (G_{-1}) is involved in the base pairing interaction of U1 RNA 5′ end with the 5′ splice site (Séraphin & Kandels-Lewis, 1993; our unpubl. results) and is moderately conserved in yeast introns (Spingola et al., 1999). The U in position 4 of the donor site, on the other hand, is conserved in the selected sequences and in naturally occurring donor sites (Rymond & Rosbash, 1992; Long et al., 1998) in spite of not being involved in a Watson–Crick base pairing (the corresponding nucleotide on the U1 side of the helix is a pseudo-uridine, Ψ). Finally, positions 2 and 6 (U in both cases) are examples of a well-conserved (U_2) and a relatively nonconserved position (U_6).

When position U_4 was mutated to an A (enhUG UAaGU) in the context of the rp51b intron, the enhancer effect was severely reduced. However, mutation of the same nucleotide to a C (enhUGUAcGU) had little effect. Interestingly, a C is frequently present in this position both in the selected sequences (Fig. 2) and in naturally occurring donor sites (Rymond & Rosbash, 1992; Long et al., 1998). Similarly, mutation of U_6 to a C (enhUGUAUGc) or to a G (data not shown) had only a modest effect, again consistent both with the selection data and 5′ splice site conservation in yeast. As expected, mutation of U_2 to a C (enhUGcAUGU) considerably reduced the splicing-enhancer effect whereas mutation of the same position to a G (enhUGgAUGU) had even a negative effect on splicing (i.e., the intron containing this sequence was spliced less efficiently than the wild-type, enhancerless intron). A similar negative effect was observed with mutation of U_{-1} to a G (enhgGUAUGU), which was more pronounced when the same mutation was coupled to the U_4-to-A change (enhgGUAaGU).

One possible explanation for this negative effect is that the presence of two Gs in the immediate proximity of the

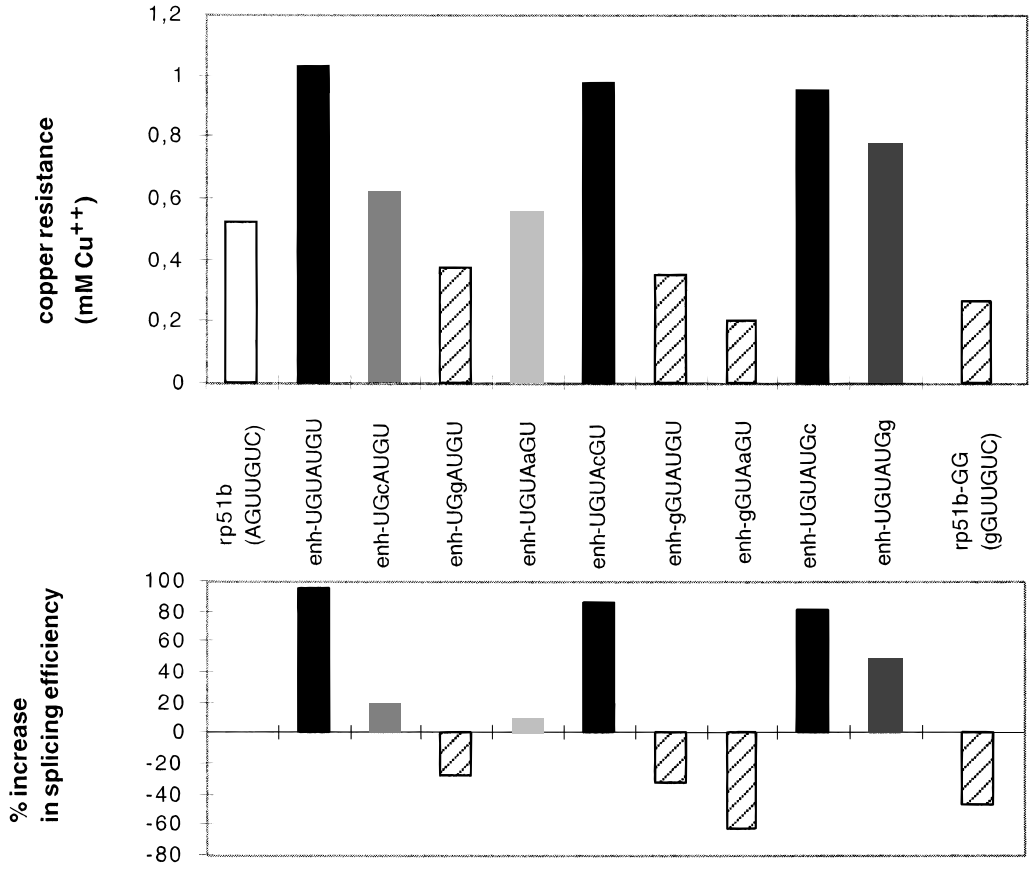


FIGURE 5. Mutational analysis of the Ψ-donor enhancer sequence. Upper panel: copper resistances associated with introns containing mutated versions of the enhancers. Bottom panel: percentage increase in splicing efficiency conferred by the various enhancers to the wild-type rp51b intron. The A-to-G mutation in rp51b-GG is located at the same position with respect to the branchpoint as the U-to-G mutation in enh-gGUAUGU. Percentage increase is $(C_{enh} - C)/C$ where C is the copper resistance of the enhancerless version and C_{enh} is the copper resistance of the enhancer-containing intron. Note that this value is negative if the sequence actually decreases splicing efficiency of the wild-type intron (hatched histograms). Experiments performed in a mud2Δ background.

branchpoint (sequences enhUGgAUGU, enhgGUAU GU, and enhgGUAaGU) is somewhat deleterious for splicing. We decided to test this hypothesis directly. A single A-to-G mutation in the wild-type, enhancerless intron allowed the creation of a G doublet at the same position with respect to the branchpoint as in enhgGU AUGU and enhgGUAaGU (rp51b-GG, Fig. 5). This intron was spliced less efficiently than the wild-type intron (Fig. 6), consistent with a negative role for a GG dinucleotide immediately downstream of the branchpoint.

These results confirm a general agreement between the conservation of the 5′ splice site in yeast introns and the data from the selection shown in Figure 2.

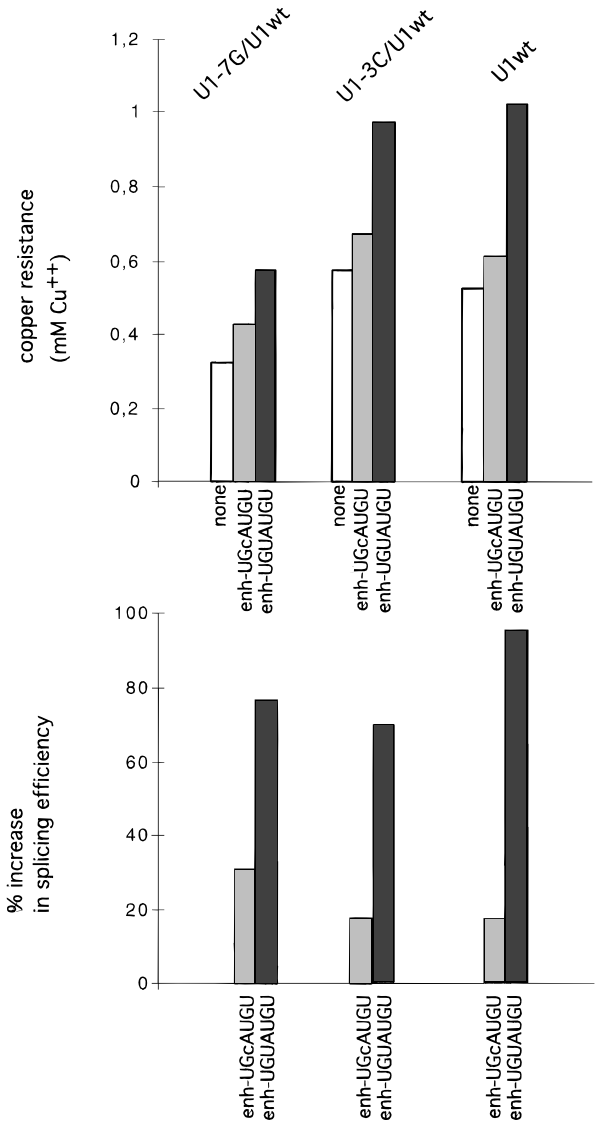


FIGURE 6. Effect of U1-7G snRNA expression (merodiploid strain for U1 snRNA) on splicing improvement conferred by the mutated sequence UGcAUGU. Upper panel: copper resistances for constructs containing the wild-type rp51b intron and enhancer sequences as indicated in the presence of additional and mutated forms of U1 snRNP (U1-7G/U1wt and U1-3C/U1wt) or without the ectopic U1 snRNA expression (U1wt) in a mud2Δ strain. The lower panel shows the corresponding enhancer effects expressed as percentage increase in splicing efficiency as in Figure 5.

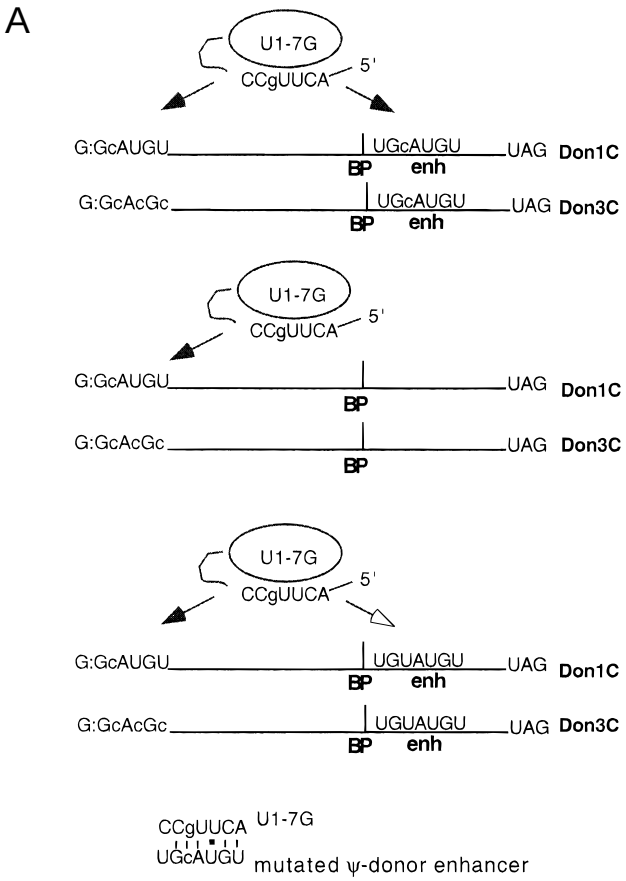
Genetic evidence of the implication of U1 snRNP in splicing enhancement

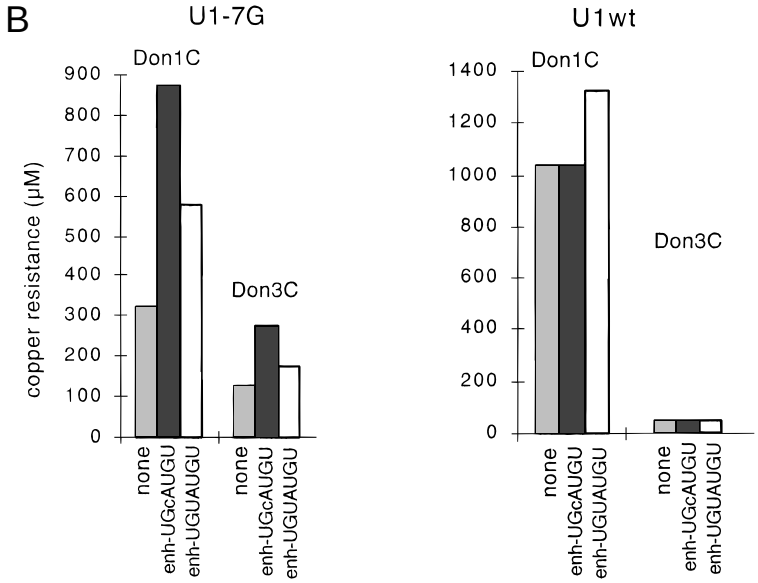
We devised a genetic complementation assay to confirm the involvement of U1 snRNP in the splicing enhancement process. Mutation U_2C in the Ψ-donor sequence almost completely abolishes splicing enhancement (Fig. 5; 16% of the nonmutated value). If U1 snRNP is the *trans*-acting mediator of this effect, expression of a U1 RNA containing an A-to-G transition at position 7 (U1-7G) should at least partially restore the enhancement effect.

Because yeast containing U1-7G as the only source of U1 RNA has a slow growth phenotype (our unpubl. results), and this mutation is likely to be lethal in a mud2Δ background, we first expressed U1-7G in the presence of a wild-type copy of the U1 snRNA gene and assessed the enhancement effect conferred to the rp51b intron by enhUGcAUGU versus enhUGUAUGU in a mud2 null background. As a control, we also coexpressed another mutated version of U1 snRNA (U1-3C), which has no obvious phenotype in a haploid strain (D. Libri, unpubl.). As shown in Figure 6, we observed a modest restoration of the splicing enhancement activity of enhUGcAUGU upon expression of the complementary U1-7G snRNA but not of U1-3C. In the presence of U1-7G, the enhancer activity of enhUGc AUGU is 30% of the enhancer activity of the nonmutated Ψ-donor sequence (the same sequence retains only 16% of the enhancer activity in the absence of U1-7G).

One possible explanation for the modest increase is that the expressed U1-7G snRNP is not as efficiently assembled as the wild-type particle.

One alternative is that the same U1 snRNP particle recognizes sequentially both the enhancer and the 5′ splice site. It is then possible that efficient enhancement would only occur when the two sequences are efficiently recognized by U1-7G. If true, more efficient splicing enhancement should be observed for an intron in which both the enhancer and the 5′ splice site contain the same U-to-C mutation. We then constructed a version of intron Don2C containing a U-to-C transition at position 2 (Don3C, sequence of the 5′ splice site G:GcAcGc) and compared its splicing efficiency in the absence or presence of the wild-type (enhUGUAUGU) or the mutated version (enhUGc AUGU) of the enhancer (Fig. 7A). Because splicing would presumably be severely affected by this triple mutation, we also constructed versions of the same introns containing only the first U-to-C transition (Don1C, sequence of the 5′ splice site G:GcAUGU, Fig. 7A). We introduced these constructions into a MUD2 strain containing U1-7G as the only source of U1 snRNA and measured splicing efficiency by copper resistance (Fig. 7B) and primer extension (Fig. 7C). As shown in Figure 7B,C, splicing was barely detect-





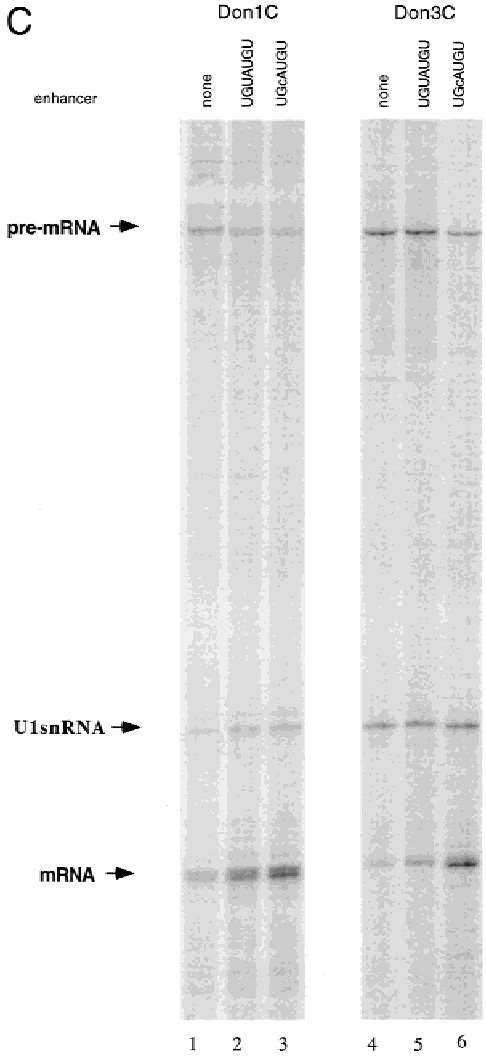


FIGURE 7. Genetic complementation of the UGcAUGU enhancer mutation by U1-7G snRNA. Experiments performed in a strain containing only the U1-7G version of U1 snRNA. **A**: The schematic drawing shows the structure of the introns used (Don3C and Don1C) with the sequences of the 5′ splice sites and the splicing enhancers (when present). The expected interactions of U1-7G snRNP are indicated by arrows; an empty arrow indicates a weaker interaction expected when U1-7G binds to the wild-type enhancer (presence of a G-U wobble pair in the RNA helix instead of a G-C). The RNA–RNA interaction between U1-7G snRNA and the enhUGcAUGU is also shown. **B**: Histogram showing the copper resistances characteristic of the various intron-containing reporters. The sequences of the enhancers as well as the introns concerned are indicated. "None" indicates the enhancerless version of the intron. The experiment was performed in a strain containing either the mutated U1-7G snRNA or U1 wild-type on a centromeric plasmid. Note that the copper growth assay for the U1-7G and the U1wt strain cannot be reliably compared because the former has a doubling time that is about two times higher than the former. **C**: Primer extension of RNAs derived from the introns shown at the top of the panels. Reverse transcription of U1 snRNA was used as a loading control. The mRNA/pre-mRNA ratios are as follows: Don1C: 0.74 ± 0.28 (n = 3); Don1C-Ψ-donor enhancer (UGUAUGU): 1.73 ± 0.73 (n = 3); Don1C-mutated enhancer (UGcAUGU): 4.26 ± 0.35 (n = 3). Don3C-Ψ-donor enhancer: 0.34 ± 0.12 (n = 2); Don3C-mutated enhancer: 2.07 ± 0.55 (n = 2). The values for the Don3C intron cannot be reliably measured.

able above background level for the Don3C intron; however, a clear splicing enhancement effect was observed in the presence of the enhUGcAUGU sequence and, to a lesser extent, of the wild-type enhancer. A strong enhancer effect was also observed for the Don1C intron in the presence of enhUGc AUGU (Fig. 7A,B, compare lanes 2 and 3). As in the case of the Don3C intron, the wild-type Ψ-donor sequence retains enhancer activity, perhaps because the binding of U1-7G to enh-UGUAUGU is still relatively efficient (G_7 of U1 snRNA is facing U in position 2 of the enhancer instead of the canonical A-U pair). Finally, we verified that the mutated enhancer (enh-UGcAUGU) does not enhance splicing of introns Don1C and Don3C in a wild-type U1 snRNP environment (Fig. 7B). [Note that the copper-growth assays in wild-type U1 or U1-7G environment are not directly comparable, because the two strains have very different growth rates (our unpubl. results) and probably different metabolism of spliced RNAs.]

DISCUSSION

Splicing enhancement and exon recognition in splicing can be regarded as two different aspects of the same phenomenon, that is, the communication between splice sites across introns or exons. The presence of enhancer sequences can make this communication (i.e., the basis for cooperative binding) more efficient, which might be required in cases of intrinsically weak and/or regulated splice sites.

The communication between the two ends of an intron is also related to splicing enhancement in yeast. We have previously shown that the base pairing interaction of two short regions near the opposite extremities of the rp51b intron can function as a general splicing enhancer (Goguel & Rosbash, 1993; Libri et al., 1995; Charpentier & Rosbash, 1996). With this work we aimed at identifying other mechanisms capable of affecting splicing efficiency in yeast. The overall strategy was to isolate from large pools of introns short sequence motifs that would enhance splicing efficiency.

We isolated two families of sequences that improve splicing efficiency. The following evidence strongly suggests that their function is independent of intramolecular base pairing: (1) the absence in the intron of long stretches of purines that could base pair with the Us; (2) the isolation in a distinct randomization-selection experiment of similar sequences (although more degenerate) in a different region of the intron (5′ end, data not shown); (3) the fact that the Ψ-donor enhancer is active in one or two copies when inserted near the 5′ end of the intron, and that the effect is additive; and (4) the Ψ-donor enhancer can be exported to an intron that contains only the last 46 out of 315 nt of the rp51b intron.

An early role for the Ψ-donor enhancer

Most of the work presented in this article concerns the Ψ-donor enhancer. The procedure itself through which these sequences have been obtained suggests that very early steps of spliceosome assembly are affected by the enhancement effect. In fact, the selection for better splicers in a mud2Δ background can be regarded as the isolation of extragenic suppressors of the mud2Δ defect. This is consistent with the hypothesis that the step in which the enhancers are functional is the same step that is rendered rate limiting by the ablation of *MUD2* and coincides with very early intron recognition (Abovich et al., 1994; Abovich & Rosbash, 1997; Berglund et al., 1997; Fromont-Racine et al., 1997; Rain & Legrain, 1997). Consistently, we show that mutations of the donor site and of the secondary structure that are greatly affected by the absence of Mud2p can be at least partially suppressed by the enhancer sequences, and in both the presence and the absence of Mud2p.

The factors involved in splicing enhancement due to U-rich sequences

One of the outcomes of our selection experiments both in the 5′ and in the 3′ regions of the intron is the isolation of a class of U-rich sequences. As in the case of the Ψ-donor enhancer, these sequences could partially suppress 5′ splice site, secondary structure defects and the effects of the absence of Mud2p (data not shown). U-rich regions have been shown to be involved in the later selection of the 3′ splice site (Patterson & Guthrie, 1991), but an earlier role similar to the one played in higher eukaryotes by the polypyrimidine stretch and suggested by our experiments has never been proposed. The splicing factor PRP8 has been genetically and biochemically shown to recognize U-rich sequences (Umen & Guthrie, 1995) in the 3′ end of introns. There is no evidence, however, that PRP8p is involved in intron recognition before the entry of the U4/U6.U5 particle in the spliceosome (Whittaker et al., 1990; Teigelkamp et al., 1995a, 1995b).

In higher eukaryotes, U-rich sequences that are often located downstream of the branchpoint are recognized by the factor U2AF (Zamore & Green, 1989). However, as U-rich sequences were obtained during selection performed in the absence of Mud2p (the likely yeast ortholog of U2AF), it is unlikely that this factor is involved in splicing enhancement through binding to U-rich sequences. Rather, our data raise the interesting possibility that U-richness in yeast introns reflects a functional selective pressure due to as yet uncharacterized binding activities as previously reported for plant introns (Gniadkowski et al., 1996).

It is interesting that, while this work was in progress, O. Puig and coworkers showed that the U1 snRNP-

associated protein Nam8 recognizes in a sequence-dependent manner the region following the 5′ splice site and contributes thereby to intron recognition (Puig et al., 1999). It is possible that Nam8 plays an additional role in splicing enhancement by recognizing the U-rich enhancer and somehow favoring the interaction of U1 snRNP with the 5′ splice site.

The role of U1 snRNP and the mechanism of splicing enhancement

The impressive similarity of the Ψ-donor enhancer with a 5′ splice site strongly suggests that the determinants of enhancement and donor-site recognition are similar or identical. It is possible that an unknown factor or a complex recognizes the GUAU/CGU sequence and mediates splicing enhancement. In this case one can reasonably expect that the same factor(s) also recognizes the bona fide 5′ splice site, whose sequence is virtually identical.

An alternative hypothesis, which we favor, is that the U1 snRNP mediates the enhancement effect. This is suggested by the following lines of evidence: (1) the general consensus agreement between the natural 5′ splice sites and the selected sequences; (2) the mutational analysis concerning nucleotides U_{+2}, U_{+4}, and U_{+6}; (3) the fact that, when inserted near the natural 5′ splice site, the Ψ-donor enhancer is actually used as a bona fide splice site, besides working to enhance use of the upstream natural site; and (4) the observation that expression of a mutant U1 snRNA (U1-7G) can restore the enhancing activity of an inactive enhancer that is mutated at the complementary position (UGcAUGU).

The latter experiments also allow some interesting considerations concerning the mechanism of splicing enhancement. In principle, the U1 snRNP particle that is bound to the enhancer could improve splicing efficiency independently of being involved in 5′ splice site recognition. This enhancement mode would closely resemble the one occurring in metazoa (Grabowski et al., 1991; Kuo et al., 1991; Watakabe et al., 1993; Cohen et al., 1994; Staknis & Reed, 1994; Chiara & Reed, 1995; Lou et al., 1995; Hwang & Cohen, 1996a, 1996b). Alternatively, the main effect of the presence of a Ψ-donor enhancer would be to increase the local concentration of the same U1 snRNP particle that will subsequently recognize the bona fide 5′ splice site; in this case one would expect that the same particle recognizes efficiently both the enhancer and the 5′ splice site. Consistent with the second model, we show that a mutated U1-7G snRNP is able to efficiently complement the U_2C mutation in the enhancer sequence (enhUGcAUGU) only when the same mutation is also present in the 5′ splice site sequence (G:GcAUGU or G:GcAcGc). This effect is specific, because the nonmutated enhancer sequence is less efficient in a U1-7G snRNP environment, whereas the mutated enhancer is nonfunctional in a U1 snRNP wild-type environment.

Because of the large difference in growth rates between the U1wt and U1-7G strain, we cannot draw any strong conclusion on the suppression of the Don1C U_2C mutation in the latter environment. Also, this intron is probably recognized quite efficiently in a U1wt environment and the rate-limiting defect is likely to reside elsewhere (Collins & Guthrie, 1999; Siatecka et al., 1999). More important are the Don3C introns, which have a more obvious 5′ splice site recognition defect and are detectably spliced only in the U1-7G strain.

Our favorite hypothesis to explain the absence in the enhancer sequences of a GGUAUGU (which interacts more strongly than UGUAUGU with U1 snRNP) is that a GG doublet in the immediate proximity of the branchpoint somehow interferes with efficient splicing; this is consistent with our mutational analysis. Although we cannot exclude that this negative effect is linked to the intron we used, GG doublets are very rarely found in yeast introns. A survey of all the annotated yeast introns (Spingola et al., 1999) reveals that the frequency of GG doublets in the 10 nt immediately downstream of the branchpoint (0.012/nt) is significantly lower than in introns as a whole (0.025/nt) (D. Libri, unpubl. results). This suggests the existence of a counterselective pressure consistent with our mutational experiments.

The mutation U_4A in the Ψ-donor enhancer almost abolished the splicing-enhancement effect and a sequence containing an A in this position (#17, UGUAaGU) was only selected once. The survival of partially or totally inactive sequences is not uncommon in selection experiments (Irvine et al., 1991; He et al., 1996; Fields et al., 1997). The inactivity of the sequence might be surprising, however, because the U_4A mutation is expected to correct a Ψ-U mismatch in the interaction with U1 [a pseudo-uridine faces in U1 snRNA intron position +4 (Massenet et al., 1999)] and conservation of position +4 of the donor site is commonly ascribed to its base-pairing interaction with U6 (Kandels-Lewis & Séraphin, 1993; Lesser & Guthrie, 1993b). We have recently shown (D. Libri, unpubl. results) that the presence of the mismatch is an important requirement for the formation of a stable complex between the enhancer sequence (or a bona fide donor site) and U1 snRNP. When the enhancer is located downstream of the branchpoint, U1 snRNP forms a complex in vitro with UGUAUGU but not (or only to a limited extent) with UGUAaGU (our unpubl. data). This is still consistent with the implication of U1 snRNP in splicing enhancement.

Natural occurrence of donor-like sequences in intron-containing genes

The Ψ-donor enhancer sequence with a mismatch allowed (e.g., NGUAUGU or UGUANGU) is present in

about half of the annotated yeast introns (119 times in 232 introns) and is often located in the proximity of the branchpoint (46 times in a window comprising 50 nt upstream of the branchpoint and the AG). This is not unexpected, as presumably some of these sequences would be used as bona fide 5′ splice sites if located in a more favorable intron position (near to the natural 5′ splice site), thus resulting in the production of out-of-frame proteins. Using the PATSER program (see Materials and Methods), we analyzed the whole set of yeast introns as well as one similarly sized randomly chosen sample of intronless genes (ORFs from chromosomes I, 60,000 nt). Interestingly, the Ψ-donor pattern was identified 1,568 times in the intron set whereas only 1,184 times in chromosome I ORF samples. Although a more detailed statistical analysis is certainly desirable (notably taking into account the differences in mono- and dinucleotide composition), it is likely that some of the natural Ψ-donor enhancer variants present in yeast introns are functional enhancers that might operate in association with inefficiently spliced introns.

MATERIALS AND METHODS

Cloning, mutagenesis, and pool constructions

Construction of the rp51b-*CUP1*-based reporter system was previously described; mutations in the rp51b intron were introduced by a PCR-based mutagenesis technique described in Libri et al. (1995). Vent polymerase was used to minimize errors and extensive sequencing was performed to verify the absence of unwanted mutations. This method was also used to construct the pools of sequences used for the randomization-selection experiments. An oligonucleotide containing 8 random nt starting at position +3 from the last nucleotide of the branch site was used to randomize the DB4 region. A plasmid containing a stuffer fragment cloned in the *CUP1* gene was used to prepare the vector for cloning of the pools. This was necessary to minimize the presence of nonmutated sequences in the pool. A limited (10 clones) random sequencing of the pool revealed some representative bias towards Gs and against Cs (Gs: 38%; Us: 30%; As: 18%; Cs: 12%) that might affect pool complexity.

Introns 3mUb1, 5mUb1, SB4/DB4, and b/bI/b.st are described in Libri et al. (1995), and intron bi/bi in Goguel & Rosbash (1993). Intron Art77 was constructed as follows: an *Nco*I site was introduced in the rp51b intron at position −9 with respect to the first nucleotide of the branch site; a DNA fragment obtained by annealing and filling in of two partially complementary oligonucleotides was then cloned between the *Nco*I and the *BamH*I site (located 15 nt upstream of the ATG of the intron-containing *CUP1* gene) to replace the 5′ portion of the intron. The overall length of the intron is 99 nt, of which 53 derive from the synthetic portion. The donor site of this intron contains the U_4C mutation. Moreover this intron contains a putative binding site for the hU1A protein (AUUG CACUCC) closed by two GC pairs and an imperfect stem afterwards, which has been recently shown (Tang et al., 1996) to be a good binding site for the yeast U2B″ protein. It is a reasonable hypothesis that binding of the yU2B″p to the intron slows down early steps recognition, which might be synthetically enhanced by the absence of Mud2p. A copy of the Ψ enhancer sequence (UGUAUGU) was introduced in introns Don2C-5′enh and Don2C-5′+3′enh by mutagenesis as described above. The endogenous sequence was replaced starting at intron position +18. Don2C-5′rnd-pool is a pool of clones containing a randomized region of the same size starting at the same intron position (+18). Intron Don2C-5′rnd is one of the clones of this pool.

Plasmids U1-7G and U1-3C (Séraphin & Rosbash, 1989) are a gift of B. Séraphin.

Analysis of the selected sequences

To analyze the selected sequences we used the programs of the CONSENSUS package (Hertz et al., 1990; Hertz & Stormo, 1999). The program WCONSENSUS was first used to align all the sequences selected in the SB4/db4 intron and mud2Δ background (flanking sequences were not included in this first analysis) and to identify a common pattern. We designated as the prior frequencies for each nucleotide either the estimated nucleotide frequencies in the starting pool or the observed frequencies in the selected pool (which is a more conservative estimate for the nucleotide bias in the pool). The program identified a pattern of 7 nt in 9 out of 18 sequences. To account for the possible occasional contribution of the flanking sequences, this pattern and the associated weighting matrix were then used with the PATSER program to score all the sequences obtained, and 2 nt of the flanking sequences on each side were added for this analysis. Because similar results were obtained for the selection winners in the SB4/db4 intron and MUD2 background, the two sets of sequences were combined. All the sequences (18 out of 39) that scored positive according to the default parameters of the program (score >0 and $\ln(p\text{-value}) > -3.52$) were retained. The final consensus was determined for this latter set of sequences (18). PATSER scores varied from 4.09 (sequence #15, $\ln(p\text{-value}) = -5.99$) to 7.67 (sequences #4,9 and W1, $\ln(p\text{-value}) = -9.86$). Sequence #12,16 was realigned manually by introducing a gap. The program GMAT-INF-GC was used to determine the information content (Schneider et al., 1986; Stormo & Hartzell, 1989) for the individual positions of the consensus sequence. The starting pool bias was included in the calculation. A small sample correction (0.12 bits) representing the average background expected from a random alignment ($n = 19$) was automatically subtracted from the information content at every position of the alignment. The standard deviation expected from a random alignment was estimated by the program to be 0.10 bits. All the positions of the alignment (with the exception of position −2) have an information content significantly different from 0. The WCONSENSUS program did not detect any additional patterns in the remaining sequences (which is, perhaps, not unexpected given the peculiar composition of the remaining sequences). The U-rich family of sequences was defined based on the extreme U-richness of its components. Sequences that cannot be obviously assigned to either family have been classified as “others.”

Note that a preliminary analysis of the selection winners in the presence of Mud2p (sequences in Fig. 2C) suggested a somewhat different family assignment that we referred to pre-

viously (Libri et al., 1995) as data not shown. We saw in sequences W10, W11, and W3 the potentiality to base pair with an intronic region immediately downstream of UB1 (roughly starting at intron position 43). However the subsequent selections performed in the mud2Δ background and described here suggested the presence of the Ψ-donor family to which we assigned W10 and W11 based on their high score (6.41 and 7.05, respectively).

The PATSER program was also used to search for the occurrence of the Ψ-donor pattern in yeast introns (Spingola et al., 1999) and randomly chosen intronless ORFs from chromosomes I. The matrix derived from the alignment of the Ψ-donor family was used as above, taking into account the nucleotide composition of every set. Bona fide 5′ splice sites have not been included in this analysis. Only motifs that scored positive according to the default parameters of the program (score >0) were retained.

Primer extension

Primer extension was performed as described previously (Libri et al., 1995) with oligonucleotides complementary to the CUP1 gene (5′-GGCATTGGCACTCATGACCTTCA-3′) or to U1 snRNA (5′-CCAAGGAGTTTGCATCAATGACT-3′).

Transformations, selections, and growth assays

The pools of sequences were transferred to bacteria by electroporation, and subsequently to yeast by the lithium acetate/PEG method as described (Gietz et al., 1992). The yeast transformants were harvested as soon as colonies were visible and plated on selective copper-containing media. Pool complexity was estimated to be as follows: 4×10^5 and 3×10^5 (selection on SB4/DB4 in a mud2Δ and wild-type strain environment respectively) and 2×10^5 (selection on rp51b in a mud2Δ environment). About 90% of the clones contained an intron variant properly integrated into the *CUP1* reporter as verified by PCR.

Growth assays were performed by spotting an aliquot of an overnight culture on plates containing various concentrations of copper. Copper resistance level of the various mutants is defined as the first concentration of copper at which growth is barely detectable after 4 days. Two to five independent copper-growth assays have been performed for each construction. Primer extension experiments have also been performed for most of the clones.

For the genetic suppression experiment in the merodiploid strain, a centromeric plasmid bearing a copy of the U1 snRNA gene containing an A-to-G transition in position 7 (kind gift of B. Séraphin) was used to transform the mud2Δ strain containing the various reporters described. For the genetic suppression assays in a U1-7G snRNA only environment, introns Don3C and Don1C with and without the wild-type and mutated enhancer sequences were directly cloned by homologous recombination into a *CUP1* reporter vector bearing a TRP marker in the DLY5 strain.

Yeast strains

The yeast strain DLY1 (*MATa*, leu2-3, leu2-112, ura3-52, trp1-289, arg4, ade2, cup1Δ, *RP51b::URA3*, *MUD2::ADE*) was obtained by deleting the *MUD2* gene in the previously described (Libri et al., 1995) Y59ΔCUPΔB strain. DLY5 strain (MATa, leu2-3, leu2-112, ura3-52, trp1-289, arg4, ade2, *SNR19::LEU2*, pCEN-*ADE2-SNR19* $^{U1\text{-}7G}$, *CUP1*::Kan^r) was derived from the BSY20 strain (kind gift of B. Séraphin).

ACKNOWLEDGMENTS

We thank K. Adelman, J. Banroques, A. Berglund, E. Brody, A. Expert-Besançon, P. Legrain, J. Marie, M. Moore, and F. Stutz for critical reading of the manuscript, and B. Séraphin for sharing results prior to publication and gift of the BSY20 strain and U1-7G and U1-3C plasmids. Thanks also to G. Hertz for help and advice with the CONSENSUS package of programs and X. Jeannin for implementing the program GMAT-INF-GC on a local server. This work was supported by the Centre National de la Recherche Scientifique, the Fondation pour la recherche Médicale, and by a grant from the National Institutes of Health to M.R.

Received June 10, 1999; returned for revision July 21, 1999; revised manuscript received November 29, 1999

REFERENCES

Abovich N, Liao XC, Rosbash M. 1994. The yeast MUD2 protein: An interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. *Genes & Dev 8*:843–854.

Abovich N, Rosbash M. 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell 89*:403–412.

Arning S, Gruter P, Bilbe G, Kramer A. 1996. Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. *RNA 2*:794–810.

Balvay L, Libri D, Gallego M, Fiszman MY. 1992. Intronic sequence with both negative and positive effects on the regulation of alternative transcripts of the chicken beta tropomyosin transcripts. *Nucleic Acids Res 20*:3987–3992.

Berget SM. 1995. Exon recognition in vertebrate splicing. *J Biol Chem 270*:2411–2414.

Berglund JA, Abovich N, Rosbash M. 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes & Dev 12*:858–867.

Berglund JA, Chua K, Abovich N, Reed R, Rosbash M. 1997. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell 89*:781–787.

Black DL. 1992. Activation of c-src neuron-specific splicing by an unusual RNA element in vivo and in vitro. *Cell 69*:795–807.

Bruzik JP, Maniatis T. 1995. Enhancer-dependent interaction between 5′ and 3′ splice sites in *trans*. *Proc Natl Acad Sci USA 92*:7056–7059.

Caputi M, Casari G, Guenzi S, Tagliabue R, Sidoli A, Melo CA, Baralle FE. 1994. A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. *Nucleic Acids Res 22*:1018–1022.

Carlo T, Sterner DA, Berget SM. 1996. An intron splicing enhancer containing a G-rich repeat facilitates inclusion of a vertebrate micro-exon. *RNA 2*:342–353.

Charpentier B, Rosbash M. 1996. Intramolecular structure in yeast introns aids the early steps of in vitro spliceosome assembly. *RNA 2*:509–522.

Chiara MD, Reed R. 1995. A two-step mechanism for 5′ and 3′ splice-site pairing. *Nature 375*:510–513.

Cohen JB, Snow JE, Spencer SD, Levinson AD. 1994. Suppression of mammalian 5′ splice-site defects by U1 small nuclear RNAs from a distance. *Proc Natl Acad Sci USA 91*:10470–10474.

Collins CA, Guthrie C. 1999. Allele-specific genetic interactions between prp8 and RNA active site residues suggest a function for

prp8 at the catalytic core of the spliceosome. *Genes & Dev 13*:1970–1982.

Cooper TA. 1992. In vitro splicing of cardiac troponin T precursors. Exon mutations disrupt splicing of the upstream intron. *J Biol Chem 267*:5330–5338.

Cooper TA, Ordahl CP. 1989. Nucleotide substitutions within the cardiac troponin T alternative exon disrupt pre-mRNA alternative splicing. *Nucleic Acids Res 17*:7905–7921.

Crispino JD, Sharp PA. 1995. A U6 snRNA:pre-mRNA interaction can be rate-limiting for U1-independent splicing. *Genes & Dev 9*:2314–2323.

Dirksen WP, Hampson RK, Sun Q, Rottman FM. 1994. A purine-rich exon sequence enhances alternative splicing of bovine growth hormone pre-mRNA. *J Biol Chem 269*:6431–6436.

Fields DS, He Y, Al-Uzri AY, Stormo GD. 1997. Quantitative specificity of the Mnt repressor. *J Mol Biol 271*:178–194.

Fouser LA, Friesen JD. 1986. Mutations in a yeast intron demonstrate the importance of specific conserved nucleotides for the two stages of nuclear mRNA splicing. *Cell 45*:81–93.

Fromont-Racine M, Rain JC, Legrain P. 1997. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens [see comments]. *Nature Genet 16*:277–282.

Fu XD. 1995. The superfamily of arginine/serine-rich splicing factors. *RNA 1*:663–680.

Gietz D, Jean AS, Woods RA, Schiestl RH. 1992. Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res 20*:1425.

Gniadkowski M, Hemmings-Mieszczak M, Klahre U, Liu HX, Filipowicz W. 1996. Characterization of intronic uridine-rich sequence elements acting as possible targets for nuclear proteins during pre-mRNA splicing in *Nicotiana plumbaginifolia*. *Nucleic Acids Res 24*:619–627.

Goguel V, Rosbash M. 1993. Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. *Cell 72*:893–901.

Grabowski PJ, Nasim FU, Kuo HC, Burch R. 1991. Combinatorial splicing of exon pairs by two-site binding of U1 small nuclear ribonucleoprotein particle. *Mol Cell Biol 11*:5919–5928.

He YY, Stockley PG, Gold L. 1996. In vitro evolution of the DNA binding sites of *Escherichia coli* methionine repressor, MetJ. *J Mol Biol 255*:55–66.

Hertz GZ, Hartzell GW, Stormo GD. 1990. Identification of consensus patterns in unaligned DNA sequences known to be functionally related. *Comput Appl Biosci 6*:81–92.

Hertz GZ, Stormo GD. 1999. Identifying DNA and protein patterns with statistically significant alignment of multiple sequences. *Bioinformatics 15*:563–577.

Hodges D, Bernstein SI. 1994. Genetic and biochemical analysis of alternative RNA splicing. *Adv Genet 31*:207–281.

Huh GS, Hynes RO. 1993. Elements regulating an alternatively spliced exon of the rat fibronectin gene. *Mol Cell Biol 13*:5301–5314.

Huh GS, Hynes RO. 1994. Regulation of alternative pre-mRNA splicing by a novel repeated hexanucleotide element. *Genes & Dev 8*:1561–1574.

Humphrey MB, Bryan J, Cooper TA, Berget SM. 1995. A 32-nucleotide exon-splicing enhancer regulates usage of competing 5′ splice sites in a differential internal exon. *Mol Cell Biol 15*:3979–3988.

Hwang DY, Cohen JB. 1996a. Base pairing at the 5′ splice site with U1 small nuclear RNA promotes splicing of the upstream intron but may be dispensable for splicing of the downstream intron. *Mol Cell Biol 16*:3012–3022.

Hwang DY, Cohen JB. 1996b. U1 snRNA promotes the selection of nearby 5′ splice sites by U6 snRNA in mammalian cells. *Genes & Dev 10*:338–350.

Irvine D, Tuerk C, Gold L. 1991. SELEXION. Systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. *J Mol Biol 222*:739–761.

Kandels-Lewis S, Séraphin B. 1993. Role of U6 snRNA in 5′ splice site selection [see comments]. *Science 262*:2035–2039.

Kao HY, Siliciano PG. 1996. Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. *Mol Cell Biol 16*:960–967.

Konforti BB, Konarska MM. 1995. A short 5′ splice site RNA oligo can participate in both steps of splicing in mammalian extracts. *RNA 1*:815–827.

Konforti BB, Koziolkiewicz MJ, Konarska MM. 1993. Disruption of base pairing between the 5′ splice site and the 5′ end of U1 snRNA is required for spliceosome assembly. *Cell 75*:863–873.

Kuo HC, Nasim FH, Grabowski PJ. 1991. Control of alternative splicing by the differential binding of U1 small nuclear ribonucleoprotein particle. *Science 251*:1045–1050.

Lavigueur A, La Branche 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.

Lesser CF, Guthrie C. 1993a. Mutational analysis of pre-mRNA splicing in *Saccharomyces cerevisiae* using a sensitive new reporter gene, CUP1. *Genetics 133*:851–863.

Lesser CF, Guthrie C. 1993b. Mutations in U6 snRNA that alter splice site specificity: Implications for the active site [see comments]. *Science 262*:1982–1988.

Libri D, Balvay L, Fiszman MY. 1992. In vivo splicing of the beta tropomyosin pre-mRNA: A role for branchpoint and donor site competition. *Mol Cell Biol 12*:3204–3215.

Libri D, Stutz F, McCarthy T, Rosbash M. 1995. RNA structural patterns and splicing: Molecular basis for an RNA-based enhancer. *RNA 1*:425–436.

Long M, de Souza SJ, Rosenberg C, Gilbert W. 1998. Relationship between “proto-splice sites” and intron phases: Evidence from dicodon analysis. *Proc Natl Acad Sci USA 95*:219–223.

Lou H, Yang Y, Cote GJ, Berget SM, Gagel RF. 1995. An intron enhancer containing a 5′ splice site sequence in the human calcitonin/calcitonin gene-related peptide gene. *Mol Cell Biol 15*:7135–7142.

Madhani HD, Guthrie C. 1994. Dynamic RNA–RNA interactions in the spliceosome. *Annu Rev Genet 28*:1–26.

Manley JL, Tacke R. 1996. SR proteins and splicing control. *Genes & Dev 10*:1569–1579.

Massenet S, Motorin Y, Lafontaine DL, Hurt EC, Grosjean H, Branlant C. 1999. Pseudouridine mapping in the *Saccharomyces cerevisiae* spliceosomal U small nuclear RNAs (snRNAs) reveals that pseudouridine synthase pus1p exhibits a dual substrate specificity for U2 snRNA and tRNA. *Mol Cell Biol 19*:2142–2154.

Moore MJ, Query CC, Sharp PA. 1993. Splicing of precursors to mRNAs by the spliceosome. In: Gesteland RF, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 303–357.

Newman A. 1987. Specific accessory sequences in *Saccharomyces cerevisiae* introns control assembly of pre-mRNAs into spliceosomes. *EMBO J 6*:3833–3839.

Parker R, Siliciano PG, Guthrie C. 1987. Recognition of the TAC TAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell 49*:229–239.

Patterson B, Guthrie C. 1991. A U-rich tract enhances usage of an alternative 3′ splice site in yeast. *Cell 64*:181–187.

Puig O, Gottschalk A, Fabrizio P, Séraphin B. 1999. Interaction of the U1 snRNP with non-conserved intronic sequences affects 5′ splice site selection. *Genes & Dev* 13:569–580.

Rain JC, Legrain P. 1997. In vivo commitment to splicing in yeast involves the nucleotide upstream from the branch site conserved sequence and the Mud2 protein. *EMBO J 16*:1759–1771.

Reed R, Maniatis T. 1986. A role for exon sequences and splice-site proximity in splice-site selection. *Cell 46*:681–690.

Robberson BL, Cote GJ, Berget SM. 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol Cell Biol 10*:84–94.

Rymond BC, Rosbash M. 1992. Yeast pre-mRNA splicing. In: Jones EW, Pringle JR, Broach JR, eds. *The molecular biology of the yeast Saccharomyces: Gene expression*, Vol. 2. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 143–192.

Schneider TD, Stormo GD, Gold L, Ehrenfeucht A. 1986. Information content of binding sites on nucleotide sequences. *J Mol Biol 188*:415–431.

Séraphin B, Kandels-Lewis S. 1993. 3′ splice site recognition in *S. cerevisiae* does not require base pairing with U1 snRNA. *Cell 73*:803–812.

Séraphin B, Rosbash M. 1989. Mutational analysis of the interactions between U1 small nuclear RNA and pre-mRNA of yeast. *Gene 82*:145–151.

Siatecka M, Reyes JL, Konarska MM. 1999. Functional interactions of prp8 with both splice sites at the spliceosomal catalytic center. *Genes & Dev 13*:1983–1993.

Sirand-Pugnet P, Durosay P, Brody E, Marie J. 1995. An intronic (A/U)GGG repeat enhances the splicing of an alternative intron of the chicken beta-tropomyosin pre-mRNA. *Nucleic Acids Res 23*:3501–3507.

Somasekhar MB, Mertz JE. 1985. Exon mutations that affect the choice of splice sites used in processing the SV40 late transcripts. *Nucleic Acids Res 13*:5591–5609.

Spingola M, Grate L, Haussler D, Ares M Jr. 1999. Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. *RNA 5*:221–234.

Staffa A, Cochrane A. 1995. Identification of positive and negative splicing regulatory elements within the terminal tat-rev exon of human immunodeficiency virus type 1. *Mol Cell Biol 15*:4597–4605.

Staknis D, Reed R. 1994. SR proteins promote the first specific recognition of pre-mRNA and are present together with the U1 small nuclear ribonucleoprotein particle in a general splicing enhancer complex. *Mol Cell Biol 14*:7670–7682.

Stormo GD, Hartzell GW. 1989. Identifying protein-binding sites from unaligned DNA fragments. *Proc Natl Acad Sci USA 86*:1183–1187.

Streuli M, Saito H. 1989. Regulation of tissue-specific alternative splicing: Exon-specific *cis*-elements govern the splicing of leukocyte common antigen pre-mRNA. *EMBO J 8*:787–796.

Stutz F, Rosbash M. 1994. A functional interaction between Rev and yeast pre-mRNA is related to splicing complex formation. *EMBO J 13*:4096–4104.

Talerico M, Berget SM. 1990. Effect of 5′ splice site mutations on splicing of the preceding intron. *Mol Cell Biol 10*:6299–6305.

Tanaka K, Watakabe A, Shimura Y. 1994. Polypurine sequences within a downstream exon function as a splicing enhancer. *Mol Cell Biol 14*:1347–1354.

Tang J, Abovich N, Rosbash M. 1996. Identification and characterization of a yeast gene encoding the U2 small nuclear ribonucleoprotein particle B″ protein. *Mol Cell Biol 16*:2787–2795.

Teigelkamp S, Newman AJ, Beggs JD. 1995a. Extensive interactions of PRP8 protein with the 5′ and 3′ splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *EMBO J 14*:2602–2612.

Teigelkamp S, Whittaker E, Beggs JD. 1995b. Interaction of the yeast splicing factor PRP8 with substrate RNA during both steps of splicing. *Nucleic Acids Res 23*:320–326.

Tian H, Kole R. 1995. Selection of novel exon recognition elements from a pool of random sequences. *Mol Cell Biol 15*:6291–6298.

Tian M, Maniatis T. 1993. A splicing enhancer complex controls alternative splicing of doublesex pre-mRNA. *Cell 74*:105–114.

Tian M, Maniatis T. 1994. A splicing enhancer exhibits both constitutive and regulated activities. *Genes & Dev 8*:1703–1712.

Umen JG, Guthrie C. 1995. A novel role for a U5 snRNP protein in 3′ splice site selection. *Genes & Dev 9*:855–868.

Valcárcel J, Green MR. 1996. The SR protein family: Pleiotropic functions in pre-mRNA splicing. *Trends Biochem Sci 21*:296–301.

Watakabe A, Tanaka K, Shimura Y. 1993. The role of exon sequences in splice site selection. *Genes & Dev 7*:407–418.

Whittaker E, Lossky M, Beggs JD. 1990. Affinity purification of spliceosomes reveals that the precursor RNA processing protein PRP8, a protein in the U5 small nuclear ribonucleoprotein particle, is a component of yeast spliceosomes. *Proc Natl Acad Sci USA 87*:2216–2219.

Zamore PD, Green MR. 1989. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc Natl Acad Sci USA 86*:9243–9247.