The EMBO Journal Vol.16 No.7 pp.1759–1771, 1997

In vivo commitment to splicing in yeast involves the nucleotide upstream from the branch site conserved sequence and the Mud2 protein

Jean-Christophe Rain and Pierre Legrain[1]

Laboratoire du Métabolisme des ARN, CNRS URA 1149, Département des Biotechnologies, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France

[1]Corresponding author

Pre-mRNA splicing is a stepwise nuclear process involving intron recognition and the assembly of the spliceosome followed by intron excision. We previously developed a pre-mRNA export assay that allows the discrimination between early steps of spliceosome formation and splicing *per se*. Here we present evidence that these two assays detect different biochemical defects for point mutations. Mutations at the 5′ splice site lead to pre-mRNA export, whereas 3′ splice site mutations do not. A genetic screen applied to mutants in the branch site region shows that all positions in the conserved TACTAAC sequence are important for intron recognition. An exhaustive analysis of pre-mRNA export and splicing defects of these mutants shows that the *in vivo* recognition of the branch site region does not involve the base pairing of U2 snRNA with the pre-mRNA. In addition, the nucleotide preceding the conserved TACTAAC sequence contributes to the recognition process. We show that a T residue at this position allows for optimal intron recognition and that in natural introns, this nucleotide is also used preferentially. Moreover, the Mud2 protein is involved in the recognition of this nucleotide, thus establishing a role for this factor in the *in vivo* splicing pathway.

Keywords: β-galactosidase/nuclear export/pre-mRNA/*S.cerevisiae*/spliceosome

Introduction

In eukaryotic cells, pre-mRNA splicing takes place in a large nuclear complex called the spliceosome. This complex is dynamically formed on nascent transcripts through recognition of conserved intronic sequences by several small nuclear ribonucleoprotein particles (U1, U2, U4/6 and U5 snRNP) and additional nuclear proteins. In the past few years, a lot of data have been accumulated on splicing reactions *per se* in the spliceosome. For example, many RNA–RNA interactions between snRNAs or between snRNAs and pre-mRNA, that occur during the formation of the mature complex in which the cleavage reactions take place, have been characterized (Madhani and Guthrie, 1994; Ares and Weiser, 1995). Nevertheless, studies on alternative splicing in mammals and other studies in the yeast *Saccharomyces cerevisiae* show that the catalysis is not the limiting step for splicing (Pikielny and Rosbash, 1985). In contrast, intron identification and stable spliceosome formation represent central elements for efficient and specific splicing *in vivo* (Hodges and Beggs, 1994). These steps have been analysed extensively *in vitro* but still remain poorly understood *in vivo*.

In vitro studies led to the definition of the commitment complexes as the earliest stable complexes formed between pre-mRNA and spliceosome constituents that can be chased into mature spliceosome in which splicing occurs (Legrain *et al.*, 1988). In the yeast *S.cerevisiae*, the commitment complex 1 (CC1) containing the U1 snRNP is first formed involving the 5′ splice site region. The CC2 complex is then formed from the CC1 complex after interactions have occurred between the conserved branch site region (the TACTAAC sequence in yeast introns) and additional splicing factors. However, the U2 snRNP (at least not the U2 snRNA–UACUAAC pre-mRNA base pairing) is not required for this complex formation (Séraphin and Rosbash, 1990, 1991). In mammals, a similar commitment step has been characterized (Michaud and Reed, 1991). Complex E formation requires the U1 snRNP, SR proteins [that could bypass the U1 snRNP requirement when present in excess (Crispino *et al.*, 1994; Tarn and Steitz, 1994)] and the U2AF factor. It involves both the 5′ and the 3′ splice site regions (Michaud and Reed, 1993). The branch site region is then required for the ATP-dependent formation of complex A. Later on, the U4/U6·U5 triple particle binds to the splicing complexes and the mature spliceosome is formed, involving base pairing of U2, U5 and U6 snRNAs with the pre-mRNA (Madhani and Guthrie, 1994; Ares and Weiser, 1995). Only then can the catalytic step I occur. The role of the branch site region in the transition from complex E to complex A is not completely understood. Most of this process seems to be sequence independent, but site-specific labelling studies show that this region is the centre of dynamic modifications (Gozani *et al.*, 1994; MacMillan *et al.*, 1994).

The targeting of U2 snRNA to pre-mRNA and the stabilization of the resulting interaction are dependent on associated protein factors. Biochemical studies in mammalian cells have identified two protein complexes, SF3a and SF3b, as part of the 17S U2 snRNP (Brosi *et al.*, 1993b). In yeast, analyses of several heat-sensitive *prp* mutants for a pre-mRNA export phenotype have led to focusing on Prp9p as a key splicing factor during pre-spliceosome assembly (Legrain and Rosbash, 1989). Identification of a *PRP21* mutant allele as suppressor of the *prp9-1* heat-sensitive mutation (Chapon and Legrain, 1992), the subsequent demonstration of physical interactions between PRP9p, PRP11p and PRP21p and the definition of multiple genetic interactions between those genes (Ruby *et al.*, 1993; Wells and Ares, 1994) led to the identification of a new complex component of the spliceosome in *S.cerevisiae* (Legrain and Chapon, 1993). Cross-reactivity of antibodies (Behrens *et al.*, 1993),

sequence homology (Bennett and Reed, 1993; Brosi *et al.*, 1993a; Krämer *et al.*, 1995), functional complementation between heterologous genes (Krämer *et al.*, 1994) and two-hybrid assays performed between heterologous proteins (Rain *et al.*, 1996) establish that Prp9p, Prp11p and Prp21p are homologous to the proteins that form SF3a, i.e. SF3a60 (SAP61), SF3a66 (SAP62) and SF3a120 (SAP114), respectively. More recently, the yeast *CUS1* gene was identified in a genetic screen for suppressors of a cold-sensitive snRNA U2 mutant (Wells *et al.*, 1996). Sequence analysis revealed that Cus1p is the yeast SF3b^{145}/SAP145 homologue (Gozani *et al.*, 1996), showing a strong structural and functional conservation between yeast and mammalian pre-spliceosome assembly.

Several proteins are found associated with the U1 snRNP, such as Snp1p and Mud1p, the yeast homologues of human U1-70K and U1A proteins, respectively (Smith and Barrell, 1991; Liao *et al.*, 1993). The Mud2 protein, which has a weak homology to the human U2AF65 splicing factor, was identified in a screen for synthetic lethality with a U1 snRNA mutation (Abovich *et al.*, 1994). Mud2p is part of the commitment complex CC2 defined *in vitro* and cross-links to the pre-mRNA only in the presence of a functional branch site sequence. The *MUD2* gene is not essential in a wild-type yeast strain, but its disruption is lethal in a *prp11-1* mutant or *MUD1*-disrupted strains (Abovich *et al.*, 1994). Since *PRP11* is associated with U2 snRNP, these results suggest that Mud2p participates in the bridge between the 5′ splice site and the branch site region recognition complexes. However, the precise requirements for the formation of a stable pre-spliceosome *in vivo* are unknown.

Several years ago, we designed an assay using an artificial short intron that allows the measurement of specific pre-mRNA nuclear retention, i.e. the analysis of interactions between nuclear factors and pre-mRNA without taking splicing activity into account (Legrain and Rosbash, 1989). Using deletion mutants, we showed that conserved intronic sequences are specifically required for active nuclear retention of pre-mRNA. In addition, our results also showed that U1 snRNP base pairing with the 5′ splice site region and several splicing factors including Prp9p, a U2 snRNP-associated protein, are implicated in this process.

Here, we use a genetic screen to isolate branch point region mutants that are defective for efficient nuclear intron recognition. We show that most if not all substitutions in the TACTAAC sequence lead to a pre-mRNA export phenotype. Interestingly, this study identifies the position preceding this conserved sequence as an element of the recognition process which also involves the Mud2 protein.

Results

Point mutations at the 5′ splice site lead to pre-mRNA export

We previously demonstrated that a mutant pre-mRNA deleted for its 5′ splice site is exported efficiently toward the cytoplasm. This result was obtained using a reporter gene containing a synthetic intron with an open reading frame (ORF) fused to the β-galactosidase-coding sequence (Figure 1A). To explore the phenotype of point mutants for pre-mRNA export, we constructed variants of this reporter. A first series of mutants in the 5′ splice site region were made by PCR (see Materials and methods), and the β-galactosidase activity of the reporter gene was measured (Figure 1B). Results are expressed as the percentage of the enzymatic activity of the same reporter gene that does not contain an intron. Various point mutations in the 5′ splice site consensus intronic sequence lead to a significant export of the mutant pre-mRNA. The strongest mutants analysed are the G1→A and G5→A substitutions. Their phenotype is similar to that of the complete deletion of the 5′ splice site region (Δ5′SS, see also Legrain and Rosbash, 1989). Mutations in two exonic positions upstream of the 5′ splice site do not show any increase in pre-mRNA export (Figure 1B). Primer extension analyses confirmed that splicing efficacy is very poor for 5′ splice site intronic mutants and that the G1→A mutant accumulates a lariat intermediate (data not shown). We also analysed two intronic mutations in the 3′ splice site (AG→AC and CG, Figure 1B). They have no effect on pre-mRNA export. A primer extension analysis of these mutant transcripts shows a block after the first step of splicing (data not shown), as already described in the literature (Rymond and Rosbash, 1985; Fouser and Friesen, 1987). Those results demonstrate that the pre-mRNA export assay detects specific substitutions in the pre-mRNA sequence and is different from a splicing assay. This first limited study suggests that this assay could be used for a screen of point mutations leading to pre-mRNA export.

Various point mutations in the branch site region lead to pre-mRNA export

No systematic analyses on TACTAAC substitutions are available in the literature, and one cannot predict how many and which substitutions lead to an export phenotype. Thus, in order to analyse the effects of point mutations in the branch site region on the export of pre-mRNA, we designed a strategy of random mutagenesis and screening. We amplified a region of the pre-mRNA sequence by PCR using oligonucleotides degenerated over 12 nucleotides, including the TACTAAC box (underlined sequence in Figure 1A). The number of mutant positions in the oligonucleotide was chosen to maximize the number of single mutations (see Materials and methods). Any of the possible single mutants can be screened for export phenotype since none of them contains a stop codon. PCR fragments were cloned into the pre-mRNA export reporter (Figure 1A). The actual substitution rate was 0.07 per position, and no major bias in nucleotide composition was observed (see Materials and methods). The yeast strain MGD353-46D was transformed with mutagenized plasmids, and 900 independent transformants were screened for pre-mRNA export phenotype.

The screening procedure could not be performed directly on cells grown on galactose, since the wild-type reporter is exported at a low rate and β-galactosidase is very stable and accumulates in yeast cells (Figure 1B, Legrain and Rosbash, 1989). Cells were grown on nylon membranes placed on glycerol- and lactate-containing medium. Then, the membranes were transferred onto galactose-containing medium for 1 h to allow reporter gene induction (see Materials and methods). Positives clones were picked up and re-assayed with calibrated cell suspensions, allowing a semi-quantitative measurement of β-galactosidase. Eighty five positive clones were isolated, their plasmids were extracted and their introns were sequenced. Forty point mutations were

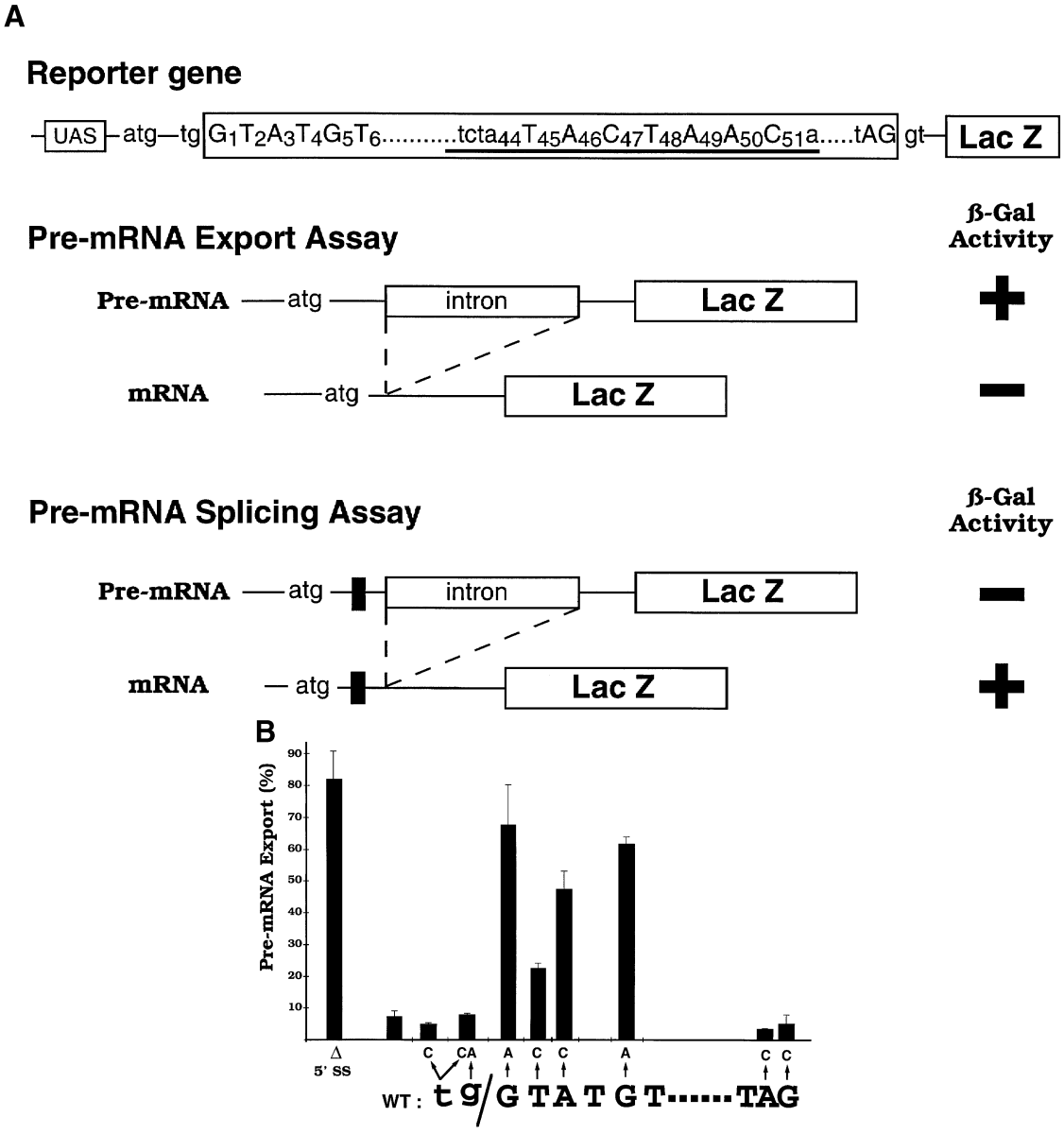

Fig. 1. Pre-mRNA export and splicing assays. (**A**) Constructs used in pre-mRNA nuclear export and splicing assays. A small synthetic intron was introduced downstream of the initiation codon in a galactose-inducible *lacZ* reporter gene (pLGSD5, Guarente *et al.*, 1982). The number of nucleotides in exon 1 allows either the pre-mRNA (pre-mRNA export assay) or the spliced RNA (pre-mRNA splicing assay) translational product to encode the β-galactosidase (see Materials and methods for details). The start codon, the sequence of exon–intron junctions and the intron sequence around the TACTAAC box are indicated. (**B**) The pre-mRNA export of 5′ and 3′ splice site mutants is expressed as the percentage of the β-galactosidase activity obtained with the pLGSD5 reporter gene. Two independent transformants were analysed in duplicate.

identified (Figure 2). All but one fall within the TACTAAC sequence and at least one substitution at any of the seven TACTAAC positions was selected. Among the 29 double mutants, none was found with the two mutations outside the TACTAAC sequence (data not shown). This result indicates that our assay selects for a subpopulation of mutagenized positions. Positions known to be important for splicing are also important for intron recognition, and point mutations are sufficient to obtain a significant pre-mRNA export phenotype.

The screening was not saturated and we performed site-directed mutagenesis for the eight additional substitutions (six in the TACTAAC and two at the position preceding it; see Figure 2). All six additional mutants in the TACTAAC sequence and the G substitution at the preceding A44 nucleotide were blue in a X-gal assay, whereas the A44→T mutant remained white (see below).

In conclusion, point mutations in conserved intronic sequences lead to the export of unspliced pre-mRNA. In particular, every single position within the TACTAAC sequence exhibits this phenotype, suggesting a key role for these nucleotides in intron recognition in addition to their role in the splicing reaction *per se*.

The strength of splicing defects is not correlated with the pre-mRNA export phenotype

The splicing deficiency of TACTAAC mutants has been tested on a limited number of substitutions with various

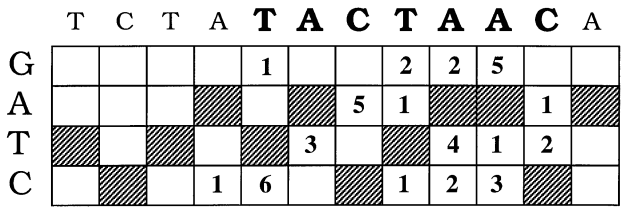

Fig. 2. Isolation of pre-mRNA export mutants. The number of mutants is scored for each substitution in the branch site region. The partially randomized sequence is given on the top and the nature of the substitution is indicated on the left. An empty box indicates that no mutant was obtained, and hatched boxes represent the wild-type sequence.

substrates (Jacquier *et al.*, 1985; Fouser and Friesen, 1986; Vijayraghavan *et al.*, 1986). The small artificial intron that we used here is a poor splicing substrate and allows for the detection of weak splicing defects contrary to efficiently spliced natural introns (Jacquier *et al.*, 1985; Pikielny and Rosbash, 1985). We subcloned all mutants described above in a splicing reporter gene (Figure 1A) and β-galactosidase activity was assayed to estimate splicing efficiency (Figure 3A).

This spliced reporter allows the detection of splicing efficiency over several orders of magnitude. Splicing activities of mutants range from 2×10^{-4} to 2×10^{-1} compared with the wild-type pre-mRNA. Two substitutions out of three at the branch site have an extremely severe phenotype (A50→T or C). Substitutions C47→T and A49→G exhibit a weak phenotype compared with other substitutions at the same positions. These substitutions still allow U2 snRNA pairing by replacing G::C and A::U pairs by a G::U pair (see insert in Figure 3A). At three positions (A46, T48 and C51) the nature of the substitution has no effect on the intensity of the splicing defect. Surprisingly, the T45→C substitution at the first position of the TACTAAC sequence almost totally abolishes splicing, whereas the other two substitutions have the weakest phenotypes among the TACTAAC mutations. To our knowledge, this mutation has never been analysed but the corresponding sequence occurs in yeast introns. We cannot exclude that the strong phenotype that we observe is specific to our splicing construct.

This collection of mutants was also analysed in a quantitative experiment for pre-mRNA export phenotype (Figure 3B; note that all constructs were first re-cloned and sequenced to eliminate any mutation in the *lacZ* coding sequence that could have occurred during the PCR, see Materials and methods). The wild-type construct has a detectable pre-mRNA export phenotype (~1% of the activity of a reporter gene without an intron; the latter construct corresponds to the highest activity expected for an export phenotype). TACTAAC mutants exhibit various export phenotypes ranging from 5 to 50 times the wild-type reporter activity. These phenotypes cannot be explained by the modifications in the amino acid residues encoded by mutant introns: (i) some of them give products identical to the wild-type (A46→T and C, C47→T, A49→G, T and C) and (ii) TACTAAC mutations combined with a non-functional 5′ splice site exhibit similar β-galactosidase activities (see Table II, and data not shown). Unlike for the splicing defect, the nature of the substitution has a minor effect on the strength of the phenotype at all positions except the last two. For those positions (the A50 at the branch site and C51), the strength of the defect does not correlate with the splicing defect (compare Figure 3A and B). In conclusion, the exhaustive analysis of the TACTAAC point mutations for splicing and pre-mRNA export phenotypes shows that these two assays are independent and measure different limiting parameters. For example, the A49 substitutions have a very weak export phenotype whereas two of them are among the strongest splicing mutants. Similarly, the A50→G substitution at the branch site exhibits a strong export phenotype with a moderate splicing defect.

We were concerned that pre-mRNA export measurements would be impaired by the activation of cryptic sites that would allow novel mRNA species to be in-frame with the β-galactosidase coding sequence. To rule out this hypothesis, primer extension analyses were performed on RNA prepared from yeast cells expressing these various mutants and grown on a galactose-containing medium (Figure 4). Multiple bands corresponding to pre-mRNA and mRNA species were observed as expected for transcripts derived from the *GAL1–CYC1* fusion promoter (Guarente *et al.*, 1982). No cryptic splice site was detected at a level likely to produce significant β-galactosidase. The β-galactosidase activity of weak splicing mutants (>1% of wild-type activity) is correlated with the primer extension analysis in spite of the instability of an out-of-frame mRNA (compare Figures 3A and 4). The A50→G mutant at the branch site exhibits a lariat accumulation as has already observed with an actin reporter gene (Fouser and Friesen, 1986). Other bands are observed under or close to the lariat position. These products do not correspond to cryptic intermediate species since they cannot be debranched *in vitro*, contrary to the lariat product itself (data not shown). No significant variation in the amount of total RNA was observed for the various TACTAAC mutants compared with RNA of the wild-type.

The position preceding the TACTAAC sequence is involved in the recognition process

Among the mutants screened for pre-mRNA export, one was located at the position preceding the TACTAAC box. We made the other two substitutions at that position and assayed them for pre-mRNA export as well as for splicing (Figure 5). Pre-mRNA export analysis (Figure 5A) shows that the four nucleotides at this position lead to different amounts of pre-mRNA export, the G having the strongest phenotype and the T an almost undetectable one (note that C and T mutants encode identical β-galactosidase fused proteins). We conclude that this position contributes to the identification of the branch site region. In contrast to mutations located within the TACTAAC box, the splicing efficiency of the four pre-mRNAs differing at this position (Figure 5B) is correlated with the pre-mRNA nuclear retention. Primer extension analyses were made with RNA prepared from cells transformed with pre-mRNA export or splicing reporter plasmids (Figure 5C). Profiles of cDNA obtained with the two types of reporter genes are similar for a given nucleotide. Slight differences in the relative amount of spliced mRNAs are observed for pre-mRNA export and splicing constructs. These differences most probably reflect a higher rate of degrada-

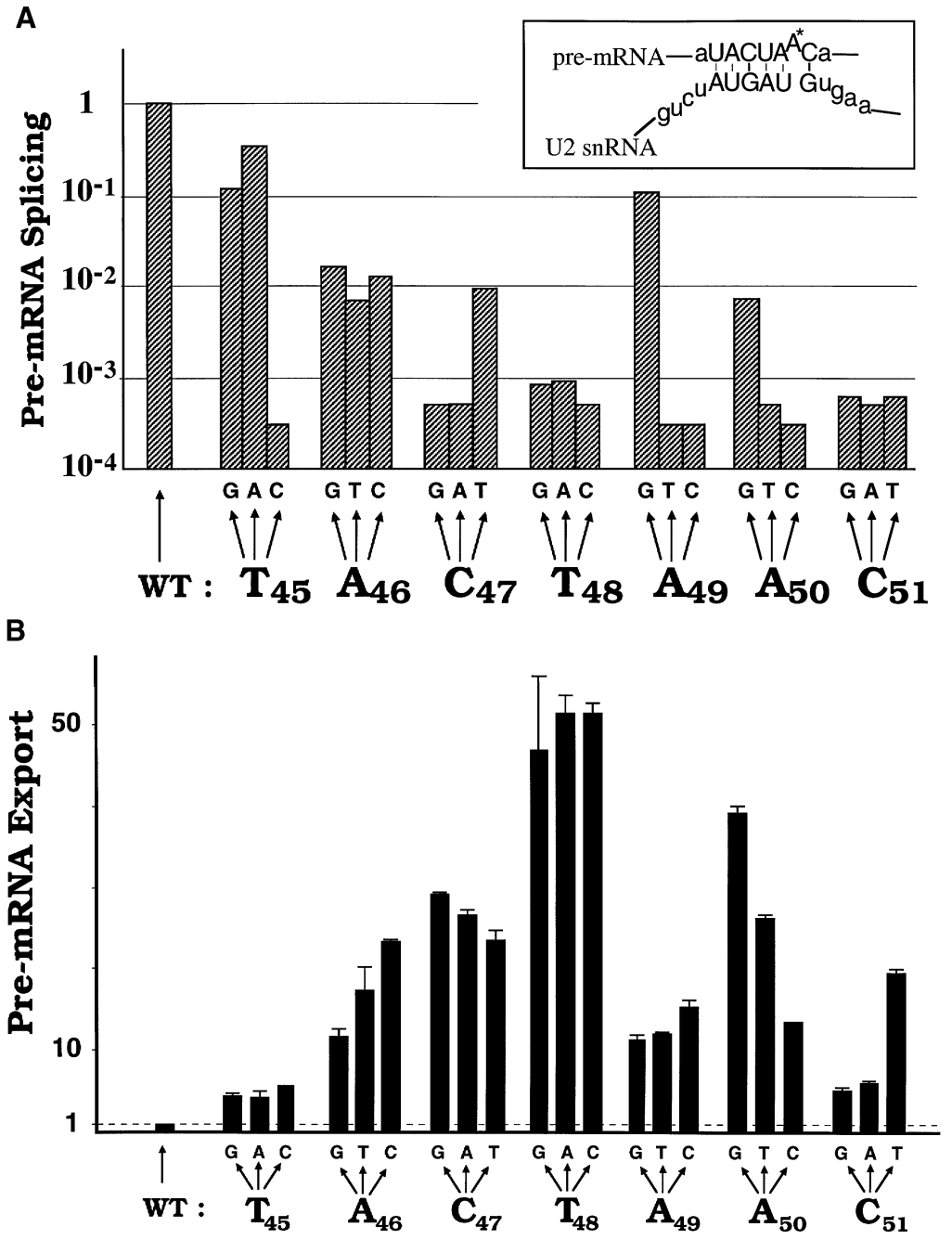

Fig. 3. Splicing efficiency and pre-mRNA export phenotype of TACTAAC box substitutions. The wild-type sequence is presented in bold under each panel and substitution is noted under the histogram. (**A**) Splicing efficiency. β-Galactosidase activity was measured under steady-state conditions (14 h after galactose induction). The splicing efficiency is expressed as the ratio to the β-galactosidase activity of the wild-type reporter and is presented on a logarithmic scale. Experiments were performed on two independent transformants. Experimental variations were within a 10% range. In the insert, the base pairing occurring in the spliceosome between the pre-mRNA and the U2 snRNA is indicated. The branch site is labelled (*). (**B**) Pre-mRNA export. β-Galactosidase activity was measured 3 h after galactose induction. Pre-mRNA export is expressed as the percentage of the β-galactosidase activity of the wild-type reporter and is presented on a linear scale. Similar relative values were observed 1 or 2 h after galactose induction. Standard deviations for two independent transformants are indicated.

tion for cytoplasmic nonsense codon-containing translated RNAs by the *UPF1*-dependent degradation pathway (compare lanes 1 and 3 with 5 and 7, respectively; Long *et al.*, 1995). The A44→T mutant has a high splicing efficiency compared with other mutants and it accumulates much less pre-mRNA (Figures 4 and 5C). This result suggests that the tTACTAAC sequence might be the one that is best recognized by the splicing machinery.

Yeast natural introns exhibit a strong bias at the nucleotide preceding the TACTAAC box

The strong influence of the nature of the nucleotide preceding the TACTAAC sequence in the recognition of the branch site region prompted us to investigate the representation of the four nucleotides in the vicinity of the branch site region in natural yeast introns. Due to the availability of the complete sequence of the yeast genome,

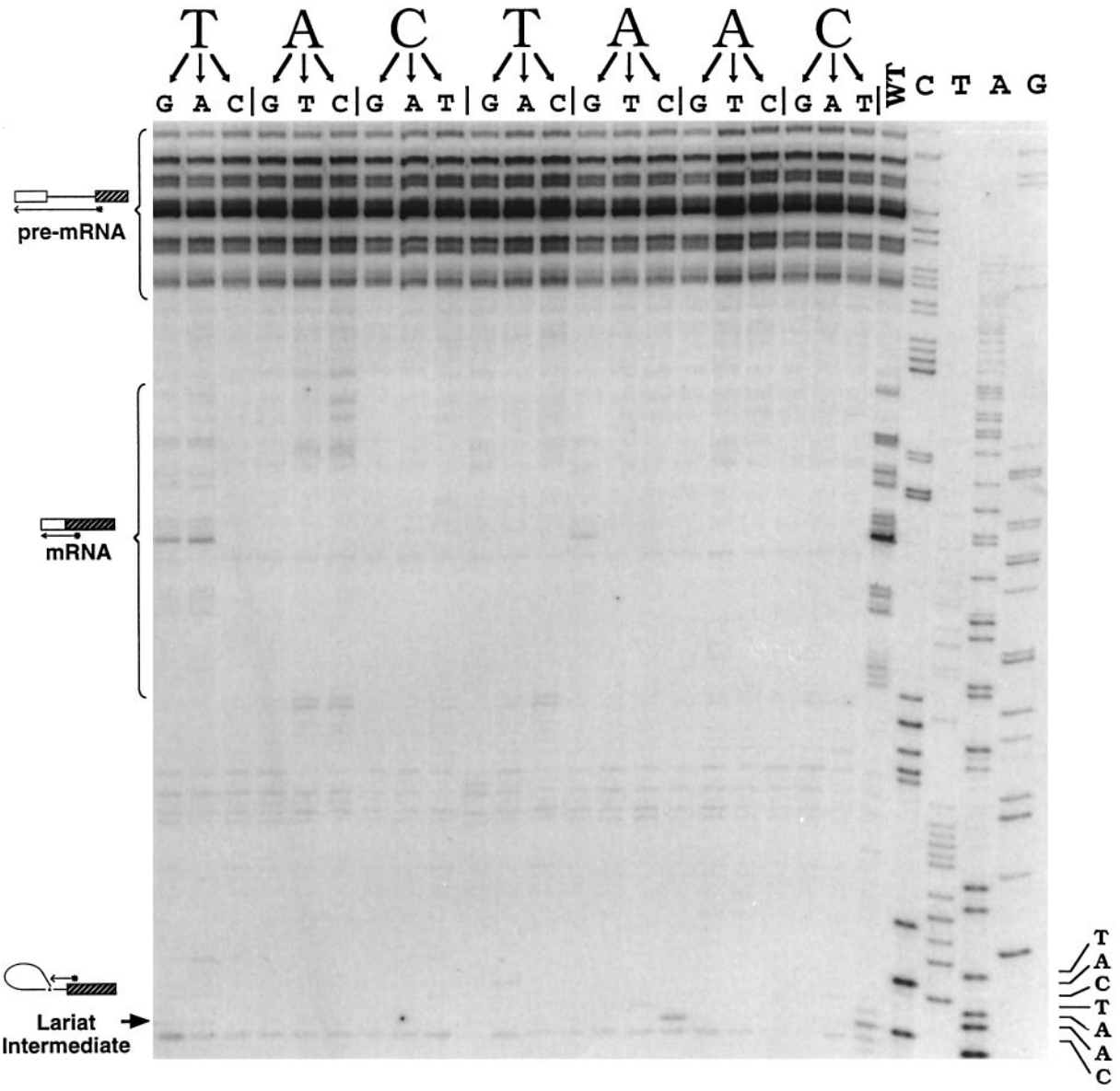

Fig. 4. RNA analysis of pre-mRNA export mutants. Primer extension analyses were performed for TACTAAC box mutants in pre-mRNA export constructs. Note that for these constructs, spliced mRNAs are out-of-frame for the β-galactosidase-coding sequence (see Figure 1A). The nature of substitution is indicated at the top of the lanes. On the right is the DNA sequence of the wild-type plasmid and on the left the various identified RNA species are indicated [the multiple bands for mRNAs and pre-mRNAs are due to multiple starts of transcription of the reporter gene (Legrain and Rosbash, 1989)].

we were able to sort out most, if not all *S.cerevisiae* introns. We have constituted a family of 224 putative introns. Of these, 159 correspond to known genes, while 65 are found associated with the newly identified ORFs. It should be stressed that the existence of any given *bona fide* intron awaits a specific experimental demonstration. However, based upon several criteria such as (i) the conserved sequences, (ii) the position in the ORF and (iii) the homology of the deduced protein sequences to proteins in other organisms, most of these introns can be considered real. The description of the complete family of the yeast introns and their various characteristics will be reported elsewhere.

We measured the occurrence of each nucleotide at the positions corresponding to the randomized sequence in our genetic screen (Figure 1A), i.e. at the four positions preceding the TACTAAC box and the one following it. These frequencies were compared with those obtained for all genomic TACTAAC sequences (1496 sequences, Figure 6). Two positions, the ones immediately preceding and following the TACTAAC sequence, show a strong difference between intronic sequences and the complete set of genomic sequences. It was already known that an A residue often follows the TACTAAC intronic sequence (63.4% of cases compared with 39.6%) (Parker *et al.*, 1987). At this position, the C residue is clearly excluded (3.1% compared with 16.2%). At the position preceding the TACTAAC sequence in introns, the frequencies of Ts and Cs are 59.8 and 5.4%, respectively (compared with 34.6 and 22.7%, respectively, in genomic TACTAAC sequences). Out of the 12 introns that contain a C at this position, six are found in new uncharacterized genes. In

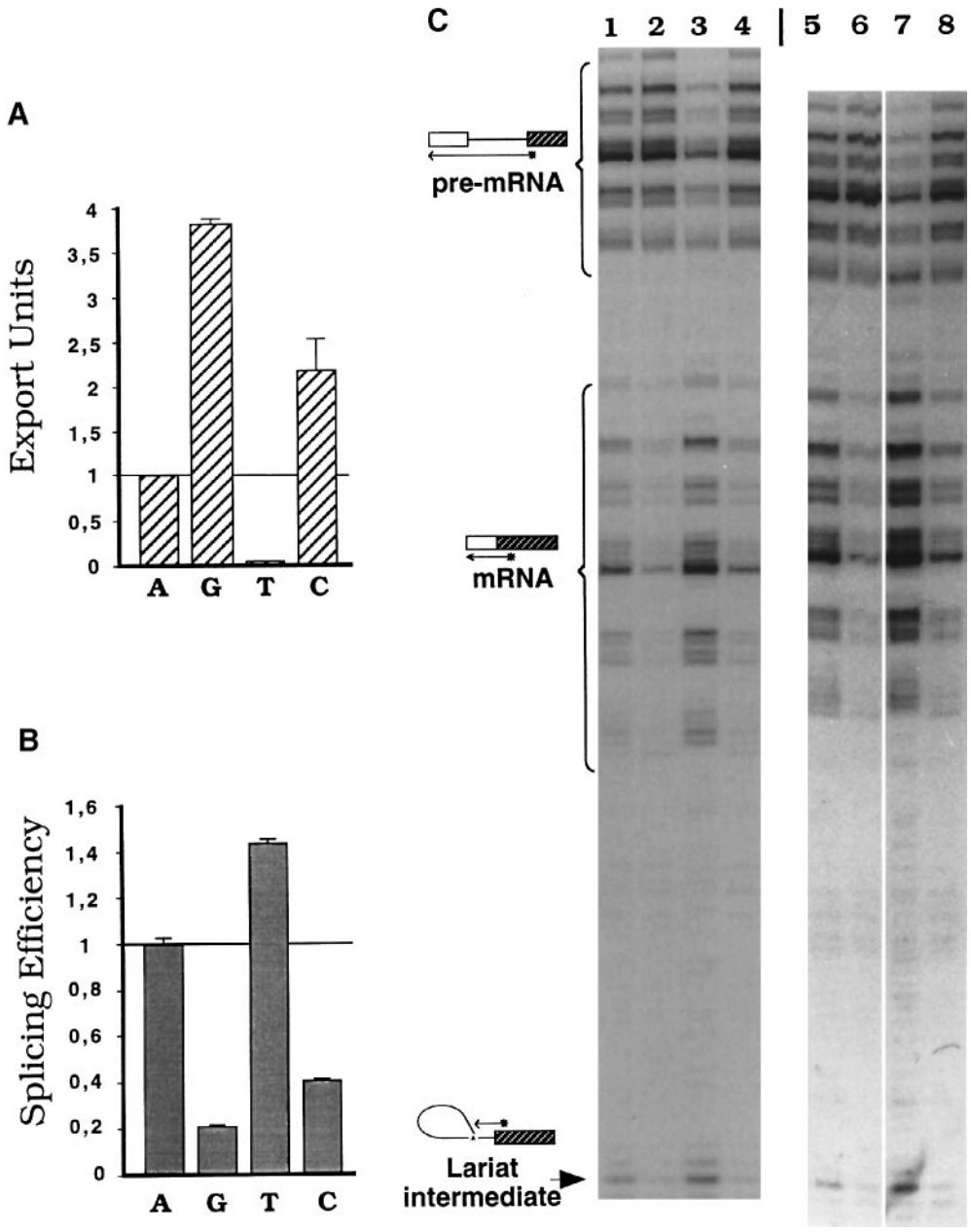

Fig. 5. The nucleotide preceding the TACTAAC box contributes to the efficacy of branch site region recognition. The effect of the nucleotide preceding the TACTAAC box on pre-mRNA export (**A**) or pre-mRNA splicing (**B**). The nature of the nucleotide preceding the TACTAAC sequence is indicated. Assays are performed as in Figure 3B and A, respectively, and results are expressed as the ratio of β-galactosidase units to the wild-type sequence (an A residue preceding the TACTAAC box). (**C**) Primer extension analyses of total yeast RNA from MG353-46D cells transformed with pre-mRNA export (1–4) or splicing (5–8) constructs differing at the nucleotide preceding the TACTAAC box (1 and 5, A; 2 and 6, G; 3 and 7, T; and 4 and 8, C). Primer extensions are performed as in Figure 4.

addition, five of them are not located in the 5′ region of the RNA where the yeast *S.cerevisiae* introns usually are. Few introns (20) are also found with a G residue at this position, but this nucleotide is naturally less represented in the total set of genomic sequences (13.8%). In addition, out of the 13 identified introns which do not contain the perfect TACTAAC consensus sequence but a variation of it (CACTAAC, GACTAAC, AACTAAC, TGCTAAC and AATTAAC), 12 contain a T residue and one an A residue at the position preceding the consensus sequence.

In conclusion, in natural introns, C and G residues occur very rarely and Ts are found preferentially at the position preceding the TACTAAC sequence. This is in complete agreement with our experimental results on nuclear export and splicing of pre-mRNAs differing at this position.

The Mud2 protein plays a role in the recognition of the nucleotide preceding the TACTAAC box

We have shown that mutations at a given position in the TACTAAC box exhibit a different phenotype for pre-mRNA export or splicing. We tested a splicing factor, Prp16p, that plays a role late in the splicing process and could affect splicing and not pre-mRNA export (Burgess *et al.*, 1990). The *prp16-101* mutation exhibits an important suppressor effect on several mutations in the TACTAAC

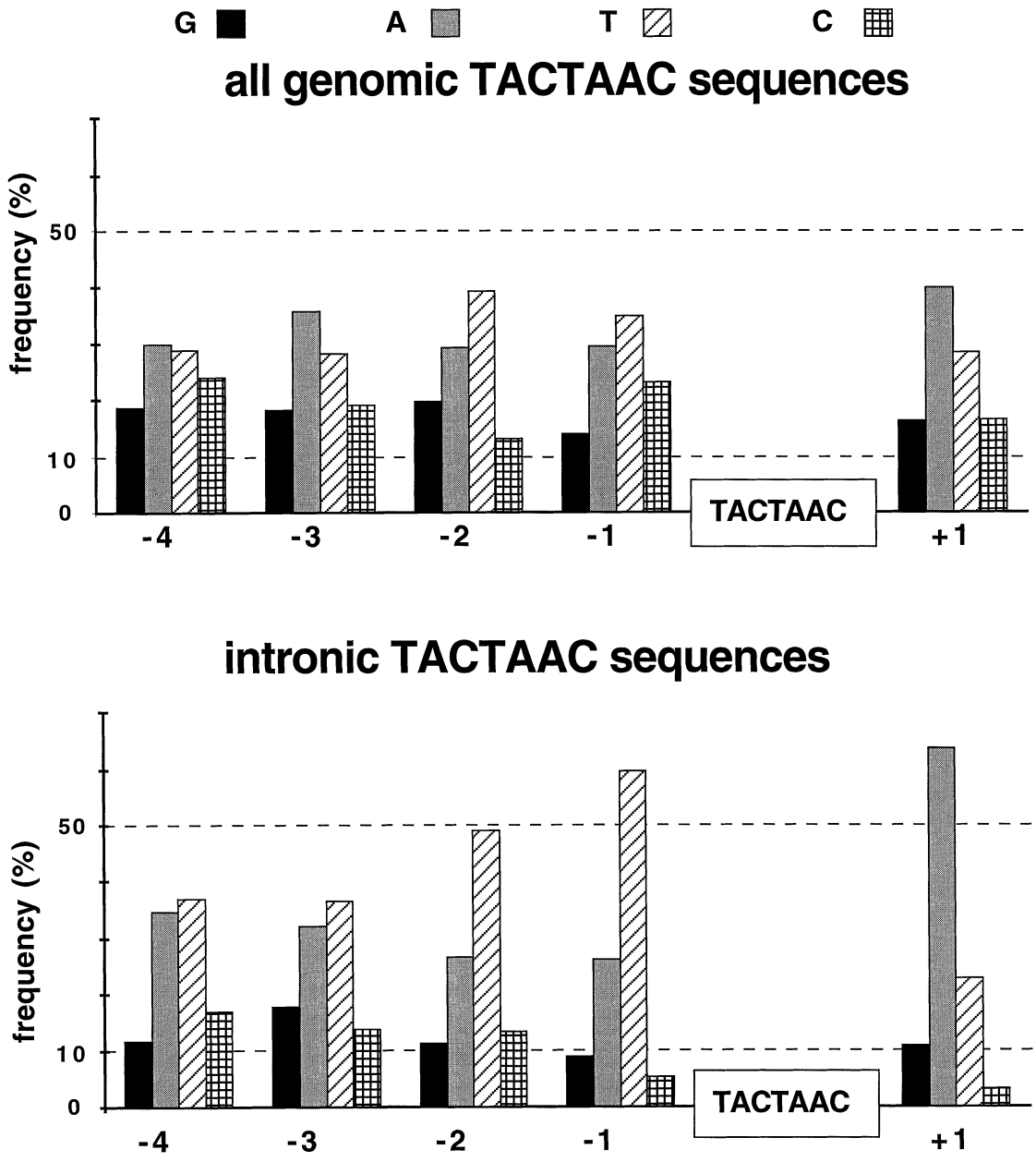

Fig. 6. Occurrence of the four nucleotides at positions surrounding yeast TACTAAC sequences. The percentage of G, A, T and C occurrences at positions –4, –3, –2, –1 and +1 relative to TACTAAC sequences are presented for all genomic TACTAAC sequences (1496 sequences, upper panel) or for TACTAAC sequences found in introns (224 sequences, lower panel).

box of the actin intron, including the branch site (Burgess and Guthrie, 1993). Wild-type 46D yeast strains harbouring the different branch site mutations in a splicing substrate were transformed with the pSE358 vector or with the pSE358 vector containing either the *PRP16* gene or its mutant *prp16-101* allele (a gift of C.Guthrie, UCSF, San Fransisco). We observed absolutely no effect of this mutation on the splicing efficiency of the reporter gene. Similarly, there was no effect of the *prp16-101* mutation on a pre-mRNA export reporter (data not shown). To our knowledge, the *prp16-101* suppressor phenotype has been described only for the actin intron mutants, and its suppressor activity could be specific for this intron. Alternatively, branch site mutations affect our synthetic intron differently compared with the actin intron, and this effect cannot be suppressed by the *prp16-101* mutation. In this regard, the absence of suppression could be related to the role proposed for the *prp16* suppressor mutants: they suppress the degradation of aberrantly branched intermediate lariats. In a *PRP16* wild-type strain, this degradation pathway leads to a decrease in the total amount of the mutant RNA (pre-mRNA, intermediate and spliced RNA) (Burgess and Guthrie, 1993). In our assay using a synthetic intron, we do not detect a significant decrease in the total amount of reporter RNA.

We then tested for *trans*-acting factors implicated early in spliceosome formation which could be involved in the recognition of the nucleotide preceding the TACTAAC box. Prp9p, Prp11p and Prp21p are associated with the U2 snRNP and participate in the formation of the pre-

Table I. Pre-mRNA export in a strain disrupted for the *MUD2* gene

Mutants	Wild-type strain	*MUD2::LEU2* strain	Ratio Δ/wt
aTACTAAC	7.3 ± 0.2	27.2 ± 1.3	3.72
cTACTAAC	9.1 ± 0.9	14.0 ± 2.8	1.53
gTACTAAC	25.8 ± 0.9	39.6 ± 2.3	1.54
tTACTAAC	0.26 ± 0.0	2.17 ± 0.0	8.47
aT**C**CTAAC	59.7 ± 11.7	60.1 ± 7.9	1.01
aTAA**A**TAAC	71.4 ± 13.7	79.1 ± 3.7	1.11
aTACTA**G**C	89.8 ± 0.6	91.0 ± 28.5	1.01
aTACTAA**T**	109.5 ± 8.2	104.8 ± 2.5	0.96

Pre-mRNA export was measured 2 h after induction of transcription and is expressed in β-galactosidase units. Two independent transformants were assayed. The ratio Δ/wt reflects the enhancement of pre-mRNA export in the *MUD2*-disrupted strain.

spliceosome (Hodges and Beggs, 1994). We overproduced each of these factors in a wild-type strain and we measured pre-mRNA export and splicing for constructs with any of the four nucleotides at the position preceding the TACTAAC box. Overexpression of these factors had no effect on splicing or on pre-mRNA export, thus suggesting that their amount is not limiting (data not shown).

The Mud2 protein is a non-essential splicing factor implicated early in spliceosome assembly (Abovich *et al.*, 1994). We compared pre-mRNA export for two isogenic strains differing solely in the disruption of the *MUD2* gene (Table I). In the disrupted strain, pre-mRNA export is enhanced significantly for pre-mRNA with any of the four nucleotides at the position preceding the TACTAAC box. The effect is more pronounced for pre-mRNA with an A or a T residue (3.7- and 8.5-fold increase, respectively compared with 1.5). However, pre-mRNA export of the T-containing substrate is still lower than for other substrates in the disrupted strain, suggesting that Mud2p is not the sole recognition factor. In contrast, the absence of Mud2p has no detectable effect on the export of pre-mRNA mutated inside the TACTAAC box (<10% increase, Table I). In addition, the *MUD2* deletion also has a marginal effect on splicing of this synthetic intron (data not shown, Abovich *et al.*, 1994).

The Mud2 protein acts independently of 5′ splice site recognition

MUD2 was isolated initially in a synthetic lethal screen with a U1 snRNA mutant (Abovich *et al.*, 1994). The observed effect of the *MUD2* disruption for the position preceding the TACTAAC box could be an indirect effect of a reduced binding to the 5′ splice site region by an altered U1 snRNP. To rule out this possibility, we measured the translation of pre-mRNA TACTAAC mutants (including the preceding position) in conjunction with a G1→A mutation at the 5′ splice site. In this context, the 5′ splice site is very poorly recognized by the splicing machinery, and β-galactosidase values are close to or above 50% of that found with the reporter without an intron (Figure 1B, Table II and data not shown). The comparison of translational values for pre-mRNA with the various nucleotides at the position preceding the TACTAAC reveals that **c**TACTAAC and **g**TACTAAC reporters are expressed at a higher rate than **a**TACTAAC or **t**TACTAAC substrates in a wild-type strain (57 and 63% compared with 31 and 48%, Table II). When the same experiment is performed in a *MUD2::LEU2*-disrupted strain, pre-mRNA export occurs at about the same rate for all four substrates (~65%, Table II). In fact, the *MUD2* disruption increases by a factor of 2 for the export of **t**TACTAAC substrate and by a factor of 1.5 for the **a**TACTAAC substrate. Once again, the effect of the *MUD2* disruption is specific for this intronic position, since experiments performed in parallel for TACTAAC mutations show no influence of the absence of Mud2p on the export of pre-mRNA mutants (Table II). From these results, we conclude that the Mud2 protein contributes significantly to the recognition of the branch site region via the nucleotide preceding the TACTAAC box, and that this interaction is not mediated through the recruitment of the 5′ splice site by the splicing machinery.

Table II. Pre-mRNA export in the absence of a functional 5′ splice site

Mutant	Wild-type strain	*MUD2::LEU2* strain	Ratio Δ/wt
aTACTAAC	47.8 ± 11.5	67.4 ± 1.8	1.41
cTACTAAC	57.0 ± 3.6	66.4 ± 2.0	1.16
gTACTAAC	63.0 ± 4.4	64.0 ± 1.4	1.02
tTACTAAC	30.9 ± 5.5	61.0 ± 6.8	1.97
aT**C**CTAAC	42.6 ± 2.6	47.3 ± 4.0	1.11
aTA**A**TAAC	65.9 ± 6.0	65.6 ± 6.4	1.00
aTAC**G**AAC	48.6 ± 4.0	46.6 ± 1.0	0.96
aTAC**A**AAC	54.3 ± 0.1	53.1 ± 3.4	0.98
aTAC**C**AAC	53.9 ± 0.5	54.3 ± 1.1	1.01
aTACTA**G**C	78.5 ± 4.0	68.3 ± 0.8	0.87
aTACTAA**T**	56.8 ± 3.0	52.5 ± 0.0	0.92

All constructs contain a G1→A mutation in the 5′ splice site. Pre-mRNA export was measured 2 h after induction and expressed as a percentage of the β-galactosidase activity of the reporter gene without an intron. Two independent transformants were assayed. The ratio Δ/wt reflects the enhancement of pre-mRNA export in the *MUD2*-disrupted strain.

Discussion

Nuclear pre-mRNA splicing is a major post-transcriptional modification that occurs in eukaryotic cells. It implies the specific recognition and excision of introns from the primary transcripts. The chemistry of the removal of introns has been studied extensively, and many key factors and parameters of the catalytic steps have been identified (Madhani and Guthrie, 1994; Ares and Weiser, 1995). In contrast, the process by which *in vivo* identification of intronic sequences in the primary transcripts proceeds is poorly understood: most studies have been performed *in vitro*, many proteins are involved and few of them seem to have a unique and essential role (Krämer, 1996). In addition, although it is firmly established that the conserved 5′ splice site and the branch site regions are involved in the intron recognition process, precise requirements at this step have not been identified. The base pairing of the 5′ splice site region with the U1 snRNA is an early event during the formation of the spliceosome (Rosbash and Séraphin, 1991), but the requirement for the branch site region–U2 snRNA base pairing is questionable. Available data suggest a predominant role in protein–protein and protein–RNA interactions

during the initial steps of spliceosome formation (Fu, 1995).

Here, we present analyses of many different point mutations for pre-mRNA splicing and export. The single point mutations located in the 5′ splice site exhibit various pre-mRNA export phenotypes. The pre-mRNA export assay performed on randomly made mutants in the branch site region revealed that single mutations at any position in the TACTAAC box lead to an enhanced pre-mRNA export. Hence, the pre-mRNA export assay is very sensitive and allows the detection of subtle modifications in primary transcripts. The two mutants on the 3′ splice site exhibit no pre-mRNA export, whereas they are blocked in splicing. The G1→A and the G5→A mutants exhibit a similar pre-mRNA export phenotype contrasted to their different splicing defects (Jacquier *et al.*, 1985; Parker and Guthrie, 1985). Thus, pre-mRNA splicing and export can be measured independently, allowing the distinction between intron recogniton factors and catalytic ones. The screening procedure that we have developed can now be used to assay pre-mRNA export systematically for a collection of mutants, such as a collection of disrupted strains.

Our complete analysis of TACTAAC mutations for pre-mRNA splicing can be compared with previous studies of yeast branch point region mutants. Using an actin intron reporter, Fouser and Friesen (1986) made a complete set of a single substitution per nucleotide in the TACTAAC sequence and Vijayraghavan *et al.* (1986) made three additional substitutions (ta**C**taac→A and T; tac**T**aac→A). Weak phenotypes are observed for mutants at the two first positions in the TACTAAC sequence (T→G and A→C); at the third position, the T substitution has a weaker phenotype than the A substitution; at the fourth position, the T→A substitution has a strong phenotype; at the branch site the A→G substitution leads to accumulation of intermediate lariat and the A→C substitution is among the strongest mutations tested in the TACTAAC. In addition, two mutations in the TACTAAC sequence were made in a *RP51A* reporter (t**A**ctaac→C and ta**C**taac→T) and exhibited a weak splicing defect (Jacquier *et al.*, 1985). All these results are similar to ours obtained with the small artificial intron. However, three substitutions in the actin mutants (tac**T**aac→C, tact**A**ac→C and tactaa**C**→G) have a rather weak phenotype; these substitutions in the present study have a strong phenotype and this discrepancy could reflect intron specificities. Taken altogether, the results show that specific mutations in the TACTAAC affect the splicing efficiency of the transcript differently. These effects can be correlated with the base pairing with the U2 snRNA (insert in Figure 3A). In this regard, the branch site region consensus in mammalian cells (YNCURAY, see Krämer, 1996) corresponds to the most acceptable nucleotides in the splicing assay. Contrary to U1 snRNA base pairing with the 5′ splice site, the U2 snRNA base pairing with the branch site region is necessary during the formation of the active catalytic site for splicing (Madhani and Guthrie, 1994). The conserved sequence in mammals could be considered a minimal catalytic requirement selected during evolution. The various pre-mRNA export phenotypes for most TACTAAC mutants do not correlate with their splicing defects. This strongly suggests that U2 snRNA–pre-mRNA base pairing is not required during early steps of intron recognition.

The position preceding the TACTAAC is important for pre-mRNA nuclear retention. An A or a T residue at this position allows a much better pre-mRNA recognition than a C or G. Contrary to the TACTAAC mutants, the splicing efficacy of the mutants at the preceding position is correlated directly to the pre-mRNA nuclear retention, suggesting that this position plays only a minor role later on in the splicing process. This finding is backed up by the analysis of natural introns: G and C residues are almost absent at the position preceding intronic TACTAAC boxes, and a positive selection of T residues is observed. This selection cannot be explained by an optimal pairing with the U2 snRNA. To our knowledge, it is the first identification in yeast of a conserved nucleotide that plays a role only during early steps of spliceosome assembly.

Previous studies performed *in vitro* demonstrated that the commitment to the spliceosome assembly pathway implicates the U1 snRNP and additional protein factors including Mud2p (Legrain *et al.*, 1988; Séraphin and Rosbash, 1989; Abovich *et al.*, 1994; Colot *et al.*, 1996). Here we describe an *in vivo* phenotype for the *MUD2* disruption: it increases pre-mRNA export and this effect is sensitive to the nature of the nucleotide at the position preceding the TACTAAC. Moreover, the Mud2p effect can be observed even on a pre-mRNA with a mutated 5′ splice site, showing that the interaction between the branch site region and the Mud2p protein occurs independently of an efficient 5′ splice site recognition. Mutations located within the TACTAAC sequence are not sensitive to the Mud2p depletion. This strongly suggests that Mud2p does not play a role in splicing catalysis *per se* but that it is required for efficient intron recognition by the splicing machinery. Many factors contribute to this process *in vivo*, rendering the Mud2p function dispensable in a wild-type strain. However, as it has been shown previously, the *MUD2* disruption is lethal when assayed in mutants containing a destabilized U1 snRNA or a heat-sensitive Prp11p that are early factors during the spliceosome assembly (Abovich *et al.*, 1994).

In higher eukaryotes, 5′ and 3′ splice site complexes have been identified independently. The 3′ splice site complexes comprise the U2 snRNP and several proteins including PTB, PSF and U2AF bound to the polypyrimidine track (reviewed in Krämer, 1996). The connection between 5′ splice site and 3′ splice site complexes occurs later on. A different situation occurs in yeast. The definition of the 3′ splice site is a late event that occurs after the first cleavage reaction. Commitment to spliceosome assembly simultaneously requires a 5′ splice site and a branch site region (Séraphin and Rosbash, 1989), even if an independent binding to each region is not ruled out (Séraphin and Rosbash, 1991). The Mud2 protein was characterized *in vitro* and proposed to be the initial branch point region-binding factor (Abovich *et al.*, 1994). We now demonstrate that the Mud2 protein contributes *in vivo* to the recognition of the branch site region via the nucleotide preceding the TACTAAC sequence, even in the absence of a functional 5′ splice site. Thus, independent binding to conserved intronic sequences appears to be a feature common to yeast and mammalian spliceosome assembly. This finding can be taken together with the fact

that Mud2p is proposed to be the U2AF^{65} yeast homologue (Abovich *et al.*, 1994). Very recently, it has been shown that the RS domain of this human splicing factor interacts with the branch point region and promotes U2 snRNA base pairing (Valcàrcel *et al.*, 1996). The Mud2 protein could exert a similar function in yeast. However, Mud2p is a dispensable factor, suggesting that it interacts with the branch point region in association with other factors. Genetic interactions were found between *MUD2*, *MUD1* and *PRP11* (Abovich *et al.*, 1994), but the precise molecular links between Mud2p and the U1 and U2 snRNPs remain to be unravelled. Using a two-hybrid assay, we found that the previously described Mud2p–Prp11p interaction exhibits a comparable or lower β-galactosidase activity than other weak interactions such as Mud2p–Prp9p (data not shown). In addition, neither Prp9p nor Prp11p were found among partners in a two-hybrid screen with Mud2p. In contrast, a novel factor was identified in such a screen (M.Fromont, J.-C.Rain and P.Legrain, submitted). This factor exhibits a strong homology with SF1, a recently cloned human splicing factor (Arning *et al.*, 1996). The bridge between U1 and U2 snRNP during *in vivo* spliceosome assembly requires a Mud2p-containing complex that we are currently investigating.

Materials and methods

Strains and plasmids

The yeast strain MGD353-46D (*Mat***a** *ura3-52 trp1–289 leu2-3, -112 ade2 arg4*) was used. The disruption of the *MUD2* gene was made by the one-step procedure (Rothstein, 1983) using a fragment of a pTZ18 MUD2::LEU2 plasmid provided by M.Rosbash (Brandeis University, Waltham). The disruption was confirmed by Southern blot analysis. Yeast transformations were performed using the standard protocol (Transy and Legrain, 1995). Reporter plasmids used for pre-mRNA export and splicing assays are derived from the pPlint plasmid family (Legrain and Rosbash, 1989). The 2 µm-based plasmids pYembl31 (Baldari and Cesareni, 1985) and pFL46S (Bonneaud *et al.*, 1991) were used for protein overexpression. *Escherichia coli* TG1 strain was used for every step of subcloning and sequencing work.

Construction of the pre-mRNA mutant library

We synthesized three partially randomized oligonucleotides:
JCR 10 5′-GGTTAACGTCGACACCGTGTTTTTGATA**i**C**i**j**i**j**C**i**j**j**C**j**G-GCCTTTTAATA-3′ JC11 5′-GGTTAACGTCGACACCGTGTTTTT-GATA**T**k**T**j**T**j**k**T**j**j**k**j**GGCCTTTTAATA-3′ JCR12 5′-GGTTAACGTC-GACACCGTGT-TTTTGATA**i**k**i**A**i**A**k**i**AA**k**AGGCCTTTTAATA-3′ where **i**, **j** and **k** are mixed phosphoramidites with a proportion of 7/8 thymidine, adenine and cytidine, respectively and 1/24 of each of the three other phosphoramidites. This degeneracy was chosen to maximize single mutants. All combinations of double mutants can be obtained. JCR10, JCR11 or JCR12 in combination with JCR1 (5′-CGGCGCT-CCACAGTTTCGGG-3′) whose sequence is located 900 nucleotides downstream in the *lacZ* gene on the reverse strand were used for PCR on the Nde°Acc° plasmid (Legrain and Rosbash, 1989). PCR fragments were blunted with T4 DNA polymerase, digested with *Sal*I and *Cla*I and ligated with pPlintΔClaI (pPlintΔClaI was derived from pPlint by partial digestion with *Cla*I to remove the site in the 2 µm cassette). TG1 cells were electroporated and 10 clones of each library were sequenced; no bias was observed in the number and nature of mutations. Colonies were pooled and three DNA libraries were prepared from a total of 1350 *E.coli* clones.

Screening of the yeast library

Yeast MG353-46D were transformed with 100 ng of each DNA library. We streaked 300 independent clones for each yeast library on minimal medium supplemented with glucose for storage and on a nylon filter (Amersham N) for X-gal assay. The filters were placed on a plate containing the same minimal medium supplemented with glucose 0.05%, glycerol 2% and lactate 2% (–ura GGL plate). Several controls were added to each filter, i.e. yeast cells transformed with known contructions: Nde°Acc°, SD5, ΔD, ΔC (Legrain and Rosbash, 1989). After 48 h of growth, filters were placed on plates supplemented with galactose (2%) for 1 h to induce the reporter gene. Filters were then processed for X-gal coloration as described previously (Transy and Legrain, 1995). Blue colonies were recovered and a second screening was realized using ~105 yeast cells spotted on a Nylon filter in duplicate. In this second assay, X-gal coloration was very reproducible. A few white clones were discarded at this step. Yeast DNA was prepared and plasmids were rescued in *E.coli* by electroporation.

Construction of mutants

Mutants of the 5′ splice site were made by PCR with specific oligonucleotides (e.g. JCR15 for the G5→A mutant: 5′-ACCGGATCCATGG-AACGGTAT**A**TTAAT-3′) and the JCR1 oligonucleotide on the Nde°Acc° template. Similarly, mutants of the 3′ splice site were obtained by PCR using upstream JCR4 (5′-ATGACCGGATCCATGGAATG-3′) and downstream JCR5 or JCR6 (5′-TCCGGATCCCGTATG-AC**G**TATTAAAAG-3′; 5′-TCCGGATCCCGTATGACC**C**ATTAAAAG-3′) on the Nde°Acc° template. PCR fragments were digested with *Bam*HI and ligated with SD5ΔClaI. New mutants in the TACTAAC region were obtained by PCR using a specific oligonucleotide such as JCR19 for the T45→A mutant (5′-GGTTAACGTCGACACCGTGTTTTTGATAT-CTAA**A**ACTAACAGGC-3′) and JCR1 on the Nde°Acc° template. PCR fragments were then cloned as for library PCR fragments. To avoid any additional PCR mutation located in the *lacZ* gene, we subcloned all single mutants in SD5ΔClaI digested with *Bam*HI. After sequence verification, these mutants were used for all further experiments.

The splicing constructs were derived from the pre-mRNA constructs by modificaton of exon1 (Figure 1A). We subcloned the exon 1-containing *Xba*I–*Sal*I fragment of the Sty° construct (Legrain and Rosbash, 1989) in the SK$^+$ vector (Stratagene). The resulting plasmid was digested with *Bam*HI, blunted and self-ligated. In all pre-mRNA export TACTAAC region mutants, the initial *Xba*I–*Sal*I fragment was replaced by the new one making an ORF coding for β-galactosidase only for the spliced mRNA.

Double mutants (at the 5′ splice site and in the TACTAAC region) were obtained following the same procedure but starting with the *Xba*I–*Sal*I fragment of the G1→A mutant.

RNA analysis

Yeast total RNA was prepared as described previously (Pikielny and Rosbash, 1986) with several modifications. Cells were grown in minimal medium (–ura GGL) supplemented with 2% galactose to an optical density of 0.8–1.2. Then 10 ml of cells were collected by centrifugation and were resuspended in 500 µl of RNA extraction buffer [0.1 M Tris–HCl (pH 7.4), 0.1 M LiCl, 0.1 mM EDTA] in an Eppendorf tube. Siliconized glass beads (300 µl) were added and cells were lysed by vortexing for 4 min in the cold. Ten µl of a 10% SDS solution were then added and tubes were vortexed for 15 s, 500 µl of phenol were added and tubes vortexed again three times for 15 s. After centrifugation, the supernatant was re-extracted twice with phenol–chloroform (1:1), and the RNA was ethanol precipitated, washed and resuspended in 10 µl of TE. Primer extensions were performed as described previously (Chanfreau *et al.*, 1994) with the oligonucleotide DT 320 (5′-CACCA-GTGAGACGGGC-3′) which is complementary to the beginning of exon 2.

Acknowledgements

This publication is dedicated to the memory of A.Kalogeropoulos who initiated systematic searches for yeast introns in databases and who contributed to many fruitful discussions during the early stages of this work. We thank N.Abovich, M.Rosbash and C.Guthrie for providing ptz18 MUD2 and PRP16 plasmids. We acknowledge C.Marck's help in sorting the collection of genomic TACTAAC sequences. We are grateful to A.Jacquier for numerous and helpful discussions. We thank M.Meth, M.Fromont, A.Jacquier and G.Rotondo for critical reading of the manuscript. This work was supported in part by the Fondation de la Recherche Médicale and by the CEC contract CHRX-CT94-0677. J-C.R. is Allocataire de Recherche (MESR-Université Paris XI).

References

Abovich,N., Liao,X.C. and 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.

Ares,M.J. and Weiser,B. (1995) Rearrangement of snRNA structure during assembly and function of the spliceosome. *Prog. Nucleic Acid Res. Mol. Biol.*, **50**, 131–159.

Arning,S., Grüter,P., Bilbe,G. and Krämer,A. (1996) Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. *RNA*, **2**, 794–810.

Baldari,C. and Cesareni,G. (1985) Plasmids pEMBLY: new single-stranded shuttle vectors for the recovery and analysis of yeast DNA sequences. *Gene*, **35**, 27–32.

Behrens,S.E., Galisson,F., Legrain,P. and Lührmann,R. (1993) Evidence that the 60-kDa protein of 17S U2 small nuclear ribonucleoprotein is immunologically and functionally related to the yeast PRP9 splicing factor and is required for the efficient formation of prespliceosomes. *Proc. Natl Acad. Sci. USA*, **90**, 8229–8233.

Bennett,M. and Reed,R. (1993) Correspondence between a mammalian spliceosome component and an essential yeast splicing factor. *Science*, **262**, 105–108.

Bonneaud,N., Ozier-Kalogeropoulos,O., Li,G.Y., Labouesse,M., Minvielle-Sebastia,L. and Lacroute,F. (1991) A family of low and high copy replicative, integrative and single-stranded *S.cerevisiae*/*E. coli* shuttle vectors. *Yeast*, **7**, 609–615.

Brosi,R., Gröning,K., Behrens,S.E., Lührmann,R. and Krämer,A. (1993a) Interaction of mammalian splicing factor SF3a with U2 snRNP and relation of its 60-kD subunit to yeast PRP9. *Science*, **262**, 102–105.

Brosi,R., Hauri,H.P. and Krämer,A. (1993b) Separation of splicing factor SF3 into two components and purification of SF3a activity. *J. Biol. Chem.*, **268**, 17640–17646.

Burgess,S.M. and Guthrie,C. (1993) A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. *Cell*, **73**, 1377–1391.

Burgess,S., Couto,J.R. and Guthrie,C. (1990) A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. *Cell*, **60**, 705–717.

Chanfreau,G., Legrain,P., Dujon,B. and Jacquier,A. (1994) Interaction between the first and last nucleotides of pre-mRNA introns is a determinant of 3′ splice site selection in *S.cerevisiae*. *Nucleic Acids Res.*, **22**, 1981–1987.

Chapon,C. and Legrain,P. (1992) A novel gene, *spp91-1*, suppresses the splicing defect and the pre-messenger-RNA nuclear export in the *prp9-1* mutant. *EMBO J.*, **11**, 3279–3288.

Colot,H.V., Stutz,F. and Rosbash,M. (1996) The yeast splicing factor Mud13 is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. *Genes Dev.*, **10**, 1699–1708.

Crispino,J.D., Blencowe,B.B. and Sharp,P.A. (1994) Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science*, **265**, 1866–1869.

Fouser,L.A. and Friesen,J.D. (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.

Fouser,L.A. and Friesen,J.D. (1987) Effects on mRNA splicing of mutations in the 3′ region of the *Saccharomyces cerevisiae* actin intron. *Mol. Cell. Biol.*, **7**, 225–230.

Fu,X.D. (1995) The superfamily of arginine/serine-rich splicing factors. *RNA*, **1**, 663–680.

Gozani,O., Patton,J.G. and Reed,R. (1994) A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. *EMBO J.*, **13**, 3356–3367.

Gozani,O., Feld,R. and Reed,R. (1996) Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes Dev.*, **10**, 233–243.

Guarente,L., Yocum,R. and Gifford,P. (1982) A *GAL10–CYC1* hybrid yeast promoter identifies the Gal4 regulatory region as an upstream site. *Proc. Natl Acad. Sci. USA*, **79**, 7410–7414.

Hodges,P.E. and Beggs,J.D. (1994) U2 fulfils a commitment. *Curr. Biol.*, **4**, 264–267.

Jacquier,A., Rodriguez,J.R. and Rosbash,M. (1985) A quantitative analysis of the effects of 5′ junction and TACTAAC box mutants and mutant combinations on yeast mRNA splicing. *Cell*, **43**, 423–430.

Krämer,A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.*, **65**, 367–409.

Krämer,A., Legrain,P., Mulhauser,F., Gröning,K., Brosi,R. and Bilbe,G. (1994) Splicing factor SF3a60 is the mammalian homologue of PRP9 of *S.cerevisiae*: the conserved zinc finger-like motif is functionally exchangeable *in vivo*. *Nucleic Acids Res.*, **22**, 5223–5228.

Krämer,A., Mulhauser,F., Wersing,C., Gröning,K. and Bilbe,G. (1995) Mammmalian splicing factor SF3a120 represents a new member of the SURP family of proteins and is homologous to the essential splicing factor PRP21p of *Saccharomyces cerevisiae*. *RNA*, **1**, 260–272.

Legrain,P. and Chapon,C. (1993) Interaction between PRP11 and SPP91 yeast splicing factors and characterization of a PRP9–PRP11–SPP91 complex. *Science*, **262**, 108–110.

Legrain,P. and Rosbash,M. (1989) Some *cis*- and *trans*-acting mutants for splicing target pre-mRNA to the cytoplasm. *Cell*, **57**, 573–583.

Legrain,P., Séraphin,B. and Rosbash,M. (1988) Early commitment of yeast pre-mRNA to the spliceosome pathway. *Mol. Cell. Biol.*, **8**, 3755–3760.

Liao,X.C., Tang,J. and Rosbash,M. (1993) An enhancer screen identifies a gene that encodes the yeast U1 snRNP A protein: implications for snRNP protein function in pre-mRNA splicing. *Genes Dev.*, **7**, 419–428.

Long,R.M., Elliott,D.J., Stutz,F., Rosbash,M. and Singer,R.H. (1995) Spatial consequences of defective processing of specific yeast mRNAs revealed by fluorescent *in situ* hybridization. *RNA*, **1**, 1071–1078.

MacMillan,A.M., Query,C.C., Allerson,C.R., Chen,S., Verdine,G.L. and Sharp,P.A. (1994) Dynamic association of proteins with the pre-mRNA branch region. *Genes Dev.*, **8**, 3008–3020.

Madhani,H.D. and Guthrie,C. (1994) Dynamic RNA–RNA interactions in the spliceosome. *Annu. Rev. Genet.*, **28**, 1–26.

Michaud,S. and Reed,R. (1991) An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes Dev.*, **5**, 2534–2546

Michaud,S. and Reed,R. (1993) A functional association between the 5′ and 3′ splice site is established in the earliest prespliceosome complex (E) in mammals. *Genes Dev.*, **7**, 1008–1020.

Parker,R. and Guthrie,C. (1985) A point mutation in the conserved hexanucleotide at a yeast 5′ splice junction uncouples recognition, cleavage, and ligation. *Cell*, **41**, 107–118.

Parker,R., Siliciano,P.G. and Guthrie,C. (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell*, **49**, 229–239.

Pikielny,C.W. and Rosbash,M. (1985) mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell*, **41**, 119–126.

Pikielny,C.W. and Rosbash,M. (1986) Specific small nuclear RNAs are associated with yeast spliceosomes. *Cell*, **45**, 869–877.

Rain,J.-C., Tartakoff,A.M., Krämer,A. and Legrain,P. (1996) Essential domains of the PRP21 splicing factor are implicated in the binding to PRP9 and PRP11 proteins and are conserved through evolution. *RNA*, **2**, 535–550.

Rosbash,M. and Séraphin,B. (1991) Who's on first? The U1 snRNP–5′ splice site interaction and splicing. *Trends Biochem. Sci.*, **16**, 187–190.

Rothstein,R.J. (1983) One-step gene disruption in yeast. *Methods Enzymol.*, **101**, 202–211.

Ruby,S.W., Chang,T.H. and Abelson,J. (1993) Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. *Genes Dev.*, **7**, 1909–1925.

Rymond,B.C. and Rosbash,M. (1985) Cleavage of 5′ splice site and lariat formation are independent of 3′ splice site in yeast mRNA splicing. *Nature*, **317**, 735–737.

Séraphin,B. and Rosbash,M. (1989) Identification of functional U1 snRNA–pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell*, **59**, 349–358.

Séraphin,B. and Rosbash,M. (1990) Exon mutations uncouple 5′ splice site selection from U1 snRNA pairing. *Cell*, **63**, 619–629.

Séraphin,B. and Rosbash,M. (1991) The yeast branchpoint sequence is not required for the formation of a stable U1 snRNA–pre-mRNA complex and is recognized in the absence of U2 snRNA. *EMBO J.*, **10**, 1209–1216.

Smith,V. and Barrell,B. (1991) Cloning of a yeast U1 snRNP 70K protein homologue: functional conservation of RNA-binding domain between humans and yeast. *EMBO J.*, **10**, 2627–2634.

Tarn,W.-H. and Steitz,J.A. (1994) SR proteins can compensate for the loss of U1 snRNP functions *in vitro*. *Genes Dev.*, **8**, 2704–2717.

Transy,C. and Legrain,P. (1995) The two-hybrid: an *in vivo* protein–protein interaction assay. *Mol. Biol. Rep.*, **21**, 119–127.

Valcàrcel,J., Gaur,R.K., Singh,R. and Green,M.R. (1996) Interaction of U2AF^{65} RS region with pre-mRNA and promotion of branch point base pairing with U2 snRNA. *Science*, **273**, 1706–1709.

Vijayraghavan,U., Parker,R., Tamm,J., Iimura,Y., Rossi,J., Abelson,J. and Guthrie,C. (1986) Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome. *EMBO J.*, **5**, 1683–1695.

Wells,S.E. and Ares,M.J. (1994) Interactions between highly conserved U2 small nuclear RNA structures and Prp5p, Prp9p, Prp11p, and Prp21p proteins are required to ensure integrity of the U2 small nuclear ribonucleoprotein in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **14**, 6337–6349.

Wells,S.E., Neville,M., Haynes,M., Wang,J., Igel,H. and Ares,M.J. (1996) CUS1, a suppressor of cold-sensitive U2 snRNA mutations, is a novel yeast splicing factor homologous to human SAP 145. *Genes Dev.*, **10**, 220–232.

Received on October 4, 1996; revised on November 25, 1996