MECHANISM OF REGULATION OF THE *RPL30* pre-mRNA SPLICING IN YEAST

SARA MACÍAS RIBELA

PhD Thesis

Barcelona, April 2008

MECHANISM OF REGULATION OF THE *RPL30* pre-mRNA SPLICING IN YEAST

SARA MACÍAS RIBELA

Memòria presentada per optar al Grau de Doctor per la Universitat Pompeu Fabra.

Aquesta tesi ha estat realitzada sota la direcció del

Dr. Josep Vilardell al Centre de Regulació Genòmica (CRG)

Dr. Josep Vilardell **Sara Macías**

Barcelona, April 2008

INDEX

SUMMARY

The mechanisms of pre-mRNA splicing regulation are poorly understood. Here we dissect how the *Saccharomyces cerevisiae* ribosomal L30 protein blocks splicing of its pre-mRNA upon binding a kink-turn structure including the 5' splice site. We show that L30 binds the nascent *RPL30* transcript without preventing recognition of the 5' splice site by U1 snRNP but blocking U2 snRNP association with the branch site. Interaction of the factors BBP and Mud2p with the intron, relevant for U2 snRNP recruitment, is not affected by L30. Furthermore, the functions of neither the DEAD-box protein Sub2p in the incipient spliceosome, nor of the U2 snRNP factor Cus2p on branch site recognition, are required for L30 inhibition. These findings contrast with the effects caused by binding a heterologous protein to the same region, completely blocking intron recognition. Collectively, our data suggest that L30 represses a spliceosomal rearrangement required for U2 snRNP association with the nascent *RPL30* transcript.

KEY WORDS

Gene expression

Saccharomyces cerevisiae

Splicing

Spliceosome assembly

Regulation of splicing

Transcription

U1 snRNP

U2 snRNP

RPL30

L30

Intron

Pre-mRNA

5’ splice site (5’ss)

3’ splice site (3’ss)

Branch site (BS)

Poylpyrimidine tract (Ppy tract)

ABBREVIATIONS

DNA: deoxyribonucleic acid

RNA: ribonucleic acid

mRNA: mature ribonucleic acid

pre-mRNA: pre-mature ribonucleic acid

RNAP II: RNA polymerase II

CTD: C-terminal domain

CBC: cap-binding complex

snRNP: small nuclear ribonucleoprotein

U1 snRNP: Uridine-rich 1 small nuclear ribonucleoprotein

NTC: nineteen complex

5'ss: 5' splice site

3'ss: 3' splice site

BS: branch site

Ppy: polypyrimidine tract

ChIP: chromatin immunoprecipitation

PCR: polymerase chain reaction

Kb: kilobase

bp: base-pairs

CC1: commitment complex 1

CC2: commitment complex 2

SP: spliceosome

IC: inhibited complex

ICGs: intron-containing genes

INTRODUCTION

1. GENE EXPRESSION, a global view:

Eukaryotic gene expression is composed of several processes, such as transcription, splicing, polyadenylation, export, translation and RNA turnover, amongst others.

Recently the view of gene expression has changed significantly, with evidences suggesting that all these processes are influenced by one another. The emerging picture is one in which most steps are physically and functionally connected (figure 1).

After RNA polymerase II (RNAP II) initiates transcription, the nascent RNA is modified by the addition of the “cap” structure at its 5’end. This cap serves initially to protect new transcript from attack by nucleases and later serves as a binding site for proteins involved in export of the mature mRNA into the cytoplasm and its translation into protein. The process of transcription is coupled to pre-mRNA splicing that removes non-coding sequences from transcripts. When transcription finishes, the newly synthesized RNA is cleaved and a polyadenosine tail is added to the 3’end of the transcript. These processes are followed by the export of the mRNA to the cytoplasm where it will be translated to protein by the ribosomes.

In vitro systems have demonstrated interconnections between the different steps of gene expression. For instance, transcription apparatus plays an active role in recruiting the machinery that caps and processes the nascent RNA transcript (Proudfoot et al., 2002; Shatkin and Manley, 2000), and pre-mRNA splicing promotes transcription elongation (Fong and Zhou, 2001) and is required for efficient export of the resulting mRNA into the cytoplasm (Reed and Hurt, 2002). Thus, important instances of gene regulation can be achieved through the interplay of these mRNA processing mechanisms.

In particular, pre-mRNA processing reactions begin to occur during transcription (Hirose and Manley, 2000). A key player in the coupling of these processes is the domain present at the C terminus of the largest subunit of RNAP II known as the “CTD” (carboxy-terminal domain). It appears that the CTD is the platform for the ordered assembly of the different families of pre-mRNA processing machineries. The coordinating role played by the RNAP II CTD in RNA processing may also ensure that the reactions occur in the correct order and that the transitions between the reactions are efficient. This organization of events may also introduce a series of quality control mechanisms, as it ensures that no individual step is omitted.

In yeast, it is unclear that the CTD functions as a platform for RNA processing proteins. However, the CBC (cap-binding complex) has been found to play a similar role. CBC is a heterodimer formed by 2 subunits: Sto1p/Cbp80 and Mud13p/Cbp20 that binds directly to the "cap" structure added to the 5' end of RNAP II transcripts. CBC helps in processes such as mRNA export (Izaurralde et al., 1995), translation by interacting with the factor eIF4G (Fortes et al., 2000). It has been also found to play a role in 3'end formation or poly(A) addition (Flaherty et al., 1997). It is also important for nuclear RNA turnover (Das et al., 2003). Last but not least, CBC plays a critical role for splicing and cotranscriptional assembly of the spliceosome machinery (Gornemann et al., 2005)(Bragulat et al, 2008 *in preparation*). The genetic interaction between CBC and U1snRNP components is considered the mediator for these effects (Fortes et al., 1999).

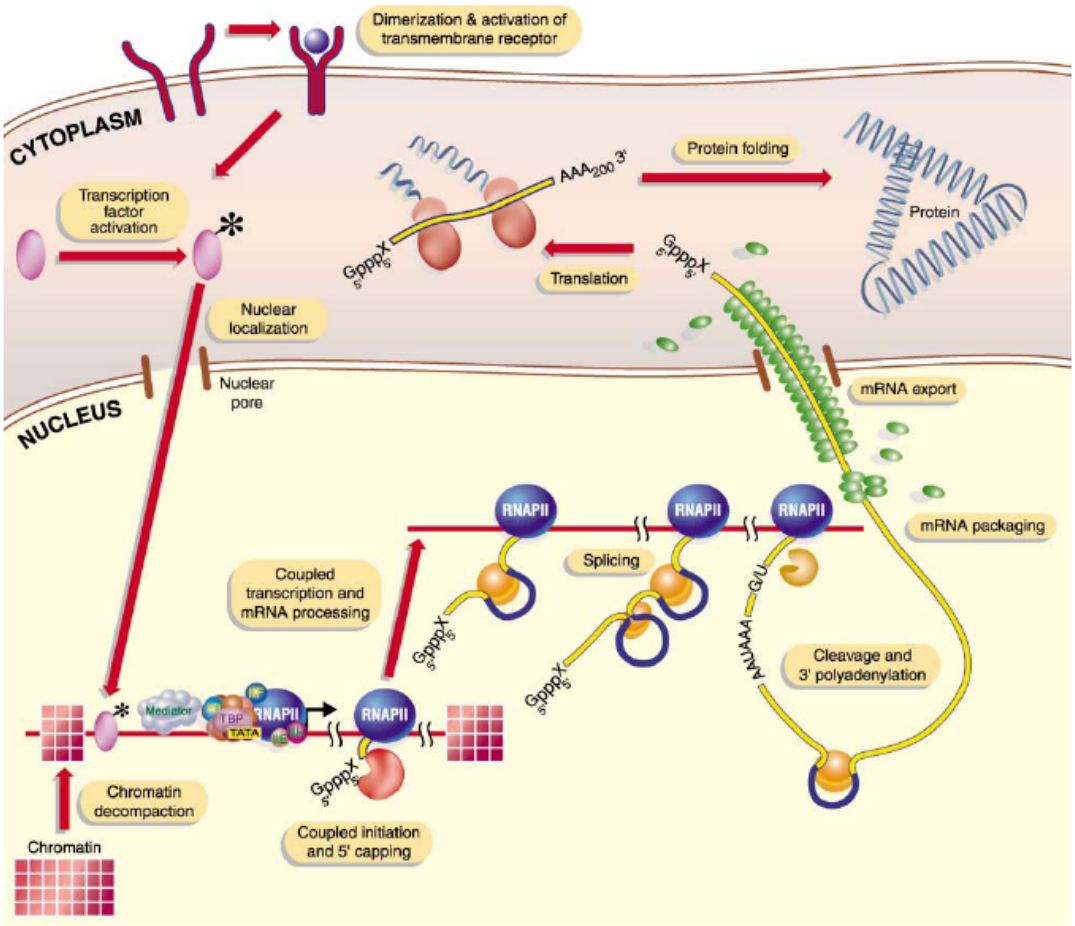

Figure 1. The new view of gene expression. Each stage from transcription to translation is physically and functionally connected to the next (from (Orphanides and Reinberg, 2002).

2. SPLICING

Primary transcripts or pre-mRNAs are often interrupted by non-coding regions, called introns. The splicing process removes introns by a two-step transesterification reaction and joins exons for the formation of the mature RNA. For that reason, splicing is essential to generate a functional message from a DNA template.

Introns contain several *cis* consensus elements, which are essential for the splicing reaction. In yeast, the 5'exon-intron junction or 5'splice site (5'ss) is marked by the consensus sequence GUAUGU (the first nucleotide of the intron is underlined). The end of the intron, the 3'splice site (3'ss), is defined by YAG (Y stands for pyrimidine, the last nucleotide of the intron is underlined). The branchpoint sequence (BS) is found upstream of the 3'splice site with a highly conserved sequence UACUAAC (underlined, the branch adenosine). The branchpoint is usually followed by a pyrimidine-rich tract (figure 2).

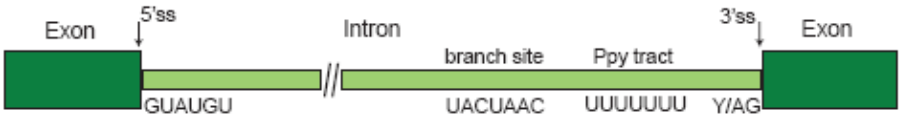

Figure 2. Pre-mRNAs are formed by introns (non-coding sequences) and exons. Introns are delimited by consensus sequences, the 5' splice site (5'ss) and the 3' splice site (3'ss). Introns contain additional information, the branch site sequence (BS) followed by a polypyrimidine-rich tract (Ppy tract).

During splicing catalysis, the 2'hydroxyl of the branch adenosine attacks the phosphate at the 5'splice site, producing a "free" 5'exon and the lariat intermediate. In the second part of the reaction, the 3'hydroxyl of the 5'exon attacks the phosphate at the 3'splice site, resulting in a ligated mRNA and a lariat intron (figure 3). Although this process is highly conserved in all eukaryotes, sequences are more degenerated in metazoans.

The splicing reaction is carried out by the spliceosome, a dynamic 60S ribonucleoprotein particle (reviewed in (Staley and Guthrie, 1998))

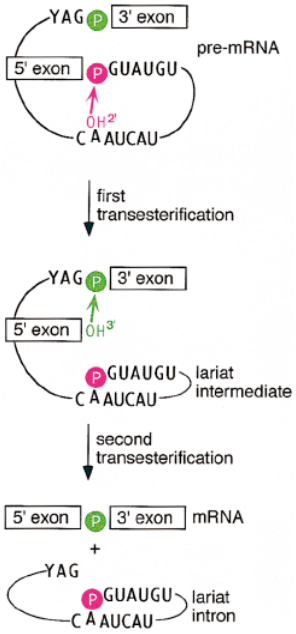

Figure 3. Pre-mRNA splicing occurs in two ATP-independent transesterification reactions. Pink, first transesterification reactants; green, second transesterification reactants (Staley and Guthrie, 1998).

3. **SPLICEOSOME MACHINERY**

The spliceosome machinery is formed by 5 U snRNPs (U-rich small nuclear ribonucleoproteins), U1, U2, U4, U5 and U6. Each of these contains a small stable RNA bound by several proteins. In addition to snRNPs, splicing requires many non-snRNP protein factors. The total number of proteins involved in splicing has been estimated to 300 (Rappsilber et al., 2002).

The spliceosome is conserved from yeast to humans, both in snRNAs and in protein components. It is a highly flexible machinery, as it can excise introns of many different lengths and many different sequences.

Studies of spliceosome assembly *in vitro*, using extracts from whole yeast cells defined an order of interaction of the snRNPs with the pre-mRNA substrate. The first to interact is the U1 snRNP particle, followed by U2 snRNP and finally, binding of the tri-snRNP U4/U6·U5.

But still there is an open discussion about how this macrocomplex, the spliceosome, is organized. Two different models have been proposed: first, spliceosome assembly as a stepwise fashion and second, the spliceosome as a holocomplex.

For the second model, Abelson and coworkers (Stevens et al., 2002) reported purification of a penta-snRNP complex from *S. cerevisiae* that represented a pre-assembled spliceosome, although at a low salt conditions *in vitro*. Suggesting that *in vivo* spliceosome engages the pre-mRNA substrate as a multi-snRNP complex.

But these results were challenged by following reports in which the holospliceosome was not detected *in vivo*. These studies raised again the model in which spliceosome assembles on 5' and 3' splice sites in a stepwise fashion (Gornemann et al., 2005; Lacadie and Rosbash, 2005). In the referenced studies, chromatin immunoprecipitation technique (ChIP) was used to follow spliceosome assembly *in vivo.* This technique has been traditionally used for detection of DNA-protein interactions but, recently, ChIP has been adapted to analyze the association of proteins with nascent RNA transcripts. ChIP consists of formaldehyde cross-linking, shearing of chromatin, and immunoprecipitation of the protein of interest. The detection of the coimmunoprecipitated DNA is made by quantitative PCR using specific oligonucleotides. Although the methodology is identical to DNA ChIP, in this particular case, the protein of interest binds to the nascent RNA chain and then directly or indirectly crosslinks to the DNA template. Direct crosslinking can occur if the nascent RNA chain is close to the template, presumably when the binding site for the protein is near RNA polymerase (see figure 4, left). Although, the extent of direct crosslinking is likely to diminish as the nascent chain is elongated. Indirect crosslinking can occur if the protein of interest crosslinks to the nascent RNA chain and RNA polymerase simultaneously crosslinks to the DNA template. In this case, crosslinking of the protein of interest is highly dependent on the integrity of the nascent RNA strand. Importantly, it is not possible, without a ribonuclase treatment step, to distinguish between direct or indirect crosslinking.

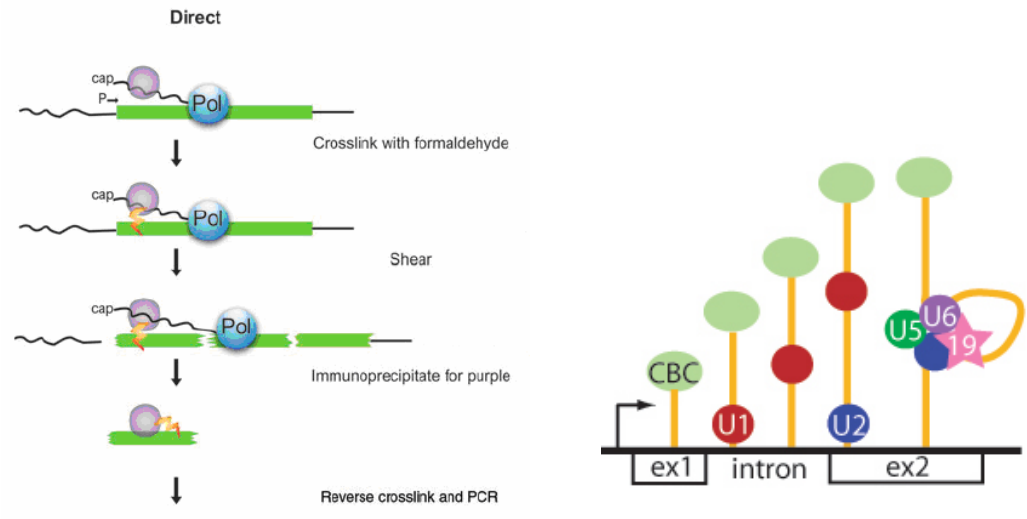

Figure 4. **Chromatin immunoprecipitation technique (ChIP) is used to follow *in vivo* spliceosome assembly.**

On the left, nascent RNA ChIP. Proteins bound to nascent RNA transcripts can be crosslinked to the template DNA. The protein of interest (purple) can be directly or indirectly crosslinked through interactions of the polymerase with the template (Nilsen, 2005).

On the right, schematic diagram of cotranscriptional spliceosomal assembly. DNA is represented by black lines and nascent RNA by orange lines. U1 (red) assembles first, followed by U2 (blue) and U5 (green) and the NTC (nine-teen complex, in pink) (Gornemann et al., 2005).

In the work of Tardiff & Rosbash the model of stepwise recruitment is reinforced combining ChIP with *in vivo* depletions of U1, U2 or U5. From their results they conclude that U1 snRNP recruitment is needed for recruitment of all subsequent snRNPs. The formation of the U1—pre-mRNA complex is independent of U2 snRNP and the tri-snRNP (U4/U6·U5), and U1—U2 pre-spliceosomes form in the absence of the tri-snRNP. The conclusion driven by this model is that snRNP recruitment to the nascent pre-mRNA predominantly proceeds via the canonical three-step assembly pathway. First, U1 binds the pre-mRNA, then U2 and finally the preassembled U4/U6·U5 tri-snRNP. However, Tardiff and coworkers (Tardiff et al., 2006) showed that full spliceosome assembly is usually completed after transcription, depending on the length of the downstream exon. Chip-on-CHIP analyses with whole-genome tiling arrays showed that, whereas U1 snRNP recruitment is independent of second exon length, U2 and U5 recruitment is dependent on that. As a consequence genes with short second exons undergo predominantly post-transcriptional splicing. This work proposes that cotranscriptional splicing can only take place on genes that have a second exon of 1Kb. If the exon is shorther than 1Kb, splicing and

3'end formation machinery will compete for the cotranscriptional recruitment to the nascent pre-mRNA.

4. SPLICEOSOME ASSEMBLY

Most of what is known about spliceosome assembly has been determined *in vitro.* The driven model by these experiments shows a series of intermediate stages of spliceosome assembly by using native gel technique. By this method, the complexes observed are: commitment complex 1 (CC1), commitment complex 2 (CC2), pre-spliceosome and mature spliceosome.

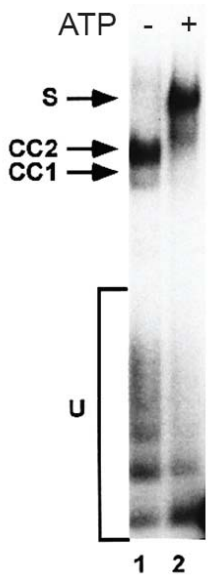

Figure 5. Spliceosome assembly can be visualized using native gel analyses.
Radiolabelled pre-mRNAs incubated in the presence of splicing extracts can assemble in CC1 (commitment complex 1) and CC2 (commitment complex 2). (U stands for unrelated to splicing complexes). Upon addition of ATP, commitment complexes disappear and a new complex with lower mobility is formed, called the spliceosome. Yeast extracts do not allow observation of pre-spliceosomes by this technique. In contrast, HeLa extracts can resolve pre-spliceosomes (called complex A).
Figure adapted from (Caspary and Seraphin, 1998).

Commitment complex:

Assembly begins with the association of the U1 snRNP with the pre-mRNA. U1 snRNP particle is formed by U1 snRNA molecule and 10 U1-specific proteins (Snp1, Mud1, Yhc1, Prp39, Prp40, Snu56, Snu71, Snu65, Luc7 and Mud15). The 5'end of U1 RNA interacts through base-pairing with the 5' splice site (Seraphin et al., 1988; Siliciano and Guthrie, 1988; Zhuang and Weiner, 1986). This first complex is called commitment complex 1, because it is accepted that once it is achieved, the pre-mRNA is committed for splicing.

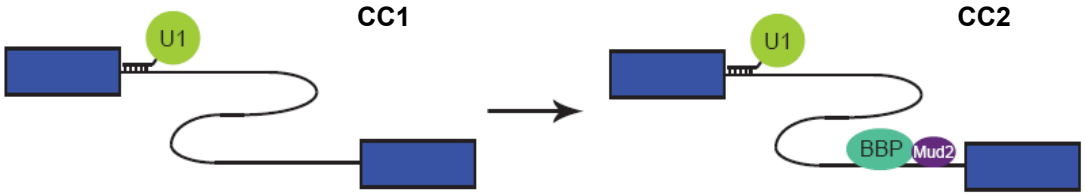

Figure 6. Commitment complex formation. First, U1 snRNP recognizes the 5'ss and forms the CC1 (commitment complex 1). Recognition of the 3'end of the intron by BBP and Mud2 forms the CC2 (commitment complex 2).

Additionally, the branch point sequence is recognized early by BBP (SF1 in mammals), the branch point binding protein, in a sequence-specific fashion (Berglund et al., 1997). This second step forms the commitment complex 2. Mud2p protein has also been detected in this complex (Abovich et al., 1994). Its function has been inferred from the mammalian counterpart, $U2AF^{65}$. This protein interacts with the pyrimidine-rich sequence that often follows metazoan branch-points (Zamore and Green, 1989; Zamore et al., 1992) and is required for U2 snRNP addition (Ruskin et al., 1988). Additionally, mammalian CC2 contains $U2AF^{35}$ that recognizes the conserved dinucleotide AG at the end of introns (Wu et al., 1999). $U2AF^{65}$ and $U2AF^{35}$ work as a dimer, but no homolog has been found for the second in yeast.

Genetic and biochemical experiments indicate a direct interaction between BBP with Mud2p and the U1 snRNP protein Prp40p (Abovich and Rosbash, 1997). This defines a bridge between the two ends of the intron, in which BBP is simultaneously linked to Prp40p and to Mud2p.

It is important to notice that all these steps can be attained *in vitro* in the absence of ATP (Seraphin and Rosbash, 1989).

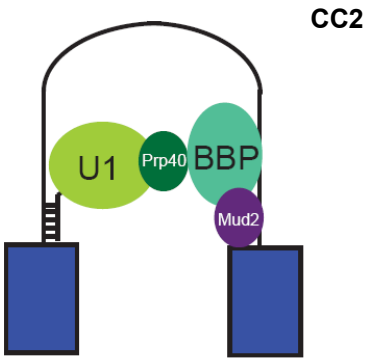

Figure 7. Cross-intron bridging interactions. Prp40p has been shown to interact with BBP in vitro, as well as, BBP can interact with Mud2p. These interactions are thought to be conserved in mammals (Abovich and Rosbash, 1997).

Pre-spliceosome:

Pre-spliceosome formation is the first ATP-dependent step in the spliceosome assembly pathway. It forms after binding of the U2 snRNP particle to the branch point sequence (Parker et al., 1987). For that purpose, U2 snRNP requires a specific conformation and the displacement of BBP from the branch point (figure 7).

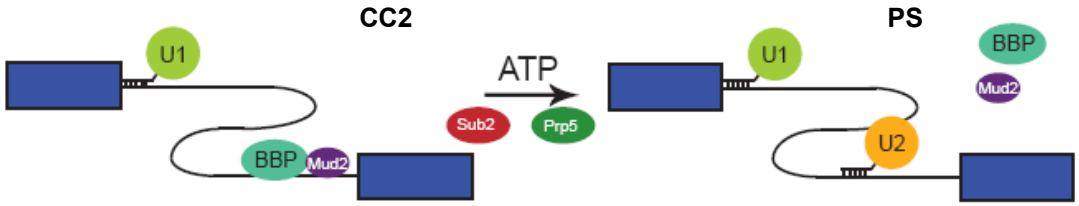

Figure 7. Pre-spliceosome formation. Helicases, such as Sub2p and Prp5p remodel the commitment complex 2 to allow stable base-pairing of U2 snRNP to the branch-site sequence.

Additional proteins are necessary for this reorganization. The ATPases Sub2p and Prp5p are essential for U2 snRNP addition. *SUB2* has been shown to genetically interact with *MUD2* (Kistler and Guthrie, 2001). Sub2p has been proposed to displace Mud2p from the polypyrimidine tract. However, it is still not clear what is the direct target of Sub2p. Recent data shows that, in fact, BBP and Mud2p are found as a pre-formed heterodimer in cells. Moreover, specific mutation in BBP can also bypass the requirement of Sub2p protein (Wang et al., 2008). Thus, the direct target of Sub2p could be the BBP-Mud2 dimer itself.

In addition, Prp5p has been involved in the conversion of U2 snRNA into an active form necessary for base-pairing to the pre-mRNA (Perriman and Ares, 2000; Perriman et al., 2003). U2 snRNA structure can be in two different conformations: U2-stem IIc, inactive for spliceosome assembly, and U2-stem IIa, active for pre-spliceosome formation (Perriman and Ares, 2007). The conversion of U2 snRNA into the active form is dependent on Cus2p and Prp5p. Cus2p stabilizes the inactive form of U2 snRNA (U2 stem-IIc) and Prp5p has been postulated to destabilize Cus2p from U2 snRNP allowing activation, in an ATP-dependent manner (Perriman et al., 2003). For this reason, deletion of *CUS2* allows pre-spliceosome formation in the absence of ATP in yeast. However, Prp5p is still necessary for other ATP-independent functions.

Spliceosome:

The next step in spliceosome assembly is the addition of the tri-snRNP U4/U6·U5 (Konarska and Sharp, 1987).

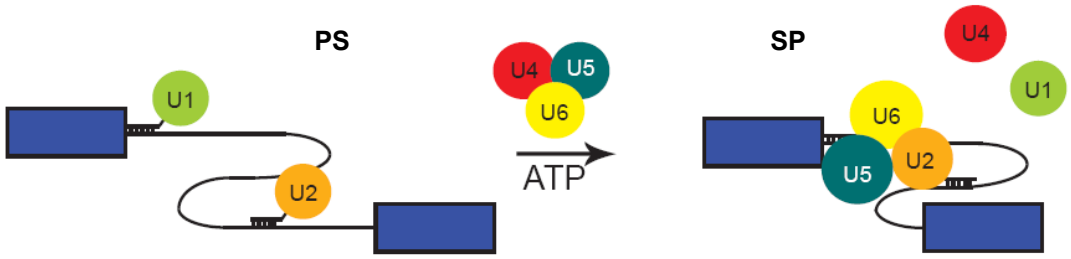

Figure 8. Spliceosome formation. During the last step of spliceosome assembly the tri-snRNP U4/U6·U5 is added and by several reorganizations of the complex the active spliceosome is formed. PS stands fpr pre-spliceosome, SP for spliceosome.

After addition of the tri-snRNP particle, rearrangements promoted by RNA helicases juxtapose the 5' and 3' splice sites and form the catalytic core. Specifically, the U4-U6 duplexes unwind (Lamond et al., 1988), and the U4 and U1 snRNPs are displaced, which allows U6 to form base-pairing interaction with the 5'ss (Wassarman and Steitz, 1992) and with a region of U2 that is near to the U2 branch-site duplex (Datta and Weiner, 1991; Hausner et al., 1990; Madhani and Guthrie, 1992; Wu and Manley, 1991). The U5 snRNP has been shown to base-pair with sequences in both the 5' and 3' exons, and is believed to position the ends of the two exons for the second step of splicing (Newman and Norman, 1992; Sontheimer and Steitz, 1993; Wassarman and Steitz, 1992; Wyatt et al., 1992). After the second step of splicing has been completed, the ligated exons and a lariat intron are released, and the spliceosomal components dissociate and are recycled for further rounds of splicing.

Different proteins are needed for all these rearrangements. Prp28p, that is essential, is necessary for the displacement of the U1 snRNA by the U6 snRNA at the 5'splice site, destabilizing directly the U1·pre-mRNA interaction or by displacing the U1-C protein (component of U1 snRNP), which stabilizes this interaction (Chen et al., 2001; Staley and Guthrie, 1999). Furthermore, Prp28p also participates in rearrangements of the U6 snRNA structure prior to binding to the pre-mRNA (Staley and Guthrie, 1999; Strauss and Guthrie, 1991).

Brr2p mediates the release of U4 snRNA from the spliceosome, likely by unwinding the U4-U6 base pairing (Raghunathan and Guthrie, 1998). Once U6 snRNP is free it can base pair with U2 snRNA for catalysis. Snu114p, the only GTPase in the spliceosome, is also involved in the unwinding of U4-U6 interaction (Bartels et al., 2002). It is an U5 snRNP component that interacts with another U5 protein, Prp8p, to carry out this process (Boon et al., 2006). In addition, Prp8p protein is also necessary for reorganizations occurring prior to the second step of splicing. This protein is the largest conserved nuclear protein spanning the eukaryotic taxa. Prp8p interacts with multiple protein and RNAs. It is considered to be performing a scaffold-like role in spliceosome, holding on to many different components (reviewed in (Grainger and Beggs, 2005)).

Prp2p has also been implicated in the structural reorganization of the spliceosome for the first transesterification step (Kim and Lin, 1996; Roy et al., 1995). Interestingly, Prp16p seems to play a similar role before or during the second transesterification reaction, possibly by promoting reformation of U2-stem IIa, also necessary for this step (Hilliker et al., 2007; Perriman and Ares, 2007). Other proteins have been implicated in the second catalytic step: Prp17p (Jones et al., 1995), Slu7p, Prp18p and Prp22p (Ansari and Schwer, 1995; Horowitz and Abelson, 1993; Schwer and Gross, 1998; Schwer and Guthrie, 1992).

The Prp19p-associated complex, or NTC (for “nineteen complex”), is associated with the spliceosome and plays an important role in mediating structural rearrangement of the spliceosome during its activation (Tarn et al., 1994; Tarn et al., 1993). NTC is required for the stabilization of U5 and U6 in the spliceosome during spliceosome activation prior to the first catalytical step (Chan et al., 2003). Stabilization of U5 and U6 by NTC is achieved in part through specifying interactions between U6 and the 5′ splice site and between U5 and the pre-mRNA (Chan and Cheng, 2005).

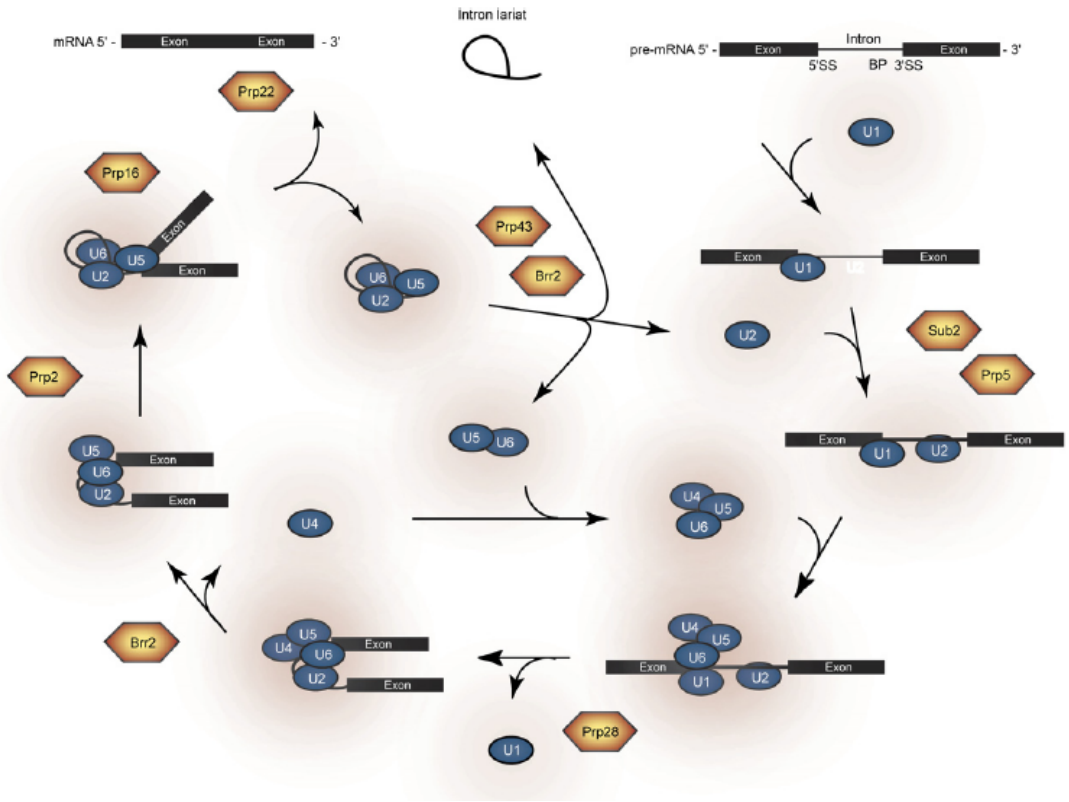

Figure 9. **RNA helicases in pre-mRNA splicing** (Bleichert and Baserga, 2007).

Postcatalytic rearrangements and recycling

After the two transesterification reactions, splicing is completed. The mRNA and the lariat intron need to be released from the spliceosome, and the spliceosome is recycled.

Prp22p and Prp43p function in spliceosome disassembly after both reactions have been carried out. Prp22p releases the mRNA from the spliceosome (Schwer and Gross, 1998).

Prp43p forms a complex with Ntr1 and Ntr2 termed NTR complex which catalyzes spliceosome disassembly (Tsai et al., 2007). Snu114p has been also implicated in the disassembly of the postsplicing complex U2/U6·U5 (Small et al., 2006).

Moreover, the mutually exclusive pairings involving U2, U6 and U4 must be restored to their original conformations. This is in part achieved by Prp24p, a RNA helicase whose function is to promote the annealing between U4 and U6 (Raghunathan and Guthrie, 1998).

5. QUALITY CONTROL IN SPLICING

Any mistake in the process of recognition and removal of introns would lead to an altered genetic message having catastrophic consequences at the level of the protein sequence. In fact, splicing signals contain low information and in some cases, as for metazoans, are poorly conserved. The mechanism that allows splicing fidelity is still not well understood, but there is an increasing number of publications reporting mutations in spliceosomal proteins that result in a loss of splicing fidelity. These evidences are helping to understand how the "splicing kinetic proofreading" process works (Konarska et al., 2006; Liu et al., 2007; Mayas et al., 2006; Umen and Guthrie, 1996; Villa and Guthrie, 2005; Xu and Query, 2007)-

Numerous DExH/D-box ATPase helicases participate in the splicing process, and each is thought to facilitate a structural transition by coupling ATP hydrolysis to a remodelling step of RNA-RNA or RNA-protein interactions.

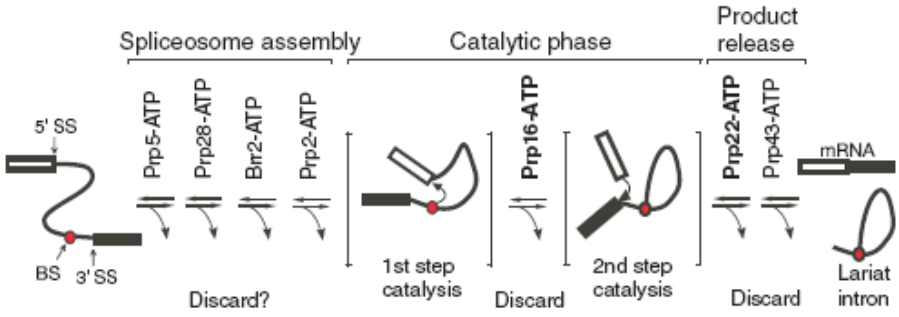

Figure 10. **Structural transitions during splicing are facilitated by DExH/D box ATPase helicases** (Query and Konarska, 2006).

The accepted model is that mutations in these helicases improve the splicing of suboptimal substrates by lowering the rate of ATP hydrolysis. For instance, Prp16p helicase was identified as an ATPase that facilitates the transition of the spliceosome from the first catalytic step to the second (Schwer and Guthrie, 1991). Mutations in the ATPase domain of Prp16p are thought to slow the rate of exit from the first step conformation (Query and Konarska, 2004) favouring catalysis in front of rejection of a suboptimal substrate (Burgess et al., 1990).

The same has been reported for other ATPases (Mayas et al., 2006; Umen and Guthrie, 1995a; Villa and Guthrie, 2005) (see figure 10).

By this view, rejection of suboptimal substrates can occur at every ATP-dependent transition along the pathway. Mutations in the ATPases affect these transitions leading to an altered fidelity.

The last example provided for this model is the Prp5p function in the transition between commitment complex 2 and pre-spliceosome formation (Xu and Query, 2007). These authors showed that decreased Prp5p ATPase activity results in improved splicing of introns with suboptimal branch regions. They proposed that U2snRNA—branch site pairing followed by conformational change of U2 mediated by Prp5p results in productive splicing. In contrast, U2 conformational change prior to pairing results in an abortive path, discarding the mutant transcript (figure 11).

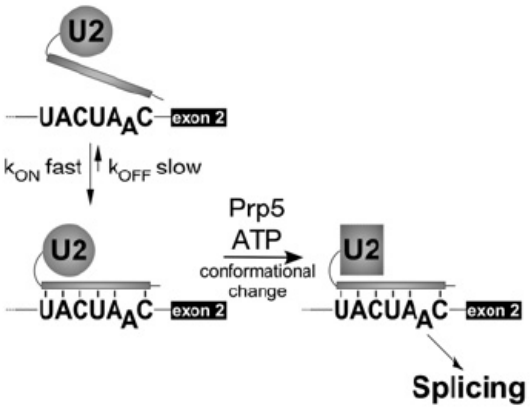

Figure 11. **Model for Prp5p-mediated spliceosomal transition** (Xu and Query, 2007)

6. REGULATION OF SPLICING

The splicing process is subjected to regulation. In metazoans, alternative splicing is the major mechanism known to be regulated, giving raise to an increase in protein diversity from a single gene. Alternative splicing includes the extension or shortening of an exon, the skipping or inclusion, and the removal or retention of an intron (figure 12, reviewed in (Maniatis and Tasic, 2002)). In general, alternative exons have suboptimal splice sites or suboptimal lengths. In addition, splicing of the regulated exons is modulated by *trans*-acting factors that recognize splicing enhancers (positive for inclusion of exon) or splicing silencers (negative for inclusion). SR proteins (serine/arginine-rich splicing factors), hnRNP family and SR-like proteins are found among these factors (Cowper et al., 2001).

While in humans at least 74% of multi-exon genes are alternative spliced, few examples are known in yeast *S.cerevisiae.*

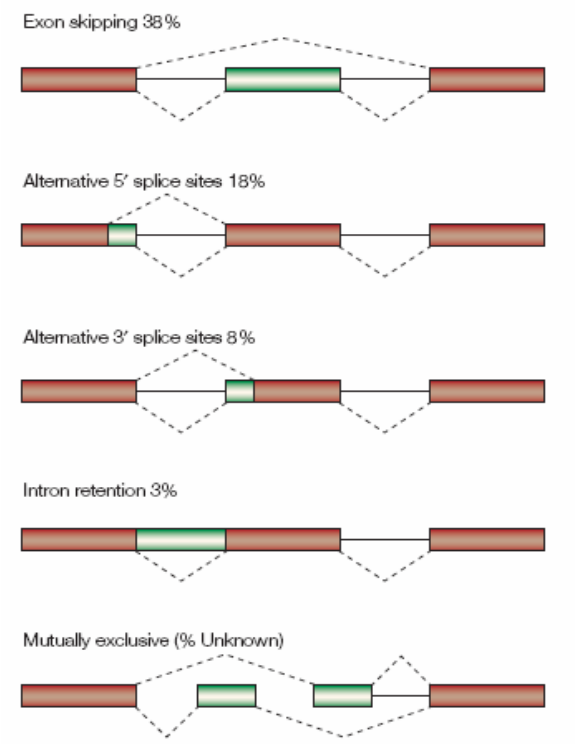

Figure 12. **Types of alternative splicing**. Constitutive exons are shown in red and alternatively spliced regions in green. Introns are represented by solid lines, and dashed lines indicate splicing activities (Ast, 2004).

7. SPLICING REGULATION IN YEAST

Budding yeast (*Saccharomyces cerevisiae*) has less than 254 spliceosomal introns in the more than 6200 annotated genes. Only 10 genes are known to have more than one intron (Spingola et al., 1999). In contrast, intron-containing genes are highly transcribed because 10.000 of the nearly 38.000 mRNA molecules made each hour are derived from this class of genes (Holstege et al., 1998). The major functional class of intron-containing genes code for ribosomal proteins. Nearly 100 introns are in ribosomal protein genes, accounting for about 90% of all mRNA intron-containing transcripts.

The fact that, in yeast most of the intron-containing genes contain only one intron reduces the possibilities of alternative splicing as it happens in vertebrates. Nevertheless, some examples of alternative splicing have been reported, consisting in a majority of cases of an intron retention event. Moreover there is no evidence for the existence of functional homologs of SR proteins or hnRNP proteins. These proteins activate or repress the inclusion of a particular exon through binding to “enhancers” or “silencers” sequences both in the intron and the exon, in metazoans.

In the group of ribosomal protein genes there are two known examples of splicing regulation. S14 protein is encoded by two different genes, RPS14A and RPS14B, whose mRNAs are found in a ratio 10:1. However, the two genes are transcribed approximately to an equal extent (Li et al., 1995). Excess S14 can bind to an RNA stem-loop structure in RPS14B pre-mRNA that is necessary for regulation, thus inhibiting its splicing and leading to its rapid degradation (Fewell and Woolford, 1999).

The other reported case is L30, whose transcript (*RPL30*) is normally spliced efficiently but in the presence of an excess L30, unspliced precursor accumulates (Vilardell and Warner, 1994).

Analogously, other examples have been found in higher eukaryotes. Human ribosomal protein S26 can also bind to its own pre-mRNA *in vitro* to control splicing (Ivanov et al., 2005). The same has been observed for another human ribosomal protein, S13 (Malygin et al., 2007).

Other important examples of splicing regulation have been detected in yeast, but in non-ribosomal protein genes. The first example is *YRA1* gene. *YRA1* encodes for a component of TREX complex, involved in mRNAs export. *YRA1* controls its own expression with a negative feedback loop in which excess levels of Yra1 inhibit splicing of its own pre-mRNA (Preker et al.,

2002). In normal conditions Yra1p protein would help to splice *YRA1* intron, because of its unusual structure (large intron and non-consensus BS sequence). In contrast, when Yra1p is in excess, it also binds its own pre-mRNA but favouring export in front of splicing (Preker and Guthrie, 2006). The unusual structure of the intron is essential for its autoregulation of splicing.

The second example concerns to meiosis-specific splicing and, in contrast to the previous examples, the splicing regulator enhances positively splicing of a subset of pre-mRNAs. Mer1p, a U1snRNP-associated protein, is expressed only during meiosis and activates the splicing of at least three pre-mRNAs (*AMA1, HFM1/MER2* and *REC107/MER103*). Mer1p specifically binds RNAs that contain a Mer1 enhancer element (Spingola and Ares, 2000). The model proposed is that transcript-bound Mer1 acts at the very first stage of spliceosome assembly to recruit the U1 snRNP to pre-mRNA (Spingola and Ares, 2000; Spingola et al., 2004).

8. REGULATION OF SPLICING BY L30

L30 is an essential ribosomal protein from *Saccharomyces cerevisiae* and it is encoded by the *RPL30* gene. L30 protein that cannot be assembled into ribosomes binds to its own transcript near the 5'splice site, preventing the complete assembly of the spliceosome (Vilardell and Warner, 1994). The ability to regulate the level of *RPL30* mRNA contributes substantially to the biological fitness of the cell (Li et al., 1996), which suggests that even a minor excess of this ribosomal protein has some deleterious effects.

When L30 protein binds to the *RPL30* transcript, this pre-mRNA adopts a structure known as a kink-turn (White et al., 2004). This RNA secondary structure consists of an asymmetric, purine-rich internal loop (2+5) (see figure 13). The mode by which L30 binds its own pre-mRNA resembles the interaction of the ribosomal protein with the kink-turn motif of helix 58 in the ribosome (Halic et al., 2005).

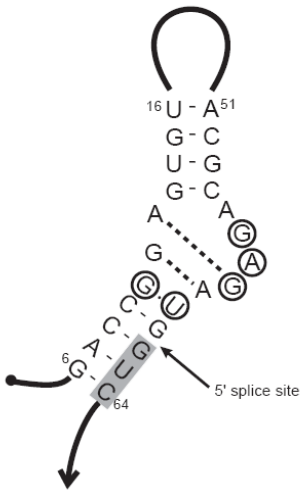

Figure 13. **Schematic representation of the kink-turn of *RPL30* pre-mRNA**. It includes the exon 1 and the first nucleotides of the intron (grey box). The cap is indicated with a solid circle, the depicted interactions in the purine loop are based on the X-ray structure, with the positions shown to contact L30 encircled (Chao and Williamson, 2004). Numbers are relative to the start of transcription.

The *RPL30* pre-mRNA contains two exons and only one intron. Exon 1 consists of 61 nucleotides and it only encodes for translation start codon (ATG), the rest is 5'UTR (untranslated region). The intron consists of 230 nucleotides. The 5'ss differs from the consensus sequence (GT**CA**GT, instead of GTATGT) although the BS sequence matches the consensus. The exon 2 consists of 510 nucleotides that encodes for the L30 protein.

The kink-turn that *RPL30* pre-mRNA adopts, partially occludes the 5'splice site (see figure 13), and it was first hypothesized that L30 stabilized this structure, thereby preventing access of U1 snRNA (Eng and Warner, 1991). However, it was shown both *in vitro* and *in vivo* (Vilardell et al., 2000a; Vilardell and Warner, 1994) that U1snRNP is associated with the *RPL30* pre-mRNA when L30 is also bound to it. This new stalled complex in the spliceosome assembly was named *inhibited complex* (IC).

Studies of the *inhibited complex* in vitro by native gel analyses revealed that its formation is independent of ATP (Vilardell and Warner, 1994). In native gels, the lower bands observed correspond to the commitment complex 1 and 2 (CC1 and CC2, figure 14, lane 2). The upper bands are the ATP-dependent spliceosomal stages, corresponding to the pre-spliceosome and spliceosome (SP, figure 14, lane 3 and 4). In the case of *RPL30* pre-mRNA, addition of L30 prevents formation of the ATP-dependent complexes (figure 14, lane 5). At the same time, L30 addition forms a new complex migrating slower than the commitment complex (the *inhibited complex*, IC) (figure 14, lane 5 and 6). In addition, a mutant pre-mRNA was used, known not to

show splicing inhibition by L30 (C9→T mutant). The mutation consists of a single nucleotidic change in *RPL30* exon 1. C9→T pre-mRNA showed spliceosome formation in the presence of recombinant L30 (lane 13). As the first ATP-dependent step in spliceosome assembly is addition of U2snRNP, it was hypothesized that L30 would be blocking a step before the association of U2snRNP (Vilardell and Warner, 1994).

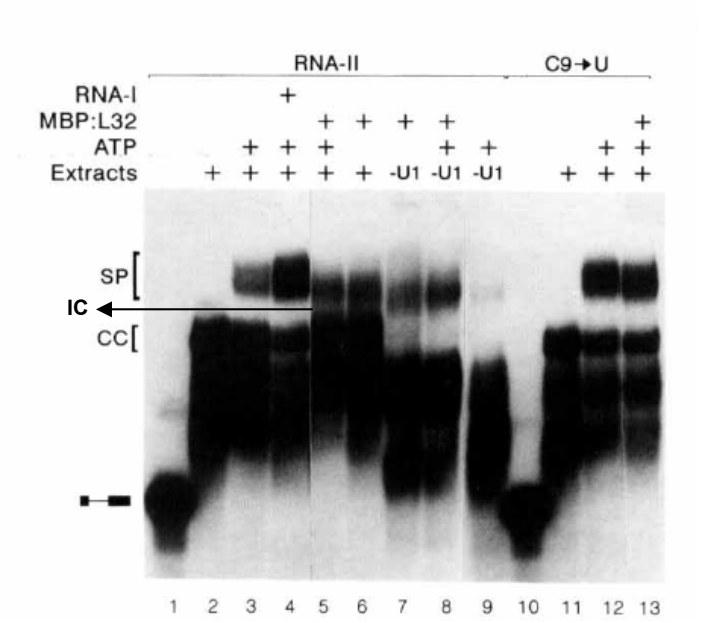

Figure 14. Formation of commitment complexes and mature spliceosomes by the *RPL30* transcript.
Labeled RNA, either WT (lanes 1-9) or the C9→T (lanes 10-13) were used. Complexes different from the CC (commitment complex) and SP (spliceosome) are formed in the presence of MBP:L30 protein independent of ATP (compared lane 5 and 6 with 3 and 4). This is called the *inhibited complex* (IC). In contrast, the C9→U pre-mRNA is unaffected by MBP:L30 (lane 13) (Vilardell and Warner, 1994).

In addition, it was studied how the position of the 5'ss affects regulation of splicing. Splicing inhibition was studied by *in vitro* splicing gels (figure 15), where the catalytical steps of splicing are visualized due to the formation of the splicing intermediates (lariats). The 5'splice site of the *RPL30* pre-mRNA is included in the kink-turn motif that serves as the binding site for L30. When the distance between the L30 binding site is increased to 12 nucleotides, L30 does not prevent splicing (figure 15, lane 8). However, an increased distance of 3 or 6 nucleotides, that is enough to move the 5'ss outside of the stem formed by the kink-turn, were not enough to prevent

regulation by L30 (figure 15, lanes 4 and 6). This result indicates that the 5'ss does not need to be in a stem to observed L30-regulation. The non-regulated transcript was named +12 and we have used it as a positive control for spliceosome assembly, because binds L30 but its splicing is not inhibited.

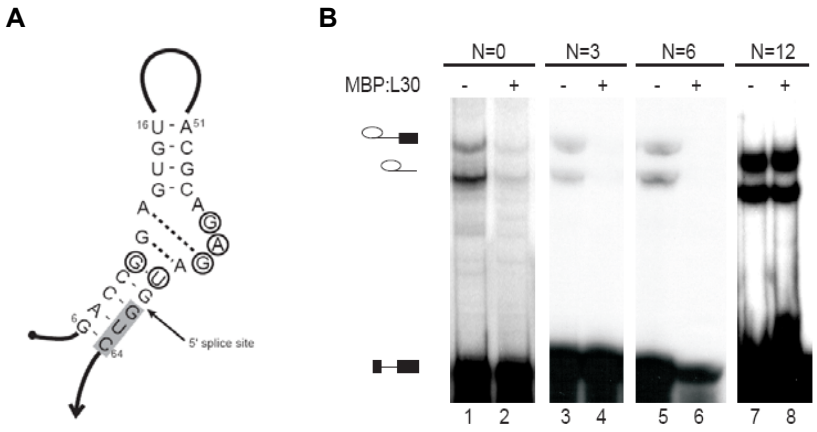

Figure 15. Effect on splicing inhibition of displacing the intron of the *RPL30* from the kink-turn fold by N nucleotides inserted at the end of the first exon while keeping strand complementarity (indicated by an arrow, panel **A**). (**B**) *In vitro* splicing was assessed in the absence (odd lanes) or presence (even lanes) of MBP:L30 fusion protein in the extracts. L30 can repress splicing even when the 5'SS is outside the helix of the k-turn (N=3, N=6). However, when he 5'SS is shifted 12 positions downstream regulation is lost (lane 8).

OBJECTIVES

OBJECTIVES

The main objective of this thesis is the study of the molecular mechanism involved in the regulation of splicing by *S. cerevisiae* L30. The *RPL30*/L30 system has been one of the most studied examples of this type of regulation. Ribosomal protein L30 can bind to its own pre-mRNA (*RPL30*) when it is in excess. L30 binding inhibits spliceosome assembly by an unknown mechanism. Inhibition of the *RPL30* splicing forms a stalled complex, also called *inhibited complex* that at least also contains U1snRNP.

Whereas most of the spliceosomal components are known, very little is understood of how their assembly can be affected. The study of the L30 regulation of splicing offers the chance to study how the early steps spliceosome assembly can be modulated by a non-splicing factor.

The specifics objectives of this work are:

1. *IN VITRO* study of the *inhibited complex* (IC):
    1.1. Study of the interaction between U1 snRNA and the *RPL30* pre-mRNA in the IC.
    1.2. Study of the spliceosomal components present in the IC.
    1.3. Study of the IC (proteins) components essential for L30 regulation
    1.4. Study of the requirements of the regulation of splicing by L30. Role of the RNA secondary structure and recapitulation of the splicing regulation system by an heterologous protein (MS2).

2. *IN VIVO* study of the *inhibited complex* (IC):
    2.1. Purification of the IC by TAP-tagging strategy (tandem-affinity purification).
    2.2. Study of the formation of the complex at the cotranscriptional level by ChIP (chromatin immunoprecipitation technique).

These questions will help to understand how L30 is inhibiting splicing of its own pre-mRNA, and how this system has evolved to be specific and highly regulable.

RESULTS

With the results showed below we describe the molecular mechanism by which the expression of the *RPL30* gene is inhibited at the level of splicing. The *RPL30*/L30 system is one of the most studied examples of splicing regulation in yeast. During this thesis we have determined at which exact level L30 inhibits spliceosome assembly, we demonstrate that it is by a specific mechanism and we hypothesize that L30 must be interfering with a necessary rearrangement for U2snRNP recruitment.

1. L30 DOES NOT PREVENT U1 BASE-PAIRING TO THE 5'SS

Binding of L30 to its own pre-mRNA (*RPL30*) stalls spliceosome assembly at an intermediate stage. Previous work on the mechanism of L30 regulation showed that the first step of spliceosome assembly, that it is U1 snRNP binding, can be observed under conditions of splicing repression by L30 (Vilardell et al., 2000a). However, it was unknown whether the 5'ss was properly recognized by base-pairing, which is relevant to understand the mechanism of repression. To test whether U1 base-pairs to the *RPL30* intron during regulation we used the psoralen-induced crosslinking technique. Psoralen is a chemical compound that intercalates between RNA double-stranded regions. Upon irradiation at 365 nm, psoralen becomes completely bound to the two strands generating a RNA specie that migrates slower in an acrylamide gel. For this experiment, a short version of the *RPL30* intron was used. This RNA contained exon1, a consensus 5'splice site and few nucleotides of the *RPL30* intron. We used consensus 5'ss (GTATGT) instead of *RPL30* 5'ss (GTCAGT) because no psoralen cross-linkings were detected with the last.

This fragment of *RPL30* pre-mRNA was radiolabelled with α-^{32}P- UTP (schematic representation figure 16A). We show that this sequence binds efficiently L30 by gel-shift (figure 16B). The radiolabelled RNA was incubated in the presence of splicing extracts, psoralen and recombinant L30 protein (expressed as MBP-L30). Figure 16C shows that the fragment of pre-mRNA used is capable to crosslink with U1 snRNA (lane 3, upper arrows), and that this crosslinking does not disappear with the addition of MBP:L30 (lane 4). We used RNaseH digestion to make sure that the crosslinked material contained U1 snRNA. After crosslinking and purification of the RNA, a DNA oligo complementary to U1 snRNA was added in the presence of RNAseH. This enzyme degrades RNA-DNA hybrids. In lanes 5 and 6 a faster-migrating band is observed, that corresponds to degradation of U1 snRNA. Finally, we also determined that U1 base-pairing occurs in the same complex where L30 is present. For this reason, crosslinking was followed by immunoprecipitation against maltose binding protein (MBP) that is fused to L30 protein. And U1 crosslinkings coimmunoprecipitated with L30 (lane 7 and 8), indicating that L30 does not prevent U1 snRNA base-pairing to the *RPL30* pre-mRNA. Moreover, we showed in figure 16D that the conditions used for MBP immunoprecipitation are the appropriate, as Snu71

and Prp40 proteins (both U1-specific proteins) can be coimmunoprecipitated with the *inhibited complex* (lane 3) as well as with a positive control (lane 4, also named +12, because the 5'ss has been moved 12 nucleotides downstream, allowing L30 binding but preventing splicing regulation, see introduction figure 16).

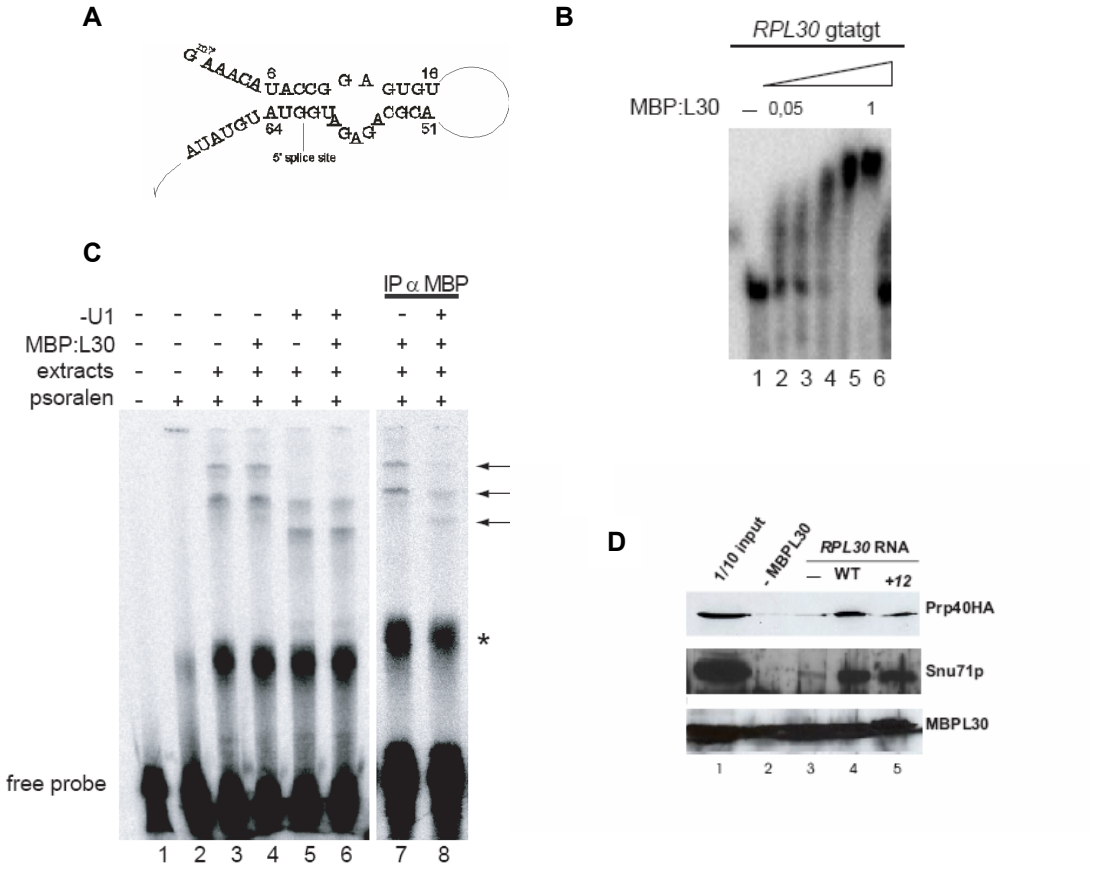

Figure 16. **U1 base-pairs to the *RPL30* 5'ss in conditions of repressed splicing.**

(**A**)Fragment of *RPL30* intron used for psoralen crosslinking experiments. The 5'SS was modified to contain the consensus sequence (GUAUGU). To maintain stem formation, complementary mutations were also introduced in exon 1.

(**B**) Gel shift analyses of this *RPL30* fragment. This RNA binds with high affinity to MBP:L30 protein. Increasing amounts of fusion protein were added (from 50ng to 1µg of protein).

(**C**) *RPL30* 5'SS can crosslink to U1snRNA (lane 3). The specific crosslinks are indicated by arrows on the left (asterisks indicate an unrelated crosslink). U1 crosslinkings are maintained in the presence MBP:L30 (lane 4), and can also be immunoprecipitated with antibodies against MBP tag (lane 7). RNaseH digestions against U1 RNA show that crosslinking bands contain U1 (lane 5, 6 and 8).

(**D**) Immunoprecipitation conditions used in panel C allow to coimmunoprecipitate Snu71 and Prp40 with the inhibited complex (WT, lane 4) as well as in the positive control for binding of L30 but lack of regulation (+12, lane 5)

2. COMPOSITION OF THE INHIBITED COMPLEX AT THE snRNA LEVEL:

To help defining at which step of spliceosome assembly L30 inhibits, we analyzed the composition of the inhibited complex at the level of snRNAs.

In the past, mainly native gel analyses were used to monitor the composition of the IC. As it was previously published (Vilardell and Warner, 1994), the IC contains U1 snRNA and cannot be formed when U1 is inactivated in splicing extracts. However, this technique has some limitations because labile interactions cannot be detected. For this reason we used the coimmunoprecipitation technique to test whether other snRNPs are present in the IC.

For this purpose, we study the snRNA composition of the inhibited complex assembled on a *RPL30 WT* pre-mRNA, containing exon 1, intron and few nucleotides of exon 2. As a positive control for snRNPs coimmunoprecipitation, the *RPL30 +12* pre-mRNA was used. This RNA showed binding for L30 but lack of splicing inhibition, becoming a suitable positive control in the coimmunoprecipitation experiments.

Apart from these two RNAs, we constructed pre-mRNAs where the branch site sequence was deleted (for both transcripts *WT* and *+12*). These RNAs allow determining whether the coimmunoprecipitated components interact with the branch site sequence.

All these transcripts were incubated with splicing extracts supplemented with MBP:L30 protein. An immunoprecipitation against MBP was performed and the coimmunoprecipitated material was analyzed by Northern blot.

Only U1 snRNA could be detected under conditions of repression (figure 17, lane 4) and addition of ATP does not overcome inhibition of the *WT* substrate (lane 8). In contrast, ATP enables progression of the *+12* substrate as evidenced by association of additional snRNAs (lane 10), U2, U5 and U6. The *+12 BSΔ* confirms that U2snRNA coimmunoprecipitation is dependent on the presence of branch site sequence.

In addition, we also observed that L30 is not interfering with the cross-intron interactions formed during the CC2 stage (Abovich and Rosbash, 1997). U1 copurification diminishes when the *WT BSΔ* was used (compare lane 3 and 5). This substrate cannot bind BBP protein, a component of the cross-intron interactions that helps to stabilize U1 at the 5'ss by an interaction with Prp40p (U1component).

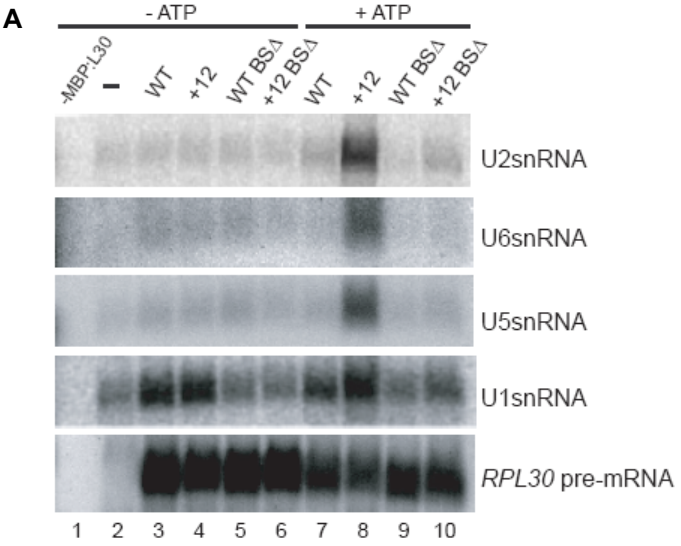

Figure 17. snRNA compositions of the spliceosome stalled by L30 ("inhibited complex" or IC). L30-containing complexes formed on different pre-mRNAs were immunoprecipitated from *in vitro*-assembled reactions and the RNA content was analyzed by Northern blot. Only U1snRNA can be associated with the IC (lane 7), in comparison with the "+12" RNA (lane 8) that can coimmunoprecipitate all the snRNAs from the spliceosome. U4 cannot be tested in this approach as it binds L30 unspecifically, probably because it also forms a canonical kink-turn.

3. U1 STABILITY IN THE *INHIBITED COMPLEX* (IC)

The binding site of L30 in *RPL30* pre-mRNA adopts a kink-turn structure upon protein binding. The 5'ss is occluded in this structure, but contrary to what we initially thought, L30 binding does not interfere with proper base-pairing of U1. However, it remained as a possibility that U1 stability might be compromised in the *inhibited complex*.

For that reason we determined the amount of U1snRNA coimmunoprecipitated with the IC compared to that of a normal commitment complex.

Splicing complexes were formed in the absence of ATP on the *RPL30 WT* and the *+12* substrates and MBP:L30. After coimmunoprecipitation against MBP, increasing salt-washing conditions were used to determine the stability of U1, ranking from 200 to 500 mM KCl. Figure 18A shows that more U1 is coimmunoprecipitated at high salt-washing conditions in the IC than in the CC (400 and 500mM KCl, lane 6 and 8). Although only significant differences were observed at 500mM KCl when different experiments were performed (figure 18B). This suggests that L30 regulation leads to the hyperstabilization of U1 snRNP.

Interestingly, this stability remains in the presence of ATP (figure 18C and 18D), arguing for the independence of L30 from an ATP-dependent remodelling step of the complex necessary for

U2 snRNP binding. This result is also consistent with the lack of U2 association in the IC (figure 17).

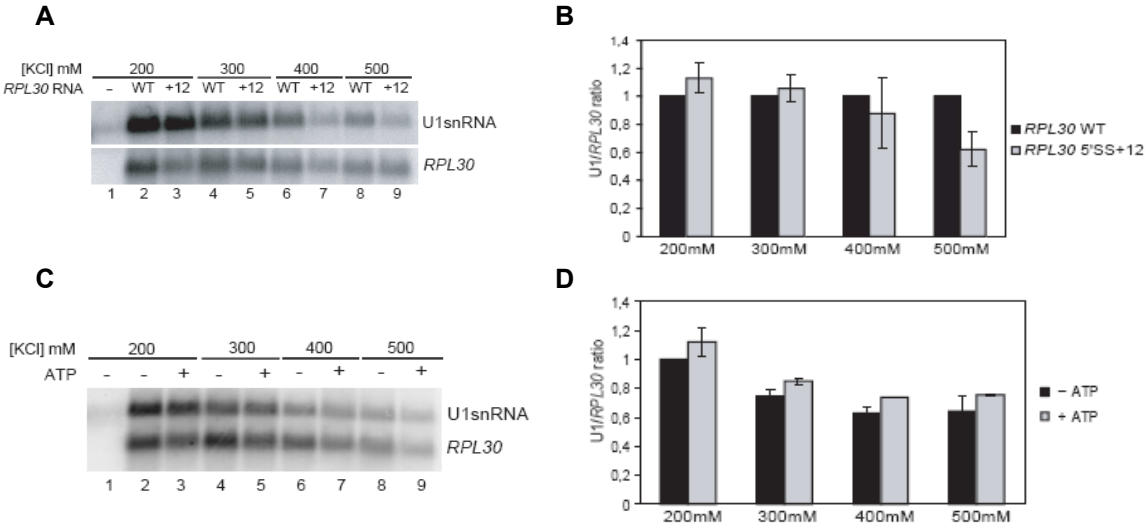

Figure 18. U1snRNP coimmunoprecipitation with L30/*RPL30* at increasing salt concentrations.

Levels of U1snRNA that coimmunoprecipitated with L30 under increasing salt wash conditions were determined by RNA extraction and subsequent Northern analyses. (**A**) Comparison of the stability of U1snRNA in the *inhibited complex* (IC, even lanes), or the commitment complex (CC, odd lanes). U1 is more resistant to high salt washes when it is in the IC compared to a non-regulated situation (lanes 6,7; and 8,9, respectively).

(**B**) Quantification of 3 different experiments as indicated in (A). The U1/*RPL30* ratio of the "*+12*" transcript was set relative to the corresponding *WT RPL30* ratio. The differences are only significant at 500mM KCl.

(**C**) Effect of the ATP on U1snRNA stability from the IC under increased salt washes conditions. Reactions without ATP (even lanes) were compared to those with ATP (odd lanes).

(**D**) Quantification of three different experiments as indicated in (C). The ratios U1/*RPL30* were normalized to the one observed at 200mM in the absence of ATP.

4. L30 DOES NOT INTERFERE WITH RECOGNITION OF THE 3' REGION OF THE INTRON:

Our results clearly show that L30 does not interfere with U1 binding whereas it inhibits U2 snRNP incorporation. Prior to U2 incorporation, BBP recognizes the branch site sequence and Mud2p the polypyrimidine tract. We asked whether the mechanism by which L30 inhibits U2 binding is by interference with the binding of these two proteins to the intron. To this end, we used coimmunoprecipitation and protein cross-linking techniques on different transcripts represented in figure 19A.

We incubated the *RPL30 WT*, *+12* and *BS-U* (*WT* with point mutation in the BS) substrates with splicing extracts containing HA-tagged BBP protein and supplemented with MBP:L30. Immunoprecipitation against MBP was performed and copurification of BBP-HA was detected by Western blot analysis. Figure 19B shows that BBP can be coimmunoprecipitated with the IC (lane 2) as well as in our positive control (*+12*, lane 3). In contrast, the point mutation in the branch site sequence (*BS-U*) abolishes BBP binding (already described in Berglund, 1997). This experiment clearly demonstrates that L30 does not interfere with the binding of BBP to the branch site sequence in vitro. In addition, it indicates that L30 is not able to interact directly with BBP.

Furthermore, we followed Mud2p binding to the *RPL30* pre-mRNA in the IC by RNA-protein cross-linking. The following substrates: *RPL30 WT*, *+12*, *WT(1-47)* and *+12(1-47)* were labelled with α-^{32}P-CTP and 4-thiouridine. 4-thio-U increases the amenability of the RNA to crosslink with other RNAs or proteins. All these substrates were incubated with splicing extracts containing a TAP-tagged Mud2p protein and in the presence or absence of MBP:L30. After cross-linking the Mud2TAP protein was immunoprecipitated and subjected to SDS-PAGE gel to detect an interaction with the substrate of interest, in which case it will become radiolabelled. Figure 19C, shows that Mud2TAP protein can crosslink both to the *WT* and *+12* substrates, in the absence or presence of MBP:L30 protein (lane 1 to 4). Moreover, Mud2TAP is interacting with the 3'end region of the intron, because interaction is lost when short-version mutants were used as substrates (lanes 5 and 6).

Together, these data indicate that the recognition of the *RPL30* transcript by the components of commitment complex 1 (CC1) and commitment complex 2 (CC2) is not prevented by L30.

However, there is still the possibility that the CC2 formed with L30 is not in the proper configuration to incorporate U2snRNP, leading to the inhibition of the spliceosome assembly.

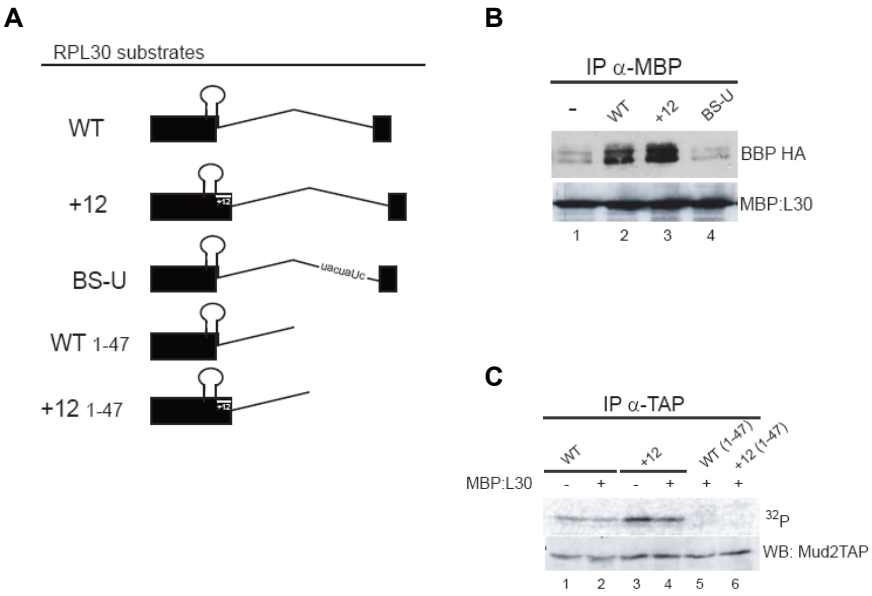

Figure 19. BBP and Mud2 associate with the *RPL30* intron during L30 repression of splicing.
Different substrates for these experiments were used (**A**). Schematic representation of *WT*, *+12* and mutant *BS-U* pre-mRNAs. *WT(1-47)* and *+12(1-47)* correspond to shorter versions of the *WT* and *+12* pre-mRNAs, respectively.
(**B**) BBP recognizes the branch site (BS) sequence in the IC. Extracts from a strain expressing HA-tagged BBP were supplemented with MBP:L30 and coimmunoprecipitated BBP-HA was observed in the *WT* substrate as well as in the positive control (*+12*). BBP fails to coimmunoprecipitate when the BS has the mutation A to U known to abolish BBP binding (UACUAuC) (lane 4), or in the absence of substrate (lane 1).
(**C**) Mud2p crosslinks to the *RPL30* intron in the IC. Mud2p interacts with *RPL30* substrate (lane 1) and this is not affected by L30 addition (lane 2), as well as in the *+12* positive control (lanes 3 and 4). The crosslink is lost when the RNA ends at position 14 in the intron (lanes 5 and 6). Western blot shows (bottom panel) that in all cases Mud2TAP protein was immunoprecipitated.

5. ALTERATIONS IN THE PROGRESS OF PRESPLICEOSOME FORMATION DO NOT AFFECT L30 INHIBITION

Critical components of the commitment complex 2 (CC2) seem to be present in the inhibited complex, although U2snRNP cannot be recruited. One explanation is that L30 hyperstabilizes cross-intron interactions disrupting U2 incorporation. For this reason, we asked if weakening of cross-intron interactions can bypass the L30 effect, and U2 recruitment can be observed.

Mud2p is a non-essential component of the interactors that are supposed to stabilize CC2. Extracts without Mud2p protein can be prepared from *mud2Δ* cells. Thus, we tested whether lack of Mud2p would alter L30 repression. Figure 20 shows that the inhibition of U2 recruitment, detected by lack of U2 co-immunoprecipitation with L30 in the presence of *RPL30* transcripts, was maintained in *mud2Δ* extracts (Figure 20B, lane 2). And, as expected, there was no effect of *mud2Δ* on the *+12* transcript (lane 3). Therefore, we concluded that the deletion of *MUD2* does not affect L30 inhibition of U2 snRNP recruitment.

Neither Mud2p nor BBP are detectable in yeast prespliceosome complex *in vitro* (Rutz and Seraphin, 1999). And before U2 snRNP is recruited, Sub2p has been proposed to remodel CC2, since deletion of *MUD2* suppresses lethality of *sub2Δ* (Kistler and Guthrie, 2001). Next, we hypothesized that L30 inhibits *RPL30* splicing by interfering with a remodelling or other function of Sub2p during assembly. In that case, deletion of *SUB2* would render L30 unable to inhibit. To test this possibility, extracts from the strain yCG472 (*mud2Δ sub2Δ*, (Kistler and Guthrie, 2001) were incubated with a synthetic RNA from *RPL30 WT* or *+12*, MBP:L30 was added and the co-immunoprecipitated U2 snRNA with MBP was analyzed. As Figure 20C shows, L30 could repress U2 recruitment in *mud2Δ sub2Δ* extracts (lane 2), while it had no effect on the *+12* substrate (lane 3). We concluded that the L30 repression of *RPL30* is not dependent on Sub2p.

We also asked if L30 could be interfering with the function of Cus2p protein, also important in U2 snRNP recruitment. As mentioned in the introduction, Cus2p negatively affects the

incorporation of U2 because, in its absence, ATP is no longer necessary for this step in vitro. To address the possible role of Cus2p in *RPL30* autoregulation, U2 snRNP coimmunoprecipitation was tested in extracts prepared from *cus2Δ* cells. As expected, incorporation of U2 to the *RPL30* intron was found to be independent of ATP (figure 20D, lane 3, *+12* substrate), consistent with previous reports (Perriman and Ares, 2000). But repression by L30 was still maintained in these extracts (figure 20D, lanes 2 vs. 3 and 4 vs. 5) both in the presence or absence of ATP. Consequently, our results point at a scenario where binding of L30 in exon 1, in close proximity to the 5'ss, blocks a key step in U2 recruitment, likely to be the base-pairing interaction.

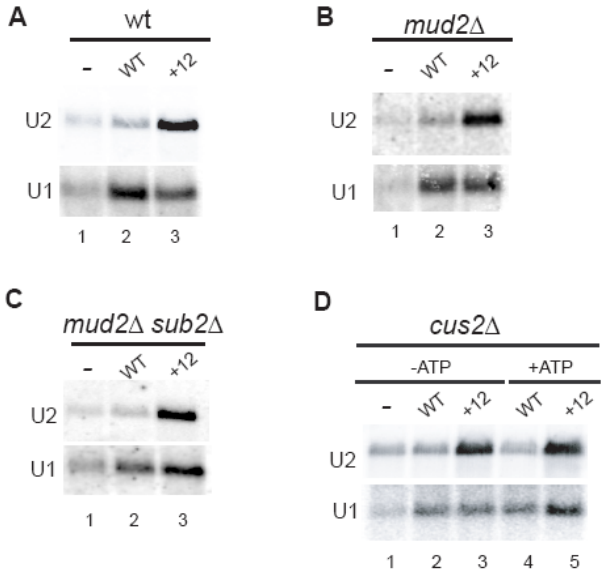

Figure 20. Control of U2snRNP recruitment by L30.

Spliceosome assembly was performed *in vitro* using the indicated extracts supplemented with MBP:L30, plus or minus ATP as shown. Reactions were immunoprecipitated using anti-MBP antibodies and subjected to Nothern analyses. The results for U2 (upper panels) and U1 (bottom panels) are shown. As a positive control for immunoprecipitation, the *+12* RNA was used. In all panels, lane 1 indicates a typical reaction but without any substrate added.

(**A**) L30 prevents U2 incorporation in wt extracts (lane 2) but not in a *+12* substrate (lane 3). (**B**) Repression of *RPL30* splicing by L30 does not require Mud2p, since association with U2 snRNA is still blocked (lane 2). (**C**) Inhibition of *RPL30* splicing by L30 does not require Sub2p. Coimmunoprecipitation of U2 snRNA in *mud2Δsub2Δ* extracts is blocked by L30 (lane 2), while it is efficient in the non-repressed *+12* substrate. (**D**) L30 can block the ATP-independent U2 association with the *RPL30* intron. U2 snRNA can join efficiently spliceosome assembly in extracts from *cus2Δ* cells with and without ATP (compare lanes 3 and 5). This association can be blocked by L30 in the *RPL30 WT* substrate (lanes 2 and 4).

By a totally different approach we asked if Prp5p function could be affected by L30 regulation. Prp5p is an essential RNA-helicase required for U2 snRNP recruitment (see figure 9, introduction). To assess the role of Prp5p in L30 regulation we took advantage of mutants developed in the lab of C. Query (Xu and Query, 2007). The mutants generated increased splicing efficiency of suboptimal RNA substrates containing mutations in the branch site. Therefore, we asked if these *prp5* mutant alleles can bypass the effect of L30 on U2 recruitment. Assuming that the CC2 formed in the presence of L30 is not suitable for U2 incorporation, we hypothesized that slowing down the Prp5p activity increases the chance for the transition between CC2 and pre-spliceosome.

We assayed *RPL30* splicing in the following *prp5* mutants strains: *prp5*-N399D, *prp5*-SAW and *prp5*-TAG. These mutant strains were transformed with the plasmid pMB15, containing *RPL30* intron fused to *CUP1* gene. The ability of these cells to grow in increasing copper concentrations is a result of the splicing efficiency of the construct, correlated with the level of mRNA.

Two of the *prp5* mutant forms used (SAW and TAG) were mutants in the SAT motif. SAT motif or motif III in DExD/H RNA helicases is thought to be required for coupling of ATP hydrolysis to U2-conformational change. While SAW mutant increases inhibition of splicing of a suboptimal substrate, the TAG mutant improves its splicing. In addition, we used the N399D mutant. Mutation is placed outside of the SAT motif, but it has been shown to be able to improve splicing of a suboptimal substrate (Xu and Query, 2007).

In normal conditions of growth, splicing of *RPL30-CUP1* reporter allowed cells to grow until 0.5-0.6mM concentration of copper in the media (figure 21, first column). When splicing of the same reporter was assayed in the mutant strains, the same behaviour was observed (figure 21, three next columns). In order to overexpress L30 *in vivo*, we constructed a plasmid (pSM35) that only contains the second exon of the *RPL30* gene under a constitutive and highly active promoter. Thus, this plasmid will produce L30 able to regulate the *RPL30-CUP1* construct, but not itself, as it lacks the exon 1. In this situation, overexpression of L30 clearly inhibits *RPL30-CUP1* splicing in all tested strains, as growth is only supported until ~0.1mM of Cu^{2+}. Thus, none of the *prp5* mutants tested could bypass the inhibitory effect of L30 on splicing, suggesting that Prp5p proofreading activity is not affected by L30.

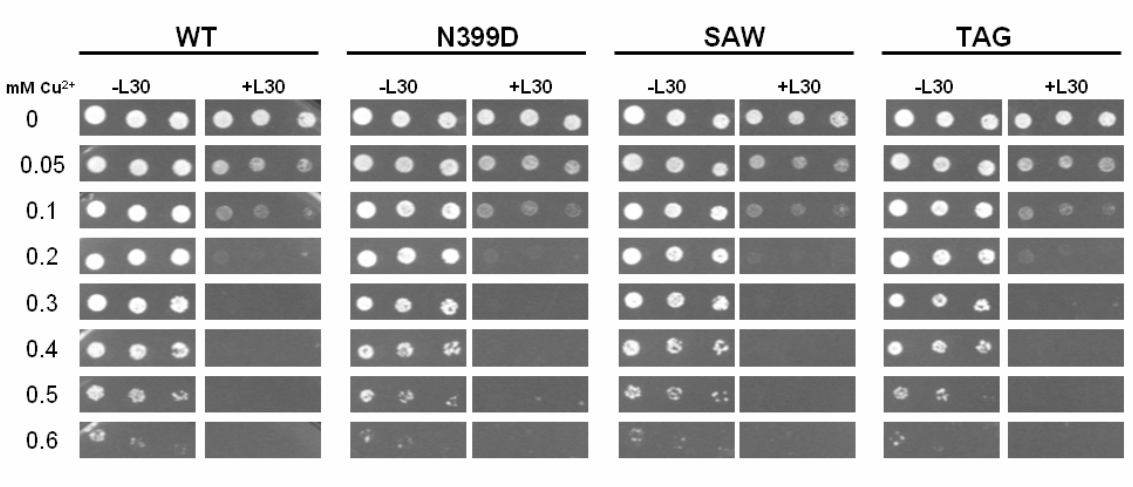

Figure 21. *prp5* mutants in ATPase domain cannot bypass L30 inhibitory effect on U2 recruitment.

In vivo splicing efficiency of the *RPL30-CUP1* pre-mRNA was monitored by cell growth in increasing copper concentration plates. Splicing efficiency was checked in a wt strain (WT), and three different *prp5* mutant strains (N399D, SAW, TAG). In all cases, in the absence of overexpressing L30 (**left column** for each strain) cells could support growth until 0.5-0.6mM of copper. When L30 was overexpressed (**right column**, indicated by "+L30") splicing of the reporter gene was repressed, as growth was only supported until ~0.1mM of copper.

6. MUD2, A ROLE IN SPLICING FIDELITY?

Mud2p protein needs to be displaced from the intron during the transition of the commitment complex 2 (CC2) to pre-spliceosome (PS). The remodelling action of Sub2p on Mud2p and, probably, BBP (Wang et al., 2008) will enable U2 base-pairing to the branch site sequence (Kistler and Guthrie, 2001).

During our analyses on the formation of the *inhibited complex* in different genetic backgrounds (deletion of *MUD2*, *SUB2* or *CUS2*) we observed a surprising effect of *mud2Δ* on U2 snRNP recruitment. Analogously to what happens in *cus2Δ* extracts, *mud2Δ* allowed partial U2 snRNP recruitment in the absence of ATP (figure 22A). The U2 ATP-independent recruitment was partial because it was not at the same extent as in the presence of ATP (figure 22A, compare lanes 3 vs. 6).

We next asked if the U2 ATP-independent recruitment observed in *mud2Δ* extracts was mediated by base-pairing. It could be possible that U2 was being recruited without reorganization of the CC2, where BBP would be still sitting on the branch site (BS). To answer this question we pre-treated extracts from *mud2Δ* cells with a 2'O-methyl DNA oligonucleotide against the U2 region that base-pairs to the BS. 2'O-methyl DNA oligonucleotides only block the accessibility to this region but do not activate RNaseH and degradation. The experiment in figure 22B clearly shows that the ATP-independent U2snRNP recruitment was mediated through a base-pairing interaction, or, in other words, that the region of U2 that base-pairs to the BS was necessary for the U2 recruitment in the absence of ATP in *mud2Δ* extracts.

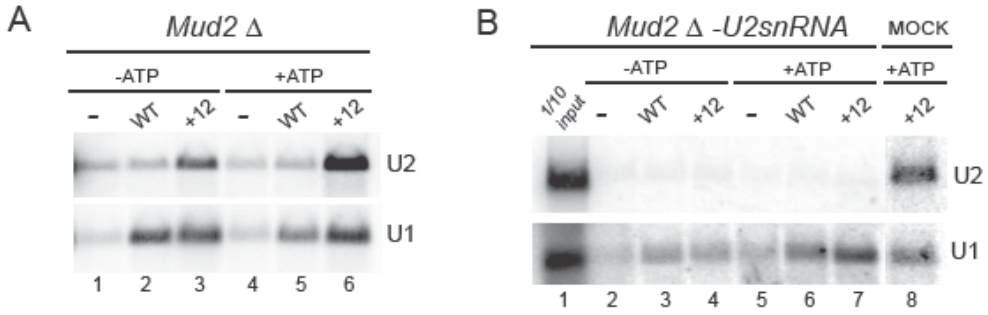

Figure 22. U2 recruitment in the absence of ATP in *mud2Δ* extracts.
(**A**) Deletion of the non-essential gene *MUD2* partially bypasses the necessity of ATP for U2 snRNP recruitment (lane 3), compare to the U2 that can be recruited in the presence of ATP (lane 6).
(**B**) Splicing reactions were preincubated with a 2'O-methyl oligonucleotide complementary to the U2 snRNP region that base-pairs to the BS. U2 recruitment, both in the presence or the absence of ATP (lane 4 and lane 7) is dependent on this region, suggesting recruitment through a base-pairing interaction. Lane 8 corresponds to a mock-treated sample, where U2 can be still coimmunoprecipitated.

This result is reminiscent of the situations where mutations on spliceosomal RNA-helicases change the constant of equilibrium between one stage of the spliceosome assembly and the next. In this particular case, the deletion of *MUD2* favours the formation of pre-spliceosomes (U2 binding), analogue to what happens with mutations in the ATPase motif of Prp5 (Xu and Query, 2007). If the CC2 and the pre-spliceosome complexes are in certain equilibrium, deletion of *MUD2* would move the equilibrium towards pre-spliceosome formation. Because Mud2p and BBP need to be recycled during this transition, in part by the action Sub2p, it could be possible that deletion of *MUD2* favours this transition through the destabilization of BBP in the BS

allowing U2snRNP base-pairing. If that is the case, analogously to Prp5, could have a role Mud2p in splicing fidelity? Could deletion of *MUD2* enhance splicing of suboptimal substrates for splicing?

To answer this question we performed *in vivo* studies that allow the measurement of the splicing efficiencies on different mutant substrates and the effect of the deletion of *MUD2* on splicing fidelity. As candidate substrates we chose mutant pre-mRNAs on the BS and 3'ss because Mud2p protein binds near these splicing signals.

The mutant pre-mRNAs were transformed in WT and *mud2Δ* cells. The splicing efficiency of the substrates was measured by growth in increasing concentrations of copper in the media. Figure 23 shows that for the group of BS mutants there were two types of behaviour: one, whose splicing was repressed upon deletion of *MUD2* (upper panel in figure 23B) and others, whose splicing was improved by deletion of *MUD2* (lower panel in figure 23B).

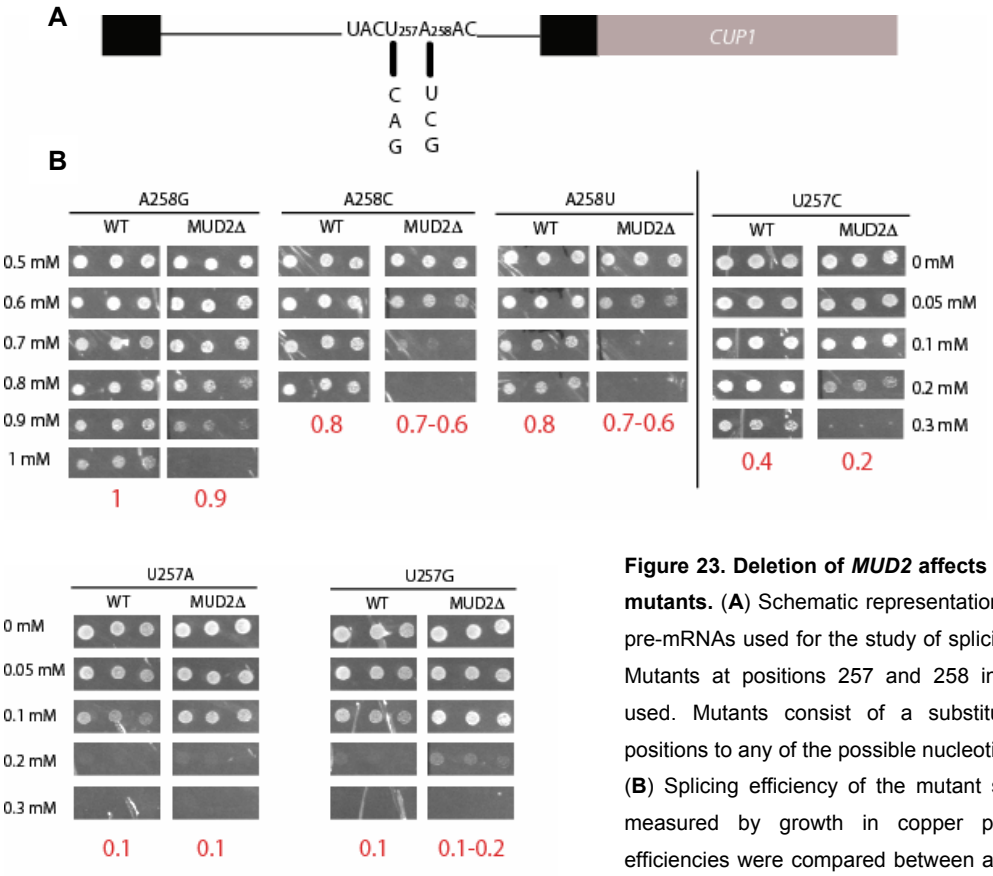

Figure 23. Deletion of *MUD2* affects splicing of BS mutants. (**A**) Schematic representation of the mutant pre-mRNAs used for the study of splicing efficiencies. Mutants at positions 257 and 258 in the BS were used. Mutants consist of a substitution of those positions to any of the possible nucleotides.

(**B**) Splicing efficiency of the mutant substrates was measured by growth in copper plates. Splicing efficiencies were compared between a WT strain (left column of every pair) and a *mud2Δ* strain (right column). In red, at the bottom part of every column, is the maximum concentration of copper in which cells can grow in this particular situation.

For the group of mutants whose splicing was affected upon deletion of *MUD2* gene (upper panel in figure 23B), we concluded that Mud2p protein was necessary for splicing of these pre-mRNAs. Surprisingly, another characteristic of this group was that also included those mutants whose splicing efficiencies were relatively good (growth in copper until 0.8-1 mM).

On the second group, we found those mutants whose splicing was improved by deletion of *MUD2*: mutants U257A and U257G (figure 23B, lower panel). In this particular case, the studied BS mutations were deleterious for splicing (only support growth at 0.1mM of copper) and deletion of the *MUD2* gene could only slightly improve this defect.

The mild enhancement in splicing efficiencies of suboptimal substrates by *mud2Δ* was expected. Mud2p would be acting on the first remodelling step of spliceosome assembly. During the whole spliceosome assembly process other proteins with proofreading activity (RNA helicases) will partially suppress the positive effect of *mud2Δ*.

In the same line of investigation we studied whether Mud2p could have a proofreading effect of the 3'ss. The putative ortholog of Mud2p in metazoans is $U2AF^{65}$ that works in a heterodimer together with $U2AF^{35}$. $U2AF^{35}$ has been proposed to bind to the 3'ss during the first steps of the spliceosome assembly (Wu et al., 1999). As no $U2AF^{35}$ homolog has been found in yeast, we asked if Mud2p could be also sensing the 3'ss in order to define the intron from the beginning of the spliceosome assembly cycle. Similar to the previous experiment, we measured splicing efficiencies on 3'ss mutants (only position -3, as it is depicted in figure 24A) in normal cells and *mud2Δ* cells.

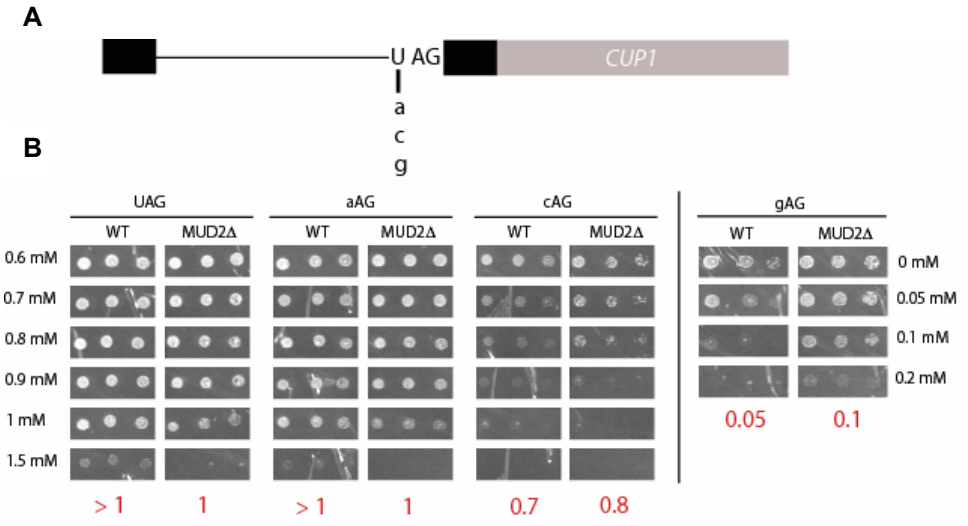

Figure 24. Deletion of Mud2 enhances splicing of 3'ss mutants.

(**A**) Schematic representation of the 3'ss mutant pre-mRNAs used for this study. Mutants consisted of a substitution of the -3 position of the intron to any of the possible nucleotides.

(**B**) Copper growth of the 3'ss mutants in a WT cell (left column of every pair) and *MUD2Δ* cells (right column of every pair). In red, maximum concentration of copper at which cells can support growth in every condition.

The canonical 3'ss is UAG. 3'ss substitutions, such as **a**AG or **c**AG, had mild effects on the splicing efficiency (compare UAG vs. **a**AG and **c**AG in WT cells, figure 24B) and *mud2Δ* did not significantly improve their splicing (right column of every pair in figure 24B). In contrast, the deleterious mutation **g**AG was significantly improved by *MUD2* deletion (from 0.05 mM to 0.1mM, figure 24B, column pair in the right). Suggesting that Mud2p could have a role in a proofreading activity affecting the 3'ss sequence. **g**AG mutants would escape the discarding pathway in *mud2Δ* cells, presumably because this deletion would allow a faster progression of the CC2 to the PS.

7. IN VIVO PURIFICATION OF THE *INHIBITED COMPLEX*

The TAP (tandem affinity purification) method was developed by the Séraphin lab (Puig et al., 2001) as a tool for rapid purification of complexes under native conditions. The TAP tag consists of two IgG binding domains of *Staphylococcus aureus* protein A (Prot A) and a calmoduline binding peptide (CBP) separated by a TEV protease cleavage site. Prot A binds tightly to an IgG matrix, requiring the use of the TEV protease to elute the material under native conditions. The eluate of the first affinity purification step is then incubated with calmodulin-coated beads in the presence of calcium. After washing, the bound material is released under mild conditions with EGTA. The protein of interest can be N-terminal TAP tagged, as well as C-terminal. TAP tag can also be split in two halves to tag two different components of the same complex. Therefore, we decided to use this method to purify the *inhibited complex* in vivo.

Proteins known to be present in the *inhibited complex* (IC) are L30 and U1 proteins (see figure 1D, results). Thus, we decided to fuse the ProtA tag to L30 and the CBP tag to Snu71p. We designed the experiment as follows: first, we constructed a plasmid as a source for overexpression of L30 containing the ORF of *Sulfolobus acidocaldarius* L30 fused to Prot A under the control of a galactose-inducible promoter (see figure 25). We used *Sulfolobus acidocaldarious* L30 for two reasons: first, because it inhibits *RPL30* splicing as *S. cerevisiae* L30 does, and second, because it cannot join the ribosome, thus avoids purifying ribosomes. (Vilardell et al., 2000b). However, overexpression of the SaL30 is lethal for the cell, because inhibits endogenous *RPL30* expression and, in consequence, no L30 is produced to form functional ribosomes. For this reason, we had to engineered a yeast strain whose endogenous *RPL30* gene contained the C9→T mutation, known to abolish splicing regulation by L30 *in vivo* and *in vitro* (see introduction, (Vilardell and Warner, 1994)). Endogenous *RPL30 C9→T* transcript will not be splicing inhibited and cells will survive in L30-overexpressiong conditions. For this reason, it became necessary to co-transform the cells with a reporter *RPL30* gene, where the inhibited complex (IC) would be assembled. We designed two different reporters that express constitutively the *RPL30* regulatory element (exon 1 and intron) fused to *GFP* cassette as exon 2. We made a *RPL30 WT* version, where the IC will be assembled, and a *RPL30 T9*

version, as a negative control, because its splicing is not L30-regulated (see figure 14 introduction, (Vilardell and Warner, 1994)).

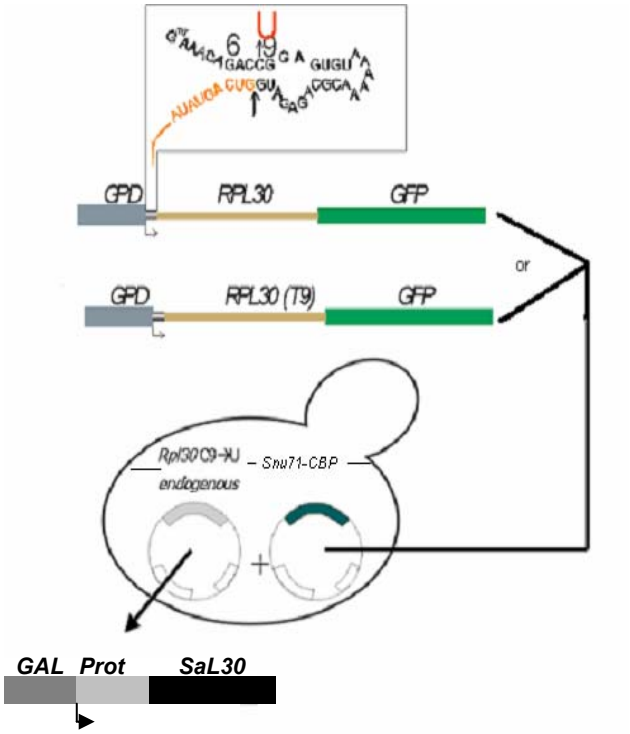

Figure 25. Model of the strategy followed for *in vivo* purification of the *inhibited complex*. Split TAP tag strategy was followed. ProtA-SaL30 is overexpressed under galactose growing conditions. *Inhibited complexes* are first purified using Prot A tag on SaL30. After TEV cleavage, the complex is repurified using calmoduline resin against the endogenous Snu71p protein CBP tagged. Purified complexes will be assembled on a *RPL30 WT-GFP* pre-mRNA or on a *RPL30 T9-GFP* pre-mRNA, as a negative control for lack of regulation.

We took advantage of the *GFP* reporter gene to check that the system generated was working as expected (figure 26). For that reason we grew cells that contain either the *RPL30 WT-GFP* reporter or the *RPL30 T9-GFP* reporter, in glucose (non-overexpressed SaL30) or galactose-containing media (overexpressed SaL30). This allowed us, qualitatively, to determine if splicing of the reporter pre-mRNAs was being repressed in regulating conditions.

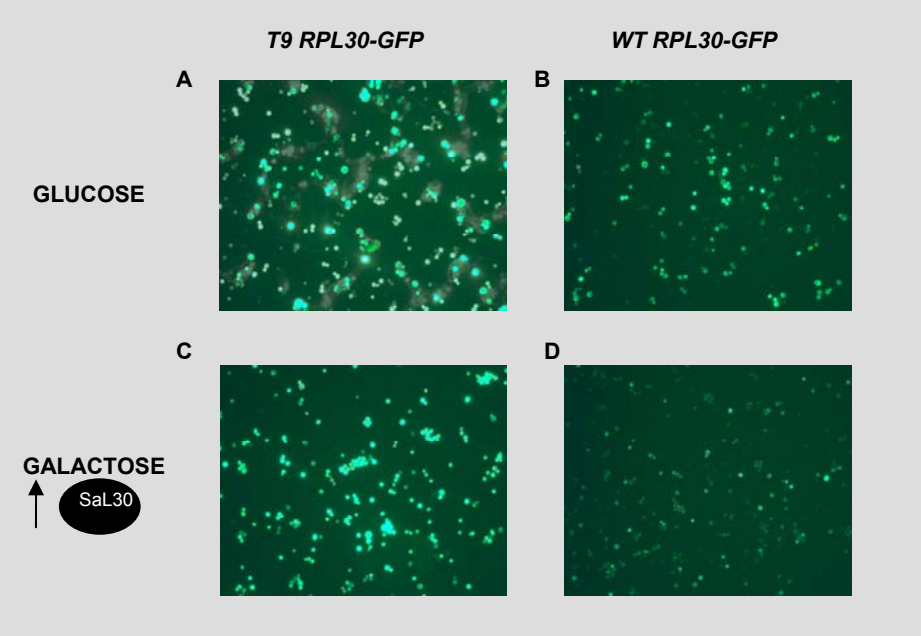

Figure 26. GFP expression in TAP-tagged strains. (A) Cells containing the *T9* version of the *RPL30-GFP* reporter show high expression of GFP in normal growth conditions. Overexpression of SaL30 (**C**) does not change GFP signal, indicating the splicing of the reporter gene is not affected.
(**B**) Cells containing the *WT* version of the *RPL30-GFP* reporter are less GFP positive compared to *T9* construct (compared B vs. A). (**D**) Overexpression of SaL30 by galactose-containing media reduces GFP expression almost to background levels.

GFP expression was a clear reflect of the steady-state level of pre-mRNA and mRNA accumulation in growing cells detected by northern analyses (figure 27). Galactose-growing conditions induced overexpression of SaL30 (figure 27, lower panel, lane 2 and 4). Overexpression of SaL30 induced the accumulation of *RPL30 WT-GFP* pre-mRNA (figure 27, upper panel, lane 2), although in basal conditions (glucose media) some accumulation could also be detected (figure 27, upper panel, lane 1). In contrast the *RPL30 T9-GFP* pre-mRNA did not accumulate in any of the growing conditions, either glucose or galactose (figure 27, upper panel, lane 3 and 4, respectively).

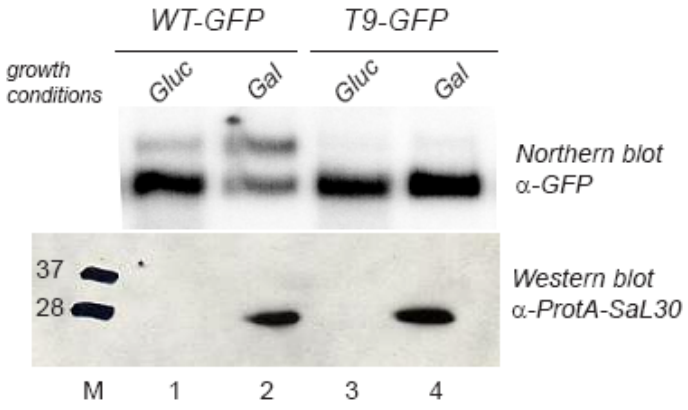

Figure 27. Splicing of the *RPL30 WT-GFP* transcript is inhibited upon overexpression of SaL30 protein. Two different strains containing either the *WT* or the *T9* copy of the *RPL30-GFP* reporter gene where grown in glucose or galactose media. Overexpression of SaL30 induces *RPL30 WT-GFP* accumulation (lane 2, upper panel) whereas *RPL30 T9-GFP* remains unaffected (lane 4, upper panel). **p** stands for precursor or pre-mRNA, **m** stands for mRNA. Lower panel, monitoring by western blot SaL30 overexpression.

An important observation from these experiments is that pre-mRNA accumulation is already detected with the *RPL30 WT-GFP* construct in normal conditions (non-overexpressing-SaL30 conditions). This observation is the consequence of using a strain that contains the endogenous *RPL30* gene C9→T mutated. This mutation abolishes L30 feedback control, and in consequence overexpresses L30. Therefore, part of the inhibited complexes that are being formed in living cells will not be purified, as endogenous L30 will be competing with overexpressed ProtA-SaL30.

The next step was to do purification at large scale (2 liters of saturated cultures). We first tried to do one-step purification, the IgG column purification for ProtA-SaL30. We compared 3 different situations: cells containing the *RPL30 WT-GFP* reporter grown in glucose, the same cells grown in galactose (overexpressing SaL30) and finally, cells containing *RPL30 T9-GFP* grown in galactose conditions (figure 28). After the first purification and TEV cleavage, only minor differences were observed between the three conditions. This means that the background of the purification was high, and not many proteins differed from the complexes assembled on the *WT* or *T9* reporters, indicating that they were not specific for the *inhibited complex*.

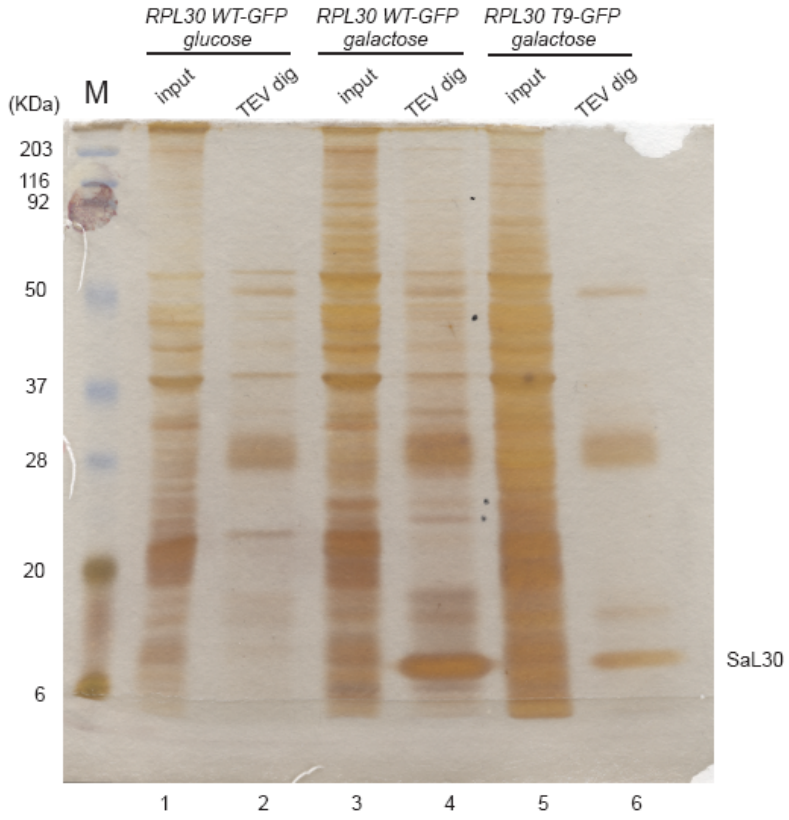

Figure 28. IgG affinity purification on WT and T9 transcripts. Cells containing the *RPL30 WT-GFP* transcript were grown either in glucose or galactose. Cells containing the *RPL30 T9-GFP* construct were grown in galactose. Extracts were prepared from the three conditions and IgG purified. After TEV cleavage, the eluted material was loaded on a 15% SDS-PAGE gel and silver-stained. TEV material from a mock purification (cells grown in glucose) was loaded on lane 2. TEV digestion of the *inhibited complex* was loaded in lane 4. In lane 6, TEV digestion of the material assembled on the *RPL30 T9-GFP* was loaded. Almost no differences were observed between lane 2, 4 and 6, except for the bands that have a dot beside.

We also performed two-step purifications, where the TEV digested material was purified through a calmodulin resin. In this case, most of the material was lost. We could only detect coimmunoprecipitation of the SaL30 with the Snu71-CBP protein (by silver-staining gels, data not shown). Although we knew that calmodulin purification worked in our hands (by calmodulin binding followed by western blot to see depletion of Snu71CBP), it gave very low yield as a second step of purification. Probably, most of the SaL30 that it was being purified by the first step was not in the same complex with Snu71CBP. This agrees with the result from figure 27, where it is shown that TEV digestions do not differ too much between the different conditions.

We concluded that the most of the complexes formed with SaL30 were unspecific and not related to splicing inhibition. Probably, the fact that part of the *RPL30 WT-GFP* pre-mRNA is already repressed in normal conditions lowers the yield of purification of the IC containing SaL30.

We could have designed other methods. One possibility was to change the calmodulin purification by a purification of the pre-mRNA (biotinylated or with MS2-binding sites). In any case, we concluded that the possible formation of the IC *in vivo* was at very low yield. We needed to increase the amount of cells to quantities difficult to manage. Currently, a person from the lab is doing the purification from an *in vitro* assembly and works with better yields.

8. COTRANSCRIPTIONAL SPLICEOSOME ASSEMBLY IS REGULATED BY L30

Chromatin immunoprecipitation technique (ChIP) has been successfully used to follow cotranscriptional spliceosome assembly. Reports from the Rosbash and Neugebauer labs (Gornemann et al., 2005; Lacadie and Rosbash, 2005) show that spliceosome assembly can be monitored cotranscriptionally, as well as splicing itself (Tardiff et al., 2006). For that reason, we chose this technique to analyze and verify *in vivo*, our *in vitro* results. ChIP allows determining binding of a spliceosomal component to a nascent pre-mRNA, as well as the position of binding. Moreover, kinetics of spliceosome assembly can also be studied with this technique; binding and release can be inferred from amplification of the immunoprecipitated DNA at different positions within the gene, assuming that distance is equivalent to time and that epitope availability does not change during the assembly process (reviewed in (Nilsen, 2005).

For these reasons, we studied how cotranscriptional spliceosome assembly on the nascent *RPL30* transcript was modified by L30. First, we monitored spliceosome assembly (U1 and U2 snRNPs) on the endogenous *RPL30* gene (fig. 29A). U1 snRNP cotranscriptional assembly was followed using a strain whose endogenous *SNU71* gene was biotin-tagged allowing protein immunoprecipitation with streptavidin beads. In addition, we monitored U2 snRNP assembly by using strain whose endogenous *LEA1* gene (U2 component) was also biotin-tagged (see methods). The DNA coimmunoprecipitated with these two biotin-tagged proteins was amplified and analyzed by real time PCR, using the primers pairs depicted in figure 29A. In addition, we monitored L30 cotranscriptional binding by cotransforming cells with a plasmid as a source of overexpressing L30 fused to the TAP tag under the control of a galactose inducible promoter. By this method, we first determined U1, U2 snRNP and L30 ChIP profiles on the endogenous *RPL30* transcript, in conditions with or without L30 excess. In normal conditions U1 snRNP is cotranscriptionally recruited to the nascent pre-mRNA, and overexpression of L30 does not prevent its incorporation (figure 29B). Furthermore, we observed that overexpressing L30 leads to a reduction on U2 snRNP recruitment, although the basal U2 levels are low to observe significant changes (figure 29C). Importantly, we could also detect cotranscriptional binding of L30 itself to the pre-mRNA. (figure 29D).

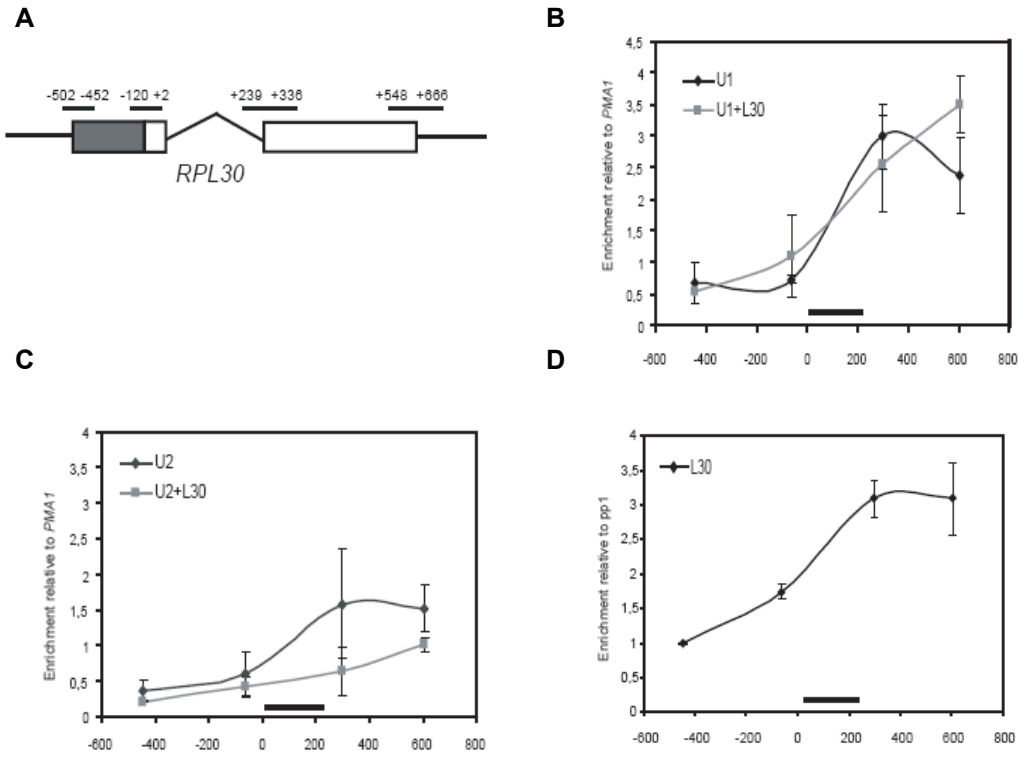

Figure 29. Spliceosome assembly and L30 binding on the nascent *RPL30* intron.

The x axis show the distance in nucleotides from the translation initiation site. The black bar indicates the location of the intron. Non-repressive conditions (absence of overexpression of L30) are indicated by a black line, while repressive conditions (excess of L30) are indicated by a grey line.

(**A**) Scheme showing the positions of the primers for the quantitative PCR analysis of the chromatin immunoprecipitated DNA (relative to the translation start site) on the endogenous *RPL30* gene. (**B**) ChIP against U1snRNP (Snu71-HTB). (**C**) ChIP against U2snRNP (Lea1-HTB). (**D**) ChIP against L30 (L30-TAP).

Tardiff and Rosbash (Tardiff et al., 2006) showed that full cotranscriptional assembly and splicing can only occur in genes containing a long second exon. They proposed that the majority of splicing happens post-transcriptionallly, as is the case of *RPL30* gene. Thus, in order to better define the cotranscriptional effects of L30 on later steps of spliceosome assembly (especially U2 binding), the *RPL30* intron was fused to the *LACZ* gene as a second exon. First, we checked by RTPCR (retrotranscription-PCR) that splicing of the chimaeric construct can also be repressed by overexpression of L30 (figure 30). In addition, we constructed a mutated

version of the *RPL30-LacZ* gene (named *RPL30*LacZ*) for binding of L30 that cannot be regulated (as shown figure 30).

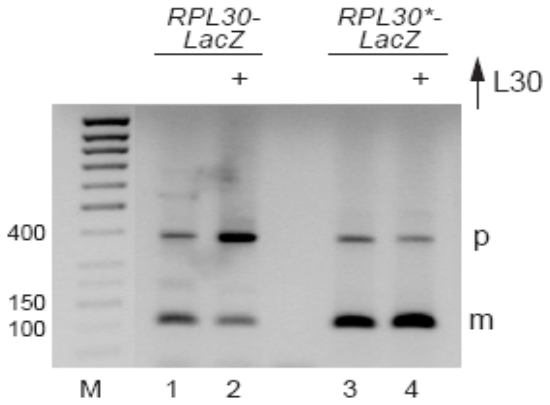

Figure 30. *RPL30-LacZ* splicing is repressed by L30.
RPL30-LacZ reporter transcript shows accumulation of pre-mRNA when L30 is overexpressed (lane 1 vs. 2). *RPL30*LacZ* transcript remains unaffected after overexpression of L30 (lane 3 vs. 4). "p" and "m" denote the PCR products from the precursor and mature transcripts, respectively.

The fusion of the *RPL30* intron to the *LACZ* cassette increased the size of the second exon by 3KB. The increase in length significantly improved ChIP signals (figure 31). In this case, we observed that overexpression of L30 leads to retention of U1 snRNP (figure 31B), and inhibition of U2 snRNP recruitment (figure 31C) as well as for U5 snRNP (figure 31D). In addition, we determined that L30 was able to bind to the nascent chimaeric transcript (figure 31E). The retention observed for U1 agrees with the one reported by Tardiff et al (Tardiff and Rosbash, 2006), with cells that have been depleted of U2snRNP particle.

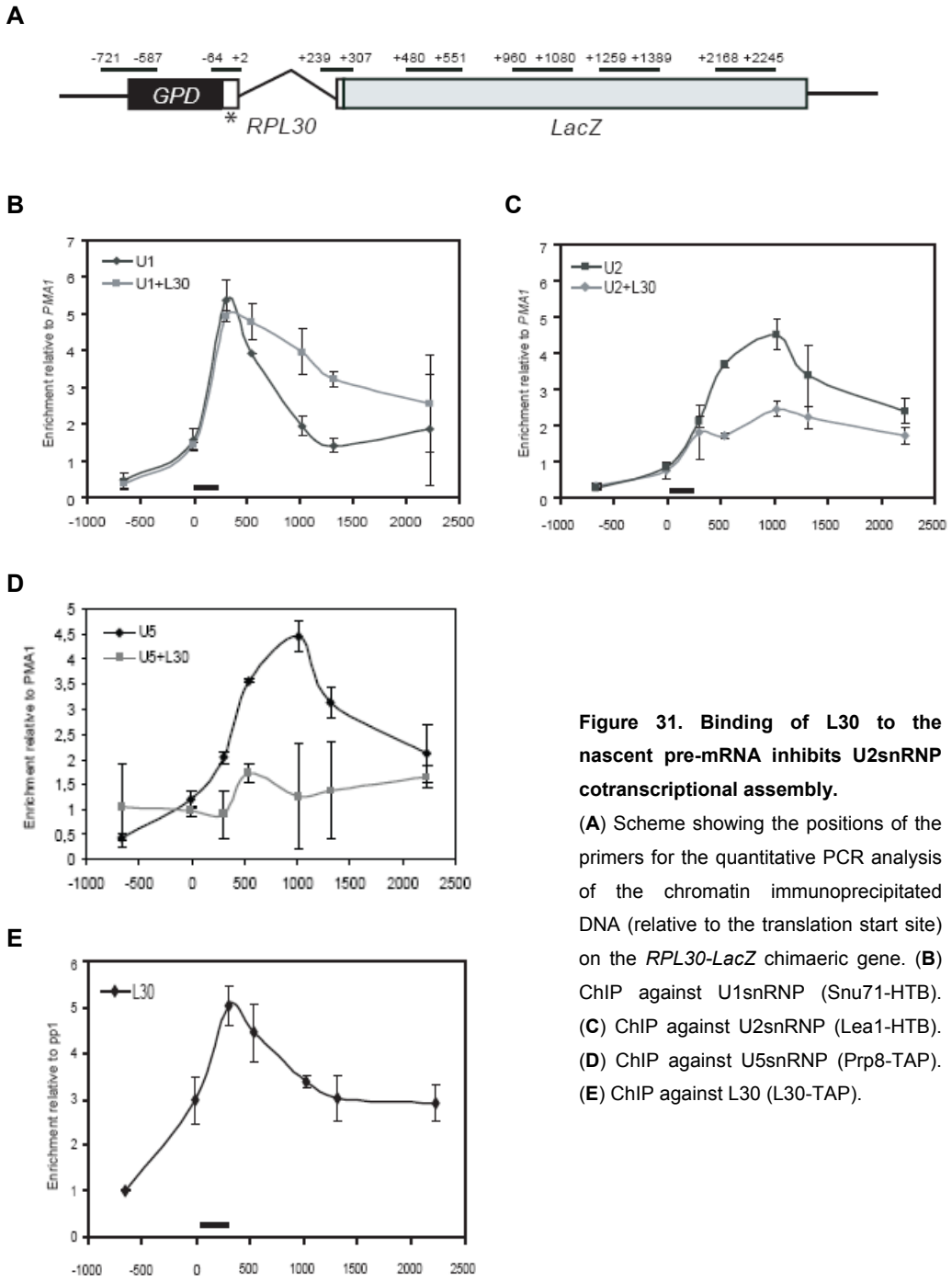

Figure 31. Binding of L30 to the nascent pre-mRNA inhibits U2snRNP cotranscriptional assembly.

(**A**) Scheme showing the positions of the primers for the quantitative PCR analysis of the chromatin immunoprecipitated DNA (relative to the translation start site) on the *RPL30-LacZ* chimaeric gene. (**B**) ChIP against U1snRNP (Snu71-HTB). (**C**) ChIP against U2snRNP (Lea1-HTB). (**D**) ChIP against U5snRNP (Prp8-TAP). (**E**) ChIP against L30 (L30-TAP).

These results showed that L30 can affect cotranscriptional assembly *in vivo* by binding to the nascent pre-mRNA. No effects were observed when ChIPs were performed on a mutant pre-mRNA that cannot be bound by L30 (*RPL30*LacZ,* figure 32A and B). As expected, L30 binding was significantly reduced (figure 32C).

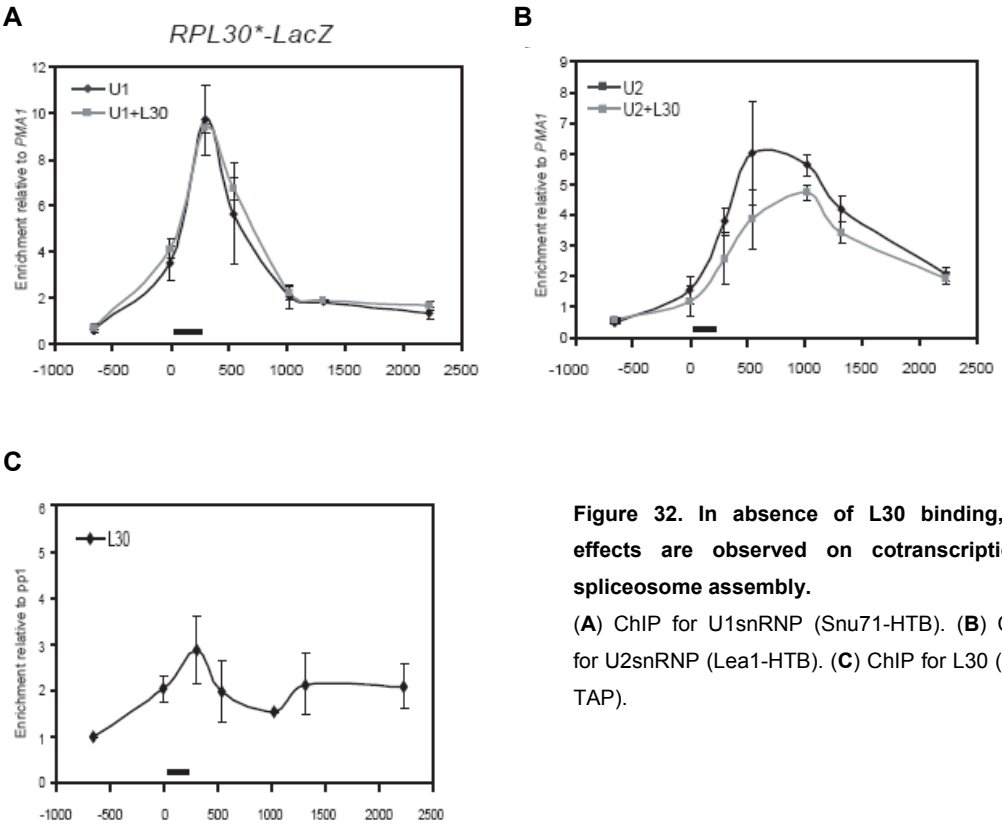

Figure 32. In absence of L30 binding, no effects are observed on cotranscriptional spliceosome assembly.
(**A**) ChIP for U1snRNP (Snu71-HTB). (**B**) ChIP for U2snRNP (Lea1-HTB). (**C**) ChIP for L30 (L30-TAP).

As a parallel control for specificity of the L30 effects on spliceosome assembly, we followed cotranscriptional assembly on the endogenous *ACT1* gene (actin), in absence and presence of overexpressing L30. As expected, overexpression of ribosomal L30 protein did not change either U1 or U2 assembly patterns (figure 33B and C).

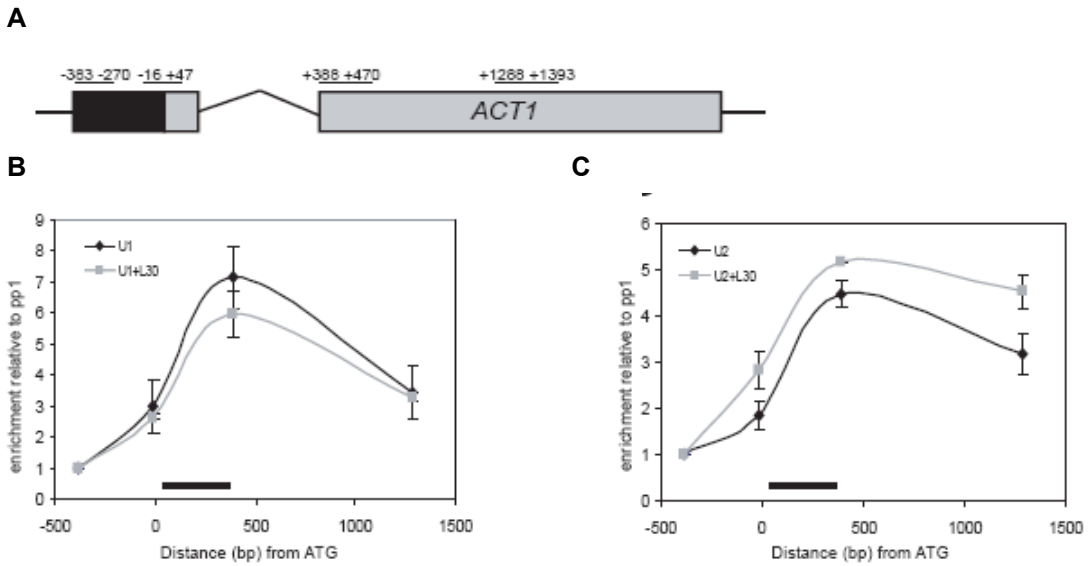

Figure 33. U1 and U2snRNP cotranscriptional recruitment on the *ACT1* transcript is not affected by overexpression of L30.

(**A**) Scheme of the *ACT1* gene and relative position to the translation start codon of the primers used for quantitative PCR analyses. (**B**) ChIP for U1snRNP (Snu71-HTB). (**C**) ChIP for U2snRNP (Lea1-HTB).

9. MIMICKING L30 INHIBITION OF SPLICING

From *in vitro* immunoprecipitation experiments and ChIP we knew that U1 is hyperstabilized in the *inhibited complex* and retained. Our data suggested that U1 snRNP hyperstabilization on the intron could be necessary and sufficient to trigger repression of U2 assembly, as L30 does. We hypothesised that if this is the cause for inhibiting U2 recruitment, then U1 hyperstabilization in a non-regulable pre-mRNA by L30 could also be stalled at the same step. We asked whether U1 stabilization is a requirement for L30 inhibition or instead it is a consequence of it. For that reason, we increased the number of potential base-pairs to U1 snRNA of the *+12* pre-mRNA that binds L30 but its splicing is not inhibited. The *+12* pre-mRNA can form 8 potential base-pairings with U1, and this number was increased to 10 to ask if this modification enables the substrate to be regulated by L30. We assembled *in vitro* splicing complexes on both pre-mRNAs (*+12* and *+12 extU1*, see figure 34) and we analyzed U2 recruitment by coimmunoprecipitation with MBP-L30. Figure 34 shows that this U1 hyperstabilization at the 5'ss is not enough to inhibit U2 snRNP binding (lane 7 vs lane 8). Considering this result, we concluded that hyperstabilization of U1 in the *inhibited complex* is likely to be the consequence of L30 inhibition of splicing but not the reason.

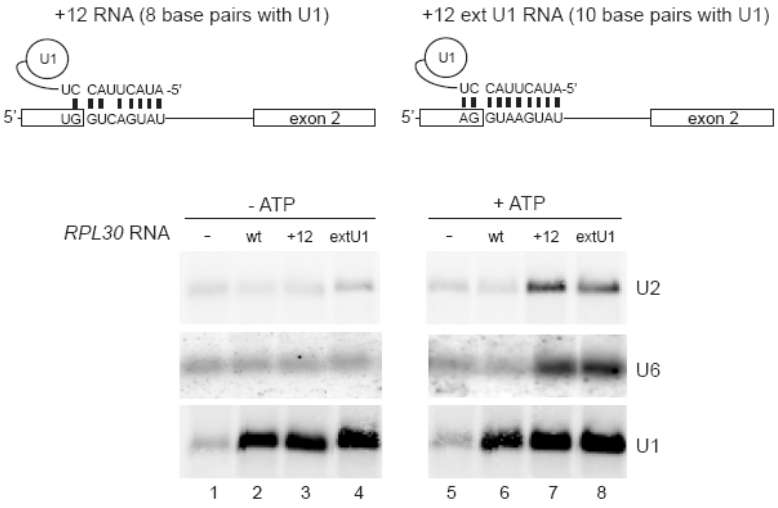

Figure 34. Hyperstabilization of U1snRNA association with the 5'ss does not mimic repression of splicing by L30.

The *RPL30 +12* transcript base pairs with U1 as indicated in top left. To produce a variant with extended base pairing with U1, mutations were introduced as shown at the top right. U2 can still associate with the *+12 extU1* transcript (lane 8, top panel). In addition there is no lack of U6 interaction with this transcript either, indicating that in our conditions the 5'splice site of the *RPL30* intron is properly recognized even 10 nucleotides susceptible to base pair with U1 snRNA.

In addition we asked if other protein binding to the *RPL30* pre-mRNA near the 5'ss could have the same effect as L30. In this particular case, we chose the MS2 protein because it is a small protein, binds with high affinity to RNA, and also binds a RNA secondary structure, as L30 does. We placed the MS2 stem-loop in front of the *RPL30* and *ACT1* intron (as shown in fig 35), containing part of the 5'ss in the RNA secondary structure. For *in vitro* splicing gels we used *ACT1* intron because it is a better substrate, and can also recapitulate L30 inhibition by adding a L30 binding site (data not shown).

Figure 35 shows that binding of MS2 protein inhibits splicing. And analogously to the *RPL30*/L30 system, displacement of the 5'ss 12 nucleotides downstream allows splicing to proceed, as we observed in the *RPL30* transcript.

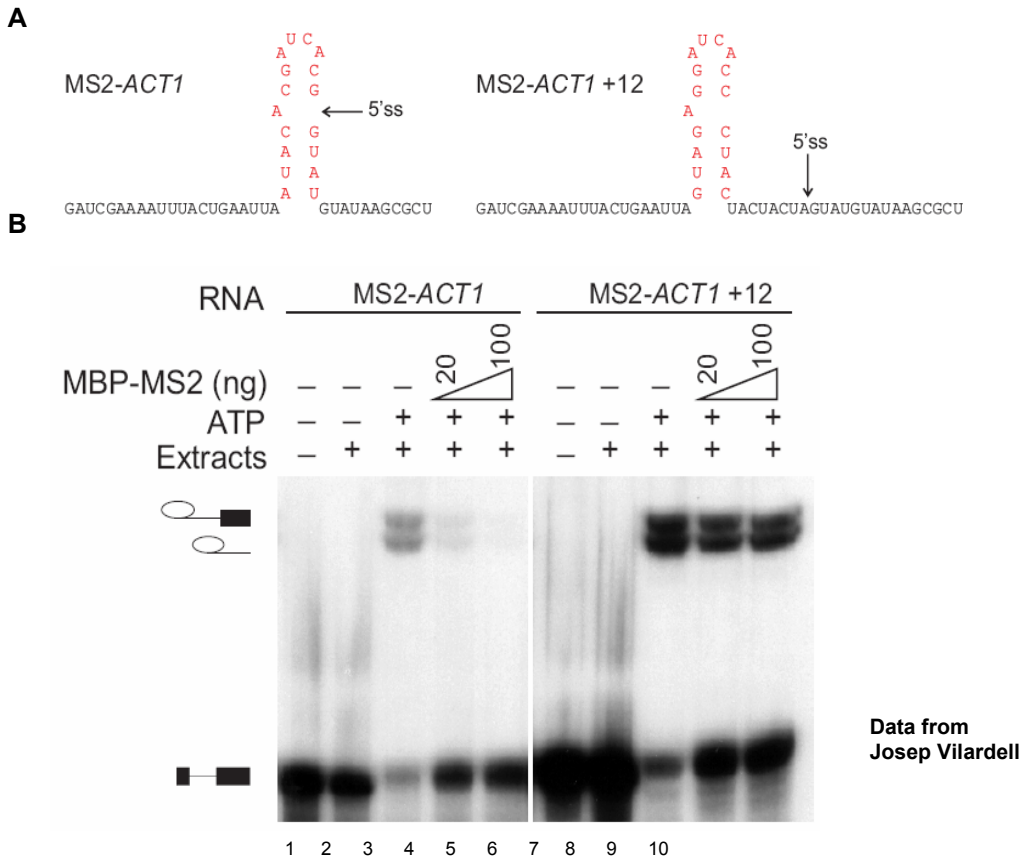

Figure 35. MS2 binding in the proximity of the 5'ss inhibits splicing.

(**A**)MS2 binding site was placed in front of the *ACT1* intron (MS2-*ACT1*) containing the 5'ss, or moving this sequence 12 nucleotides downstream (MS2-*ACT1+12)*.

(**B**) *In vitro* splicing of the MS2-*ACT1* (lanes 1-5) and MS2-*ACT1+12* (lanes 6-10). Similarly to the *RPL30*/L30 interaction, binding of MS2 blocks splicing, unless the 5'ss is at a distance from the protein binding site.

As MS2 binding to the pre-mRNA near the 5'ss inhibits splicing, we had to determine if the mechanism was similar to L30. For that purpose, we performed coimmunoprecipitation experiments to detect which are the snRNAs that copurify with the MS2/pre-mRNA complex. In this experiment we used a chimaeric RNA containing the MS2 binding site instead of the L30 binding site in front of the *RPL30* intron (figure 36A). In addition we used the pre-mRNAs mentioned above, that are constructed on the *ACT1* intron, to demonstrate that the MS2 binding effects are not intron-specific.

By this approach, we demonstrated that MS2 inhibition of splicing was at a different step of spliceosome assembly. Figure 36B, shows that while the normal *inhibited complex* (*RPL30*/L30) can only recruit U1 snRNA (panel B, lane 4), MS2 blocks this step, as no U1snRNA is detected in these conditions (panel B, lane 2). The same was observed for the *ACTIN* intron. While MS2+12*ACT* can recruit U2 and U1 snRNA in the presence of MS2 protein (panel C, lane 4), MS2*ACT* cannot recruit any of those particles (panel C, lane 2). These results argue for a totally different mechanism of inhibition. MS2 would be already blocking the first step of spliceosome assembly, possibly by steric hindrance at the 5'ss.

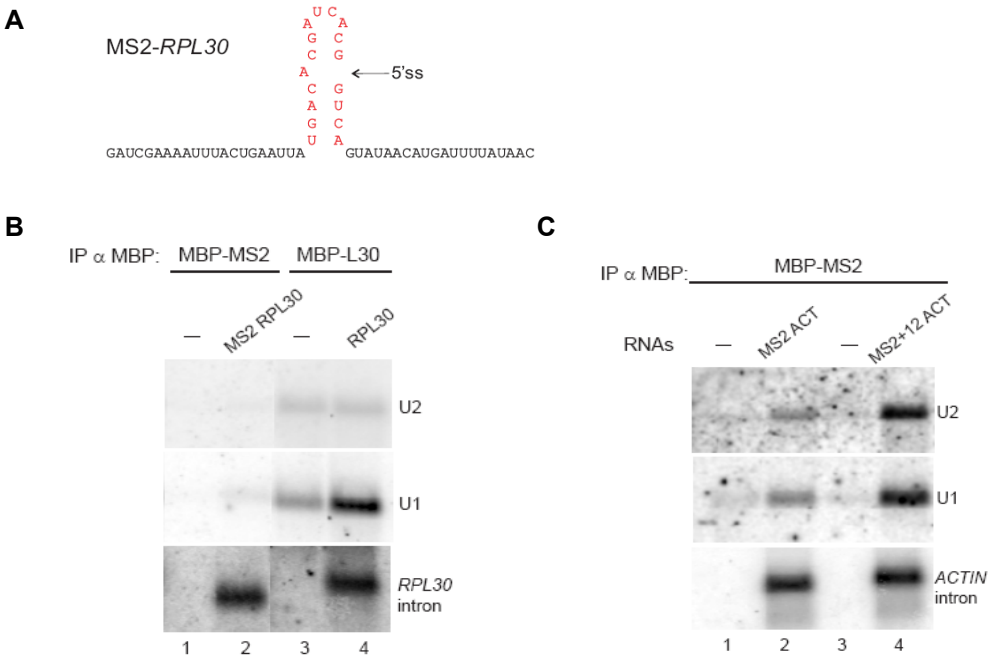

Figure 36. MS2 binding near the 5'splice site blocks transcript recognition by U1snRNA.

(**A**) MS2 binding site was also inserted in front of the *RPL30* intron, containing the 5'ss. (**B**) *RPL30*/L30 complexes can only coimmunoprecipitate U1 snRNA (lane 4), in contrast binding of MS2 protein prevents U1 snRNA association (lane 2). (**C**) U1 and U2 snRNAs can be coimmunoprecipitated with MBP:MS2 under conditions of lack of repression (lane 4). Both coimmunoprecipitations are greatly reduced when MS2 binding leads to splicing repression with MS2-*ACT1* substrate (lane 2).

DISCUSSION

In *S. cerevisiae* only 5% of the nuclear genes contain introns, the majority being monointronic. Comparison of introns from hemyascomycetus yeasts suggests that introns have been massively lost during evolution. However, both intron losses as well as gains might have occurred during this process (Bon et al., 2003; de Souza et al., 1998; Fink, 1987).

Thus the question is why a small number of genes have retained introns. In these cases, the surviving introns may confer a selective advantage to their host. In *S. cerevisiae*, mRNAs corresponding to intron-containing genes account for ~26% of all mRNA transcripts, and 90% of them are derived from genes encoding for ribosomal proteins (Ares et al., 1999; Spingola et al., 1999). The elevated proportion of introns residing in ribosomal protein genes (41%) suggests that introns are required to maintain and control the high ribosome biogenesis (Vinogradov, 2001). In addition, the presence of introns offers an additional level of post-transcriptional regulation, the pre-mRNA splicing.

The rate of accumulation of each ribosomal protein is carefully regulated by the cell to provide the equimolar ratio necessary for the assembly of the ribosome (Gorenstein and Warner, 1976; Udem and Warner, 1972). Yeast has developed different methods to control the accumulation of the products of ribosomal genes, such as translational regulation (*RPL3* and *RPL28*), mRNA processing regulation (*RPL30* and *RPL28*) or increased protein turnover (*RPS7, RPS10A* and *B*) (Warner et al., 1985).

The paradigm of the regulation of ribosomal protein synthesis is in eubacteria where the binding of a ribosomal protein structure in the mRNA blocks translation (Nomura et al., 1984). The interactions of individual ribosomal proteins with specific operons has been widely conserved among related eubacteria (Allen et al., 1999). In some cases, the mRNA structure resembles that of the site in rRNA to which the protein binds (reviewed in (Zengel and Lindahl, 1994). Similarly, yeast L30 regulates splicing of its own pre-mRNA by binding to a RNA structure resembling its rRNA binding site (Halic et al., 2005; Vilardell et al., 2000b).

In this thesis I have investigated the *RPL30* regulated intron in yeast that fails to assemble into spliceosome upon binding of a regulatory factor (L30) to the upstream exon. We have shown that the resulting inhibited spliceosomal complex cannot recruit U2 snRNP and that the regulatory factor binds and acts co-transcriptionally.

IN VITRO STUDY OF L30 REGULATION

It was previously reported that L30 can control the *RPL30* transcript at multiple levels, inhibiting spliceosome assembly at an early stage (Vilardell et al., 2000a) as well as reducing its translation (Dabeva and Warner, 1993). Although the RNA secondary structure seems to be highly conserved, there is much evidence that support the flexibility in the requirements to repress splicing. First, the 5'ss of *RPL30* does not need to be included in the kink-turn for splicing inhibition (figure 15, introduction). Second, the geometry of the binding site in *RPL30* can be altered without affecting inhibition (by swapping the strands of the asymmetric purine-rich bulge in the K-turn, data not shown). Third, a distant ortholog of L30 (*Sulfolobus acidocaldarius)* regulates *RPL30* indistinguishably, *in vivo* and *in vitro* (Vilardell et al., 2000b). Fourth, fusion of the kink-turn of *RPL30* upstream of other introns is sufficient to render these transcripts sensitive to L30 *in vitro* (data not shown), consistent with previous results using chimaeric genes containing parts of *RPL30* (Vilardell and Warner, 1997). Taken together, these evidences indicate that neither intronic sequences nor the precise configuration of L30 binding are determinant for inhibition. Nevertheless, repression cannot be recapitulated with a heterologous protein (figure 36, results). Binding of MS2 protein upstream of the *ACT1* or *RPL30* intron seems to completely block initial intron recognition (U1 binding). It is probable that the geometrical conformation of the MS2/*MS2-RPL30* complex interferes with the initial base-pairing interaction, as displacement of the 5'ss sequence 12 nucleotides downstream abolishes this effect (figure 36C, lane 4, results).

One explanation for L30 inhibition is that it prevents proper base-pairing of U1 with the 5'ss. It has been reported that U1 snRNP can initially bind the pre-mRNA in the absence of base-pairing through U1-C interaction, that later will be replaced by the canonical U1 snRNA/5'ss base-pairing (Du and Rosbash, 2002). Thus L30 would not prevent initial recognition which is in agreement with previous results (Vilardell and Warner, 1994), but rather U1 base-pairing interaction. However, L30 does not appear to be contacting either the G61 or G62 nucleotides (5'ss, figure 13, introduction) in the L30/*RPL30* complex (Chao and Williamson, 2004) arguing against this hypothesis. Furthermore, our psoralen cross-linking analyses indicate that the 5'ss of *RPL30* is base-paired with U1 snRNA during repression of splicing by L30 (figure 16C,

results). This is consistent with the requirement of the 5'end sequences of U1 for *in vitro* formation of the *inhibited complex* (Vilardell and Warner, 1994) and prompted us to assess the recognition of the rest of the *RPL30* intron during regulation (figure 19, results). From these experiments we conclude that L30 cannot inhibit the formation of the CC2 *in vitro*, because BBP and Mud2p are detected in the complex. Thus, repression might be occurring on steps involved in recruitment of U2 snRNP, while U1 remains stably associated with the intron. Moreover, the detection of Mud2p cross-linking to the *RPL30* intron (figure 19C, results) indicates that the splicing pathway is likely to follow the "Sub2-Mud2-dependent" pathway proposed by Kistler and Guthrie (Kistler and Guthrie, 2001). Since the cross-linking is maintained in the presence of L30, this pathway might not be altered. For that reason, we investigated how the binding of L30 could affect the steps involved in the formation and progression of the commitment complex to the pre-spliceosome.

First, we asked whether L30 represses *RPL30* splicing by stabilization of the CC2 complex, which is held together by a set of cross-intron interactions that include U1 snRNP (Prp40p), BBP and Mud2p. The results shown in figure 20B indicate that *RPL30* splicing is not resistant to L30 in *mud2Δ* extracts, arguing against a Mud2-dependent CC2 hyperstabilization by L30. Since Mud2p has been involved in U2 snRNP recruitment via interactions with Prp11 (Abovich et al., 1994), an interference with this connection to repress U2 association can also be discarded. It could also be possible that deletion of *MUD2* was not enough to destabilize these cross-intron interactions. An important additional experiment would be to deplete extracts of BBP and check for L30 inhibition. *In vitro* depletion of BBP has no significant effect on pre-spliceosome formation and splicing (Rutz and Seraphin, 1999), but needs to be recycled, along with Mud2p, during the transition from CC2 to pre-spliceosome.

Second, we tested if L30 repression of splicing could be related to other Mud2p-independent rearrangements in CC2. *SUB2* is an essential gene that has been proposed to remove Mud2p from the polypyrimidine tract, as deletion of *MUD2* suppresses its requirement for viability (Kistler and Guthrie, 2001). However, recent results suggest that the direct target of Sub2p is the BBP-Mud2p heterodimer (Wang et al., 2008). Therefore, since Mud2p is not involved in repression, we asked whether Sub2p is required for splicing inhibition by L30, using extracts from *mud2Δ sub2Δ* cells. The data shown in figure 20C demonstrate that L30 does not need

Sub2p to block U2 association with the pre-mRNA. Importantly, Sub2p has been shown to be required for CC2 formation (Zhang and Green, 2001), therefore this result strongly suggests that L30 does not require formation of the CC2 to block the progression of spliceosome assembly.

Third, we explored whether L30 could interfere with the interactions of U2 snRNP with the substrate. Cus2p has been implicated in the binding of U2 to the intron. Extracts from *cus2Δ* cells allow U2 recruitment even in the absence of ATP, as long as there is Prp5 activity (Perriman and Ares, 2000). Interestingly, although U2 ATP-independent association could be replicated in the *RPL30* intron, L30 can still block it (figure 20D, results). This is consistent with the model whereby L30 precludes branch site recognition by U2 snRNP.

Fourth, we assessed the role of Prp5p in L30 regulation. A model for splicing fidelity proposes that Prp5p controls optimal interaction between U2 snRNA and the intron. If this interaction does not occur correctly, Prp5p will target the complex for degradation (Xu and Query, 2007). If Prp5p ATP hydrolysis rate is lowered, it will result in retention of suboptimal substrates for pre-spliceosome formation. Therefore, we asked if low levels of Prp5p ATPase activity enable the *inhibited complex* to progress in spliceosome assembly. Mutants in *prp5* (Xu and Query, 2007) known to change U2 accessibility to introns, did not have any effect on basal *RPL30* splicing (figure 21, results). However, some of these mutants could improve or prevent splicing of suboptimal substrates (mutants in the branch site sequence (Xu and Query, 2007)). We determined by copper assays that L30 regulation is not affected by these mutants. Therefore, we speculate that L30 is inhibiting prior to the function of Prp5p in U2 snRNP recruitment. This result also suggests that the initial base-pairing between the U2-intron that is locked by Prp5p to allow splicing progression, is not taking place in the *inhibited complex*.

Taken together, our results are consistent with a model in which L30 does not prevent intron recognition, and that the regulated spliceosome can adopt the CC2 conformation. Hence, our data strongly suggest that the inhibition by L30 takes place at the next step in spliceosome assembly, which is the recognition of the BS by U2 snRNP. Probably, inhibition is achieved by blocking some necessary rearrangement for U2 incorporation, rather than through an inhibitory contact with a splicing factor (as no two-hybrid interaction has been found between L30 and a spliceosomal component). This hypothesis is supported by the fact that L30 prevents *RPL30* splicing in HeLa extracts (data not shown). Although the splicing factors involved in the early

complexes in HeLa have diverged form their *Saccharomyces* counterparts, the interactions are thought to be largely maintained (Abovich and Rosbash, 1997) and probably susceptible to the same reorganizations.

The stabilization of U1 snRNA during repression can be due to a direct action of L30 on U1, or be a consequence of blocking the association of U2. The sequences close to the 5'ss of *RPL30* are particular in two aspects. First, they differ from the consensus (G/GUCAGU in *RPL30* vs. G/GUAUGU consensus, G/G denoting the 5'ss), although this divergence is not required for regulation (data not shown and (Eng and Warner, 1991)). Second, they have the potential of forming an extended base-pairing with U1. The evolutionary conserved *RPL30* 5'end intronic sequence is G/GUCAGUAU (Eng and Warner, 1991), which contains two extra positions (underlined) that can base-pair with U1snRNA. Interestingly, in a genome-wide analysis of yeast pre-mRNA splicing, *RPL30* appears as a very effective U1 recruiter (see figure 5 of (Tardiff et al., 2006)). However, it is possible to recapitulate L30 inhibition of splicing in constructs with the AU to UC mutation (G/GUCAGUuc, data not shown) in the *RPL30* intron. Therefore, extended base-pairing to U1 seems not to be required for regulation by L30, consistent with the fact that a hyperstable binding to U1 does not bypass the need for L30 in repression (figure 34, results). The unwinding of the U1—5'ss interaction seems to take place during the tri-snRNP addition (Staley and Guthrie, 1999), for that reason we did not expect to have any effect on U2 recruitment by increasing the stability of the U1 interaction with the pre-mRNA, but rather on U6. In our hands no effects were observed by increasing the potential base-pairings to U1 of non-regulated pre-mRNAs (figure 34, lane 8, results). However, we did not show that U1 is in fact hyperstabilized in these conditions (co-immunoprecipitation of U1 in increasing salt washing conditions).

Nevertheless, due to the high degree of conservation of the *RPL30* 5'ss, this particular sequence must have evolved to provide an optimal balance between expression and regulation *in vivo*. Because the *RPL30* gene is essential, very highly expressed, and in need of a constant fine tuning, proper balance between expression and regulation is crucial (Li et al., 1996).

IN VIVO STUDY OF L30 REGULATION

In vivo purification of the inhibited complex:

The first attempt to study *in vivo* regulation by L30 was the purification of the *inhibited complex* (IC) by the TAP tagging approach. This technique would have offered interesting data, such as the components on this complex or the relative stoichiometry between them. Unfortunately the IC was assembled *in vivo* in low quantities, due to the use of a strain whose endogenous *RPL30* gene cannot be regulated producing an excess of L30 in basal conditions (C9→T mutant, see figure 27, results). For this reason, this project is now being continued in the laboratory using an *in vitro* approach. The complex is assembled *in vitro* resulting in higher yields during the purification process.

Cotranscriptional study of the inhibited complex:

The development of chromatin immunoprecipitation techniques (ChIP) to follow spliceosome assembly offered the possibility to test our model of regulation *in vivo*. In addition, it provided a different tool to analyze possible interactions that may occur during L30 inhibition of splicing but escape other systems of detection (*i.e.* native gels). Furthermore, it also allowed approaching the question of co-transcriptional control of splicing. Consequently, we followed spliceosome assembly and its regulation by ChIP studies. Our ChIP profiles on the endogenous *RPL30* transcript for U1 and U2 during lack of regulation by L30 (figure 29, results) are qualitatively identical to those previously published (Tardiff et al., 2006). Repression of splicing by L30 does not block U1 association with the intron, as seen *in vitro*. In *RPL30* there is a low detection of U2, as expected (because of the length of the second intron), and induction of excess L30 diminishes U2 detection even further. This result indicates that L30 can repress U2 engagement co-transcriptionally, and argues against some particular association with U2 during inhibition of splicing (not detectable in our approaches *in vitro*; for example, an ATP-independent association of U2 similar to that of the mammalian E complex (Michaud and Reed, 1991)). The low levels of U2 co-transcriptional engagement seem to indicate that this step must be post-transcriptional, although some transcripts may have time to recruit it co-transcriptionally.

In order to define if L30 can also inhibit co-transcriptional U2 recruitment, a fusion between *RPL30* and the *LacZ* gene was performed. This chimaeric transcript allows to study late steps of spliceosome assembly (Lacadie and Rosbash, 2005). Interestingly, the same ChIP patterns were observed in the *RPL30-LacZ* construct. The interaction between the L30 protein and the nascent transcript was also verified by ChIP (figure 31, results), producing a ChIP profile as long as the message contains the L30 binding site (compare L30 ChIP in *RPL30-LacZ and RPL30*-LacZ*), and consistent with L30 being the effector of the changes detected in assembly of the spliceosome. Figure 31D shows a decreasing L30 ChIP signal along exon 2, very similar to that of U1 in regulated conditions, indicating that the binding of these two components is being stabilized on the nascent transcript. And, in agreement with previous results, no U2 recruitment could be detected.

Why is it advantageous for L30 to interact with its transcript during transcription? U2 engagement, whether co or post-transcriptional must be very fast and efficient. A discernible population of CC in cells under normal conditions has not been observed. For this reason, L30 needs to be well-positioned in advance prior to the moment of inhibition in order to block U2 snRNP recruitment.

There is also the possibility that L30 activates the export of the pre-mRNA at the expense of splicing (in a situation reminiscent of *YRA1* (Dong et al., 2007)), which is consistent with the finding that most of the *RPL30* repressed pre-mRNA is located in the cytoplasm. The unspliced *RPL30* transcripts in the cytoplasm are largely untranslated, suggesting continued association with L30. However, these cytoplasmic molecules that migrated out of the nucleus must have been previously dissociated from U1snRNP (Vilardell et al., 2000a).

Transcription, splicing and nuclear export of mRNAs are closely interrelated processes, which, partly share common molecular machineries. Sub2p and Yra1p interact with the RNA polymerase II dependent transcriptional machinery (Strasser et al., 2002), splicing factors and the export specific transport machinery at the nuclear envelope (Reed, 2003). Specifically, Sub2p has been proposed as one of the proteins that physically bridges the splicing and mRNA export machineries (Libri et al., 2001). Cotranscriptional binding of Sub2p would be enhancing export at the expense of splicing. Sub2p binds poorly *in vivo* to intron-containing genes (ICGs) and its pattern is inversely correlated to that of Mud2p (Moore et al., 2006), but also binds to

intronless genes. The proposed explanation is that Sub2p regulates the association of splicing factors with intronless genes as well as intron-conatining genes (ICGs).

Moreover, it was shown that pre-mRNAs that fail to be committed to the spliceosome assembly pathway are efficiently exported to the cytoplasm (mutations in the 5'ss or BS sequences, (Legrain and Rosbash, 1989; Rain and Legrain, 1997)). Mutations on these sequences would be affecting binding of U1snRNP and BBP/Mud2p, thus the complex is not recognized as suitable for pre-spliceosome formation, and binding of export complexes will be favoured.

Unprocessed pre-mRNAs are usually retained in the nucleus to be processed to mRNAs and are rarely exported to the cytoplasm. Pre-mRNAs that fail to complete splicing or 3'end formation are usually degraded by the nuclear exosome (Bousquet-Antonelli et al., 2000; Torchet et al., 2002). In contrast, *RPL30* pre-mRNA and others pre-mRNAs that escape the nuclear retention system of unspliced transcripts, are exported to the cytoplasm where they are recognized by ribosomes and degraded in a translation-dependent manner by the NMD pathway (He et al., 2003; Vilardell et al., 2000a). These pre-mRNAs also include those with mutations at the 5'ss or 3'ss (Legrain and Rosbash, 1989; Rain and Legrain, 1997).

We favour a model in which L30 can be either slowing down the incorporation of U2snRNP or forming an unsuitable commitment complex (both by an unknown mechanism) allowing the export machinery to join the nascent *RPL30* transcript. An interesting experiment would be to ChIP Sub2p and Yra1p on the nascent *RPL30* in regulated conditions. On normal conditions, the TREX complex should be recruited normally but the transfer of Sub2p and Yra1p to the RNA may be hindered by the spliceosome (Abruzzi et al., 2004). If spliceosome assembly is inhibited by L30, there could possibly be an increase in ChIP of these two proteins. This would explain the presence of *RPL30* unspliced pre-mRNA in the cytoplasm (Vilardell et al., 2000a).

SPLICING REGULATION IN YEAST

While metazoans have developed a more complex system of regulation of splicing (SR proteins and hnRNP), yeast also has instances of splicing regulation despite not having apparent functional homologues for these proteins. Splicing regulation has been reported for genes related to protein synthesis (*RPL30*), meiosis (*MER1*) and export (*YRA1*).

In these examples, binding of proteins initially not related to splicing regulation, can affect positively (*MER1*) or negatively (*RPL30* and *YRA1*) the splicing of their own or other pre-mRNAs. The mechanisms underlying these observations are still poorly understood.

With our work, we have shown that when L30 ribosomal protein is in excess in the cell, it is able to bind its own pre-mRNA co-transcriptionally to control its synthesis from the beginning. We have also observed that L30 does not interfere with the initial recognition of the intron (by U1, BBP and Mud2p) but can block U2 snRNP recruitment, both post-transcriptionally (on the endogenous *RPL30*) and co-transcriptionally (*RPL30-LacZ*). However, the exact mechanism of inhibition is still unknown. We hypothesized that although binding of commitment complex factors occur, the conformation is not suitable for progression in spliceosome assembly. Thus, some proofreading mechanism has to be activated (analogously to Prp5, competition between progression in splicing and disposal of the pre-mRNA). One possibility is that the *inhibited complex* is not in the proper configuration for U2 recruitment. In support of this model, the configuration of the *inhibited complex* could enable the engagement of the export machinery because spliceosome assembly cannot progress. A competition between these two machineries has been found (Abruzzi et al., 2004).

L30 is a small protein that has a highly restricted sequence due to the high number of interactions that it establishes inside the ribosome. Therefore, we do not expect L30 to interact with any spliceosomal protein. Nevertheless, it has evolved to strongly stabilize an RNA secondary structure that interferes with the proper configuration of the spliceosome on the intron, thus changing the destiny of its own pre-mRNA towards a discarding pathway. The inhibited *RPL30* is exported to the cytoplasm and degraded by NMD, because of the presence of in frame stop codons (Vilardell et al., 2000a). Therefore, every molecule of L30 that cannot

join the ribosome will have the possibility of controlling expression of its own pre-mRNA, from the very beginning of its synthesis (transcription) and accompanying it until the end.

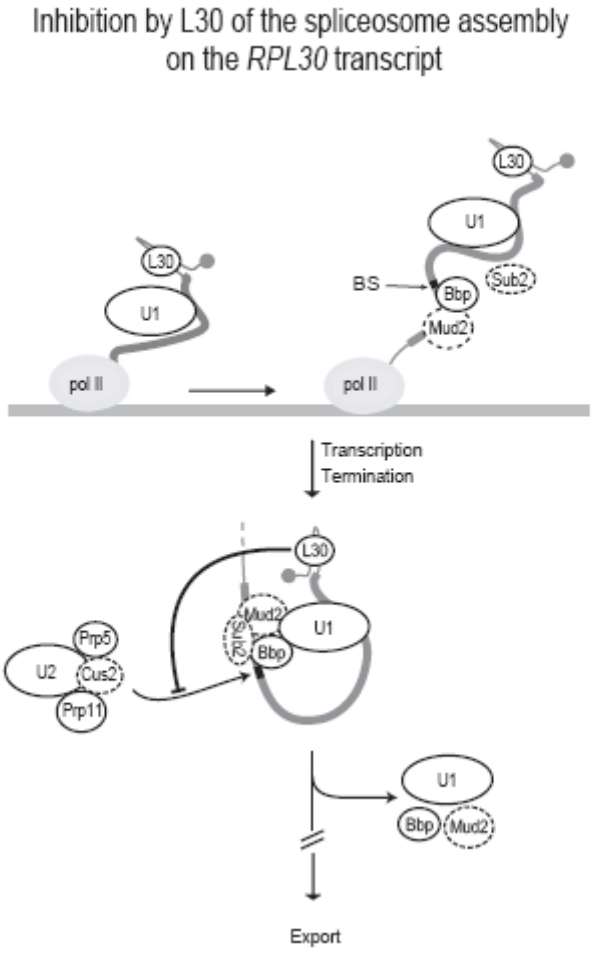

Figure 37. Model for inhibition of splicing by L30. The cap structure is represented as a solid circle, the kink-turn bound by L30 as a spike, and the intron as a thick line; all in grey. The branch site (BS) is indicated in black. The downstream exon is represented by a dotted line. DNA is in light grey. Factors shown not to be required for L30 repression are represented with dotted circles.

Under conditions of excess L30, the protein interacts early with the nascent transcript and binds a kink-turn near the start of the intron, but this does not prevent its recognition by CC2 factors. In *RPL30* most of U2 snRNP binds to the CC post-transcriptionally, but when L30 is present this binding is abolished. L30 inhibition of splicing is independent of the roles of Mud2p, Cus2p and Sub2p.

After splicing inhibition, the complex needs to be remodelled before it can be exported to the cytoplasm where *RPL30* pre-mRNA will be degraded by NMD.

Mud2p IN SPLICING FIDELITY

Mud2p protein is the yeast putative ortholog of the mammalian $U2AF^{65}$. $U2AF^{65}$ binds to the highly conserved poly-pyrimidine tract of introns (Zamore et al., 1992). However, Mud2p seems not to be essential for growth in yeast. There are several explanations for this fact. In yeast, polypyrimidine tracts (Ppy tracts) are often missing in introns, whereas the branch point is highly conserved. Therefore, in some introns Mud2p does not have a binding site. Second, Mud2p is lacking the RS (arginine-serine) domain present in $U2AF^{65}$ proposed to help base-pairing of U2 snRNP with the branch site (Valcarcel et al., 1996). These evidences suggest that Mud2p should be less necessary for progression of the spliceosome assembly.

However, Mud2p needs to be recycled to allow progression through the spliceosome assembly pathway. Sub2p has been proposed to be the helicase that helps in this recycling (Kistler and Guthrie, 2001). The action of Sub2p occurs during the first ATP-dependent step of spliceosome assembly, the transition from commitment complex 2 (CC2) to pre-spliceosome (PS).

Every ATP-dependent step in spliceosome assembly is an opportunity for regulation and maintenance of splicing fidelity (as it is the case of *RPL30* regulation of splicing). There are multiple examples that strengthen the model of the "kinetic proofreading". The paradigm is the case of Prp16p, a DEAH-box ATPase that facilitates the transition between the first and the second step of splicing. However, other examples have been reported for the different transitions that occur during the spliceosome assembly pathway. Recently, Prp5p has been implicated in the kinetic proofreading of the first transition step of the spliceosome assembly, from CC2 to PS (Xu and Query, 2007). Mutations in Prp5p that lowers its ATPase activity enable splicing of suboptimal substrates. Slowing down the transition between CC2 and PS, the chance for progression of a suboptimal substrate into the splicing pathway is increased.

During the development of this thesis we found a similar result for Mud2p. Deletion of the *MUD2* causes a partial ATP-independent recruitment of U2snRNP *in vitro* (figure 22A, results). Moreover, we determined that the ATP-independent U2 recruitment was achieved through a base-pairing interaction with the branch site (figure 22B, results), suggesting that BBP was already displaced from the BP allowing U2 base-pairing. Mud2p has been shown to interact with BBP, and BBP in turn interacts with U1snRNP (through Prp40p) (Abovich and Rosbash, 1997).

We could hypothesize that deletion of *mud2Δ* helps destabilizing BBP from the intron. This could be the explanation of why the deletion of *SUB2* is no longer lethal when combined with *MUD2* deletion (Kistler and Guthrie, 2001), although both BBP and Mud2 are the target for Sub2-recycling function (Wang et al., 2008). The interaction between BBP and Mud2p should be necessary to stabilize BBP at the BS. This interaction, amongst others, is important for the intron definition step, where pre-mRNAs will be committed to the splicing pathway. According to this model, if a commitment complex factor is missing, the pre-mRNA will fail to be spliced. In fact, this is the case for Mud2p, *in vivo* deletion of *MUD2* enhances pre-mRNA export at the expense of splicing (Rain and Legrain, 1997).

Thus the most plausible explanation for our observations is that deletion of *MUD2* promotes *in vitro* destabilization of BBP from the branch site, allowing U2 snRNP base-pairing, independent of the presence of ATP. However, we still do not know if the pre-spliceosome-like complex that it is formed in the absence of ATP is functional and can be chased into mature spliceosomes (see figure 39, discussion).

A consequence of this hypothesis is that deletion of *MUD2* will improve splicing of suboptimal substrates by favouring the transition from CC2 to PS. Next, we asked if mutations in the BS or the 3'ss sequences were bypassed by *MUD2* deletion.

The branch site sequence (BS) is highly conserved in yeast (UACUAAC). During the first step of spliceosome assembly this sequence is bound by BBP. Mutations in the BS sequence can diminish or completely abolish BBP binding *in vitro* (Berglund et al., 1997). Any mutation in the "branch" adenosine UACUAAC (underlined) completely abolishes BBP *in vitro* binding so does the mutation UAC**A**AAC (mutation in bold). In contrast, the mutation UACU**G**AC does not affect BBP binding (Berglund et al., 1997).

Two of these previously characterized mutations were used for our *in vivo* study of *MUD2* deletion. In our case, the mutation UAC**A**AAC is referred as U257A (see results, figure 23), and the mutation UACU**G**AC, as A258G. Surprisingly, the two mutations behaved differently upon deletion of the *MUD2* gene, and while U257A has been shown to abolish BBP binding, A258G has not. This observation made us propose a model for the cooperation of the BBP and Mud2p interaction in the progression of spliceosome assembly. When the BS sequence contains a mutation that slightly affects BBP interaction, splicing efficiencies are affected by *MUD2Δ*, and

this is because Mud2p protein helps stabilizing BBP at the BS (see figure 38B). In contrast, when the effect of the deletion is measured in BS mutants that completely abolish BBP binding, the deletion of *MUD2* does not do anything (figure 38C) . This raises the idea of the need of an optimal binding of BBP at the BS for U2 snRNP recruitment. This optimal binding could be achieved through the cooperation of Mud2p. Mud2p is then only necessary when the BS sequence is mutated. In general, BS sequences are highly conserved and this can be the explanation by which Mud2p is not necessary for splicing *in vivo* (figure 38A).

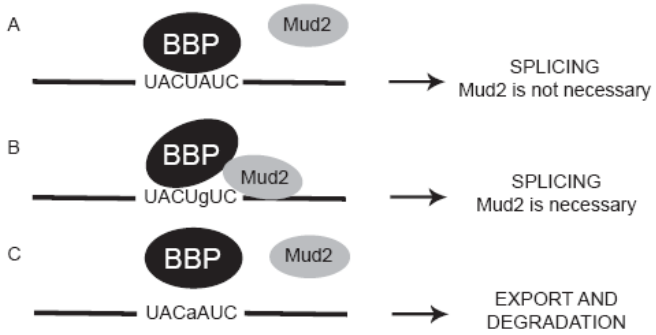

Figure 38. Model of the cooperative binding of BBP and Mud2 for spliceosome progression.

(**A**) In normal conditions (consensus BS sequence) Mud2p is not necessary as BBP binding is optimal for spliceosome progression.

(**B**) Mutations in the BS sequence that slightly affect BBP binding need Mud2p to stabilize BBP allowing the optimal binding for spliceosome progression.

(**C**) Some nucleotides are critical to preserve BBP binding. When these nucleotides are mutated BBP cannot bind the BS at all. Thus the pre-mRNA will not be committed for splicing. Unspliced pre-mRNA will be exported to the cytoplasm where it is probably degraded by NMD.

During the first steps of spliceosome assembly the conserved intronic sequences are recognized to commit the pre-mRNA into the spliceosome assembly pathway (5'ss and BS), except the 3'ss. In fact, in yeast the first catalytic step can be accomplished *in vitro* with RNAs that lack the 3'ss sequence (Rymond and Rosbash, 1985). In contrast, metazoan systems need the 3'ss sequence for the first catalytic step (Ruskin and Green, 1985) which is bound by U2AF^{35}. No homolog has been found for it in yeast. *S. cerevisiae* seems not to have developed a system to sense the status of the 3'ss until the transition from the first to the second catalytic step (Umen and Guthrie, 1995a; Villa and Guthrie, 2005). From an evolutionary point of view, it seems a waste of energy to commit a pre-mRNA for splicing when it will not be able to progress during the catalytical steps. The fact that Mud2p can crosslink to intronic regions between the BS and 3'ss (McPheeters and Muhlenkamp, 2003) and seems to be the homolog of U2AF^{65} prompted us to ask if Mud2 had any affect on the selection of the 3'ss. This is why we asked if splicing fidelity of 3'ss mutants pre-mRNAs was altered in a *mud2Δ* strain. We used mutants in the position -3 of the 3'ss (+1 is the first nucleotide of the following exon). The -3 position is preceding the dinucleotide AG and is not highly conserved. Normally the -3 position consists of an uridine (**U**AG), although in a high number of introns the combinations **A**AG or **C**AG are also found. In contrast the combination **G**AG is not found in any natural yeast intron, as it is very deleterious for splicing (see figure 24B, results). However, splicing to this sequence is the one that is significantly improved by deletion of the *MUD2* gene. This result does not imply a direct interaction between Mud2p and the 3'ss, although crosslinking experiments demonstrated that Mud2p can crosslink to the position -3 of the intron (McPheeters and Muhlenkamp, 2003). This result indicates that Mud2p has a role in the maintenance of the 3'ss fidelity. An alternative explanation arises from the observation that, *in vitro,* deletion of *MUD2* increases the incorporation of U2 snRNP in an ATP-independent way. Thus, *in vivo* when using optimal pre-mRNAs (conserved intronic sequences) there is no effect upon deletion of *MUD2,* but in pre-mRNAs whose splicing is impaired, the increase in U2 recruitment can bypass the effect of the mutations in the intronic sequences. Therefore, Mud2p controls the progression from CC2 to PS. In normal conditions, the presence of Mud2p is a negative modulator of this transition, but the progression of CC2 to PS is facilitated when Mud2p is recycled or, in this case deleted (as shown in figure 39, discussion).

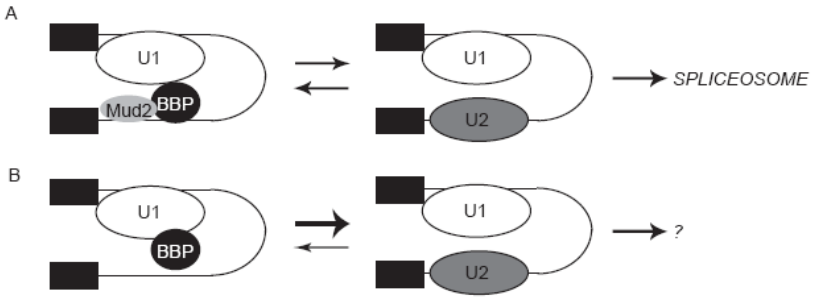

Figure 39. Deletion of *MUD2* increases the incorporation of U2snRNP independent of ATP.
(**A**) Once the commitment complex (CC) stage has been attained, the pre-mRNA is committed to the spliceosome pathway. The CC is remodelled (BBP and Mud2) to allow U2snRNP base-pairing to the BS (pre-spliceosome, PS), in an ATP-dependent manner. Both states (CC and PS) are found in an equilibrium. (**B**) Deletion of Mud2 favours the incorporation of U2snRNP. Mud2 helps stabilizing BBP in the BS. If Mud2 is absent, BBP is more easily destabilized allowing higher probability of U2snRNP recruitment. An increase of U2 incorporation is observed in the absence of *MUD2 in vitro*, an improvement of splicing is observed with suboptimal substrates *in vivo*. Pre-spliceosomes formed in the absence of *MUD2* and ATP should therefore be functional for the progression of spliceosome assembly.

CONCLUSIONS

1. L30 binding to its own pre-mRNA does not prevent U1 snRNA base-pairing interaction, although the 5'splice site is in a stem.

2. U1 snRNA is more stable in the complex with L30 compared to a normal commitment complex. U1 hyperstabilization is not affected by the presence of ATP *in vitro.*

3. L30 does not interfere with the recognition of the 3'end of the *RPL30* intron: BBP and Mud2 binding.

4. The *inhibited complex* cannot bind U2 snRNP *in vitro*

5. Neither *mud2Δ* nor *sub2Δ/mud2Δ* affects the formation of the *inhibited complex in vitro.*

6. The functions of neither Prp5p nor Cus2p in U2 snRNP recruitment are required for the formation of the *inhibited complex.*

7. *In vivo* formation of the *inhibited complex* is low to allow purification by the TAP tagging procedure.

8. L30 can regulate spliceosome assembly on its pre-mRNA co-transcriptionally.

9. Co-transcriptional binding of L30 to its own pre-mRNA leads to a retention of U1 snRNP on the *RPL30* transcript.

10. Co-transcriptional L30 binding can affect U2 snRNP post-transcriptional recruitment (endogenous *RPL30*) as well as co-transcriptional (*RPL30-LacZ).*

11. Hyperstabilization of U1 snRNP on a non-regulated L30 pre-mRNA does not mimic L30 inhibition of U2 snRNP recruitment.

12. Binding of a heterologous protein (MS2) upstream to the 5'splice site blocks splicing.

13. Binding of MS2 upstream of an intron blocks U1 snRNP recruitment.

14. Deletion of *MUD2* partially bypasses the *in vitro* ATP requirement for U2 snRNP recruitment.

15. Mud2p has a role on the splicing fidelity of the 3'splice site sequence.

MATERIALS AND METHODS

Strains

All strains used for this work are listed in the following table

Strain	Genotype	Reference
BJ2168	MATα ura3-52 leu2 gal2 prb1-1122,1 pep4-3prc1-407	
yMR1	Mat a , ura 3-52 leu2 trp1-289 ade2 HIS3 (his+) msl5::LEU2 + pGAL-BBP-HA	(Abovich and Rosbash, 1997).
yJV47	MATα, ura3-52, leu2, gal2, prb1-1122,1 pep4-3, prc1-407, Mud2-TAP-URA3	This study
FDBS007-02B(A)	Mat a, ura3-52; his3Δ200; LEU2; LYS2; trp1Δ63; YNL286w(4, 852)::KANMX4	Euroscarf ref. 10082A
yJV51	Mata, trp 1-289 ura 3-52, arg4(RV-), ade2, MUD2KO::LEU2, Leu2::KANMX6	This study
yCG472	yCG470 MUD2KO::KANMX6	(Kistler and Guthrie, 2001)
DT-128	MATa, ade2, his3, leu2, trp1, ura3, lys2,SNU71-HTBKAN	This study
DT-130	MATa, ade2, his3, leu2, trp1, ura3, lys2, LEA1-HTBKAN	This study
yYZX02 (PRP5 WT)	Mata, ade2, cup1d::ura3 his3 leu2 lys2 prp5d::loxP trp1, pRS314-PRP5(PRP5 TRP CEN ARS)	(Xu and Query,2007)
yYZX02 (PRP5 N399D)	Mata, ade2, cup1d::ura3 his3 leu2 lys2 prp5d::loxP trp1, pRS314-PRP5(PRP5-N399D TRP CEN ARS)	(Xu and Query,2007)
yYZX02 (PRP5 SAW)	Mata, ade2, cup1d::ura3 his3 leu2 lys2 prp5d::loxP trp1, pRS314-PRP5(PRP5-SAW TRP CEN ARS)	(Xu and Query,2007)
yYZX02 (PRP5 TAG)	Mata, ade2, cup1d::ura3 his3 leu2 lys2 prp5d::loxP trp1, pRS314-PRP5(PRP5-TAG TRP CEN ARS)	(Xu and Query,2007)
yJV30	MATa, ade2, ura 3-52, leu2, lys2, RPL30::T9, Snu71::CBPURA3.	This study

DTY-128 (Snu71-HTB; oligos DT410/DT411) and DTY-130 (Lea1-HTB; DT413/DT414) were generated by PCR amplification of pFA6a-HTB-kanMX6 (gift from Peter Kaiser, (Tagwerker et al., 2006)) and transformation into W303. Integration was confirmed by diagnostic PCR and expression/biotinylation confirmed by western blot analysis with streptavidin-HRP (Amersham). Primers are as follows: DT410 Forward primer 5'-AAACGATAGCTGATAGACTGTGGAGTCGTAAAGAATTTCGCTTGGGGACCCGGATCCCCGGGTTAATTAA-3', DT411 Reverse primer 5'-TTTCAGAGCGAGCCTTTCCCTTTTGGGACGCGCGCCAAGGCCCTTCTGTTGAATTCGAGCTCGTTTAAAC-3' DT413 Forward primer -5'-CTTCTTTAGAAGAGATTGCCAGGCTGGAAAAACTACTCTCTGGTGGTGTTCGGATCCCCGGGTTAATTAA-3' DT414 Reverse primer 5'-

TTTATAATTCTTTTTTTTTAAGTCATTGAACAGTCGCACTAACCAAAAGAGAATTC
GAGCTCGTTTAAAC -3'

yJV47 (Mud2-TAP; oligos JV309/JV310) was generated by PCR amplification of pBS1539 (Puig et al., 2001) and transformation into BJ2168. Integration was confirmed by PCR and western blot analysis with PAP (peroxidase anti-peroxidase, purchased form Sigma). Primers are: JV309 5'-
GTTCTGTGTACTTATATAGATGAGGACGACTTTGACATGATGGAAGCAACCCAACTTTCCTC
CATGGAAAAGAGAAG and JV310 5'-
TTAGATCTACATAATGAATACTCAATTCTTTACTTAATTTCGCTCTACAAAATAGACCATTAC
GACTCACTATAGGG.

Plasmids

pL30-5A plasmid, in which 5 adenines substitute for the nucleotides 17-50 in *RPL30*, was made by a circular PCR with primers JW504 (5'-aaACGCAGAGATGGTCAGTA-3') and JW505 (5'-tttACACTCCAGTCTGTTTGG-3') on a pGEM-3Z plasmid with the sequences corresponding to the entire transcript of *RPL30* (pL30pre). The insert was fully sequenced. pL30T6AU65 (*RPL30* with a 5'SS consensus, introducing also the appropriate complementary mutation to keep the L30 binding site) was made by circular PCR on pL30pre with primers JW518 (5'AATTGAATAAGCTGTAGGTTCTTAAACACTCCGGTaTGT-3') and JW519 (5'-AATCAATATACGCAGAGATGGTatGTATAA-3'). The product was sequenced.

pSM33 (*RPL30-LacZ*) was constructed as follows: primers JV689 (5'-ttaactagtGCCGTCGTTTTACAACGTCGTGAC-3') and JV687(5'-ttagtcgacTTATTTTTGACACCAGACCAACTG-3') were used to amplify the *LacZ* gene from YEp357, digested with SpeI/SalI and inserted in a pRS316 containing the *GPD* promoter (from the start of transcription to 605 bp upstream), *RPL30* exon1, intron, and 12 nucleotides of exon 2. It also includes 496 bp of the *RPL30* terminator.

pJMB1 (*RPL30*-LacZ*) is based on pSM33 with point mutations in exon1 from *RPL30* thatabolish the kink-turn formation, and therefore L30 binding and regulation. Mutations were introduced by PCR using the oligo JV700 S (5'-

gacggatccaaacactggtcgcattaagAGAACCTACAGCTTATTCAAT-3') and JV90 (5'-gtcgacctgcagactagtTTTAACTGGGGCCTGTT-3'), digested with BamH1/Spe1 and cloned into pSM33. The construct was verified by sequencing. These mutations do not affect the annealing to the primers used in the ChIP analyses.

pMB73 was made by fusing *ACT1* leader sequences (pos. -64 to -26 respective to the start of translation) to the *RPL30* exon 2 and 23 nt of its terminator. Primers were JV200 (5'ctaggatccccggcgactcttttagatttttcacgcttcactgctttttttcttcccaaatgGCCCCAGTTAAATCCCA-3') and JV120 (5'-actgtcgACAAATCGTTTGAACTTACCTTA-3'), used on a plasmid containing *RPL30* sequences. This, flanked by BamHI-SalI sites, was cloned into pG14(Lesser and Guthrie, 1993). This chimaeric gene is not subjected to any L30 feed-back control of expression.

pSM34 was made by cloning into pRS314 a fragment containing the *GAL* promoter and Protein A, from pBS1761 (Euroscarf), using primers JV166 (5'-atcgcggccgcGTACGGATTAGAAGC-3') and JV267 (5'-aggactagtAAGCTTATCGTC -3'). This plasmid also contained the RPL30 terminator, obtained by PCR with primers JV89 (5'-atggtcgacTAAGGTAAGTTCAAACGA-3') and JV88 (5'gtagggcCCTTCCATACCTTCCC-3'). The L30 ORF was made by PCR with primers JV894 (5'-cttactagtgtcagtggaGCCCCAGTTAAATCCCAAG-3') and JW456 (5'-ctcgtcgACCTTATTTAAGCCAAGGT-3') and inserted in between the Protein A and the terminator, using SpeI and SalI sites.

pSM35 was based on pMB73. The same insert: GPD-L30 exon 2-PGK terminator was inserted in pRS316 using XbaI and HindIII sites.

pMBP-MS2 was constructed by inserting an MS2 ORF PCR fragment in the StuI HindIII sites of pMalc (New England Biolabs). The cloned MS2 contains de following mutations to avoid multimerization: V30I, positions 68-80 deleted (VATQTVGGVELPV), and the position A81G. All mutagenesis were verified by sequencing.

RPL30 WT-GFP (pSM17) and RPL30-T9 GFP (pSM18) were constructed as follows: they contain the GPD promoter, followed by exon 1 intron and first 12 nucleotides of exon 2. A PCR amplifying the GFP ORF (JV270'-ttaactagtAGTAAAGGAGAAGAAC-3', and JV271'-ttagtcgacTTATTTGTATAGTTCAT-3') was inserted SpeI/Sal I between the *RPL30* second exon and the *RPL30* terminator. The pSM14 to overexpress *S. acidocaldarius* L30 was cloned in a

pRS314 containing the GAL promoter fused to Protein A tag (NotI/SpeI from pBS1761 (Puig et al., 2001)) followed by the ORF of SaL30 and the *RPL30* terminator.

In addition we created the pSM15 is a subclone of pBS1539 (Puig et al., 2001) that was digested NheI/EcoRI and klenow treated, in order to eliminate the protein A tag. A PCR amplifying CBP tag plus URA3 marker with tails complementary to the Snu71 gene, was used to generate the yJV30 strain by homologous recombination.

Plasmids bearing mutations in the BS and 3'ss were from (Query and Konarska, 2004).

RNAs and T7-templates

Primer T7 and JW839 (5'-GACTTGATAACCAAAGCC-3', position 55 in *RPL30* exon 2) were used to amplify the *RPL30* gene from the pL30-5A plasmid to generate the wt template. A PCR with primers T7 and JW765 (5'-GCCTTCTTGCTAATCCC-3') on the pL30 T6AU65 was used to generate a shorter version of the *RPL30* RNA for psoralen cross-linking experiments. These primers were also used to generate the *RPL30* WT "1-47" and the *RPL30* "+12 1-47" RNA. For a template to transcribe *RPL30* "+12" RNA primers JW1185 (5'-cctaatacgactcactataggAAACACATCGGAGTGTAAAAAACGCAGAGATGatgatgatgGTCAGTATAAC-3') and JW839 were used.

Branch site sequence (TACTAAC) was deleted from the *RPL30* WT and "+12" templates using a PCR-based deletion method. Point mutation of the BS (TACTAuC) was done on the *RPL30* WT template using PCR methods. *RPL30 +12ext to U1* was made from a PCR JV724 (5'-cctaatacgactcactataggAAACACATCGGAGTGTAAAAAACGCAGAGATGatgatgatgaagGTAAGTATAACTGA-3') and JW839.

RNAs containing MS2 binding sites were designed as follows: *MS2actin* from a PCR with JW1353 (5'-actaatacgactcactataggGATCGAAAATTTACTGAATTAAtacacgatcacggtatgtataAGCGCTTGCACCATC-3') and JW1181 (5'-CACATACCAGAACCGTTATC-3') on actin gene. *MS2 +12actin* from a PCR with JW1366 (5'-actaatacgactcactataggGATCGAAAATTTACTGAATTAgtagaggatcaccctactactactagtatgtataAGCGCTTGCACCATC-3') and JW1181 on actin gene. And *MS2Rpl30* with a PCR JW1361 (5'-

actaatacgactcactataggGATCGAAAATTTACTGAATTAtgacacgatcacGGTCAGTATAACATGATT TTATAAC-3') and JW839 on rpl30 gene.

RNAs were transcribed as described (Vilardell and Warner, 1994).

Radiolabelled pre-mRNAs are synthesized with: 1mM CAP, 30mM DTT. 0.5mM ATP, 0.5mM CTP, 0.2mM GTP, 0.15mM UTP, 0.5µl RNase inhibitor, ~100 ng PCR template and 0.01mM α-^{32}P-UTP.

Recombinant proteins, yeast extracts, and *in vitro* splicing.

MBP:L30 fusion protein was purified as described previously (Vilardell and Warner, 1994). MBP:MS2 protein was a gift of Nuria Majós.

Yeast extracts were made as described in (Umen and Guthrie, 1995b). MBP:L30 or MBP:MS2 proteins were added to the extracts at a ratio of 0.1 - 1 µg per 4 µl of extracts, depending on the experiments. *In vitro* splicing reactions were performed as described in (Vilardell and Warner, 1994).

Yeast competent cells and transformations

Yeast competent cells were made following the lithium acetate protocol (Gietz et al., 1992) and transformed by the PEG/LiAc-Te method.

Copper assays

Cultures were grown to midlog phase in -Leu, -Trp -Ura medium, diluted to 0.3 OD, and equal volumes were dropped onto -Leu, -Trp, -Ura plates containing 0, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6 until 1.5mM of $CuSO_4$ (Lesser and Guthrie, 1993) with serial dilution 1/3, 1/9 and 1/27. Plates were scored and photographed after 3 days at 30ºC.

RNA extraction and RT-PCR

RNA was obtained from 15ml yeast cultures (0,3-0,8 OD_{600}). Pellets were processed by the Hot/Acid-phenol method. RNAs were quantified using Nanodrop, and quality was checked by agarose- ethidium bromide gels.

5 µg of total RNA was DNAsed and retrotranscribed using the JV697 oligo for the *RPL30LACZ* analysis and JW839 for endogenous *RPL30* RNA. Retrotranscribed material was amplified with 25 cycles of PCR with oligos JV169 (5'-GGATCCAAACAGACC-3')/JV697 for *RPL30LACZ* and JV169/JW839 for endogenous *RPL30*.

Psoralen crosslinking

Psoralen crosslinking was performed as previously described (Du and Rosbash, 2001). Commitment complexes were formed under splicing conditions (10 µl) and placed on ice. AMT-psoralen (SIGMA) was added to a final concentration of 20 µg/ml and samples were irradiated at 365nm for 10 min. U1snRNA digestion was performed after psoralen crosslinking using an anti-U1 oligonucleotide JW1411 (5'-GAATGGAAACGTCAGCAAACAC-3') in the presence of RNAse H enzyme during 30min. at 37ºC. Co-immunoprecipitations of the cross-linked material with MBP-L30 were performed as described in (Vilardell and Warner, 1994). Polyacrylamide gels were visualized using Phosphorimager.

Immunoprecipitations

Splicing complexes containing MBP:L30 or MBP:MS2 were isolated by immunoprecipitation with antibodies against MBP. When indicated, extracts were depleted of ATP by addition of 0.2 mM glucose and subsequent incubation at 25ºC for 10 min. Splicing reactions (10 µl) were set up with 1 pmol of cold RNA and incubated 20 min at 23ºC. After spliceosome assembly, 0.5 µl of monoclonal antibody anti-MBP (New England Biolabs) was added and the reaction was placed on ice 10 min. 200 µl of SPL buffer (60mM KPi, 2.5mM MgCl2, 1mM DTT, 150mM KCl, and 0.05% NP40) containing Protein A-Trisacryl beads (Pierce), previously washed 3 times with the same buffer, were added. After 2 hr rotating at 4C, beads were washed 3 times with 500 µl of SPL-200mM KCl. During the last wash beads were transferred to a new tube.

For protein analysis of the immunoprecipitated product, 20 µl of 2x Laemmli buffer were added to the beads and analyzed in 10% SDS-PAGE gels. Gels were transferred to a nitrocellulose membrane and blotted with different antibodies. Antibody anti-HA was purchased from Roche, to detect TAP proteins PAP antibody from Sigma was used. For RNA analysis, the washed beads were resuspended in 150 µl of 0.3M sodium acetate pH 5.2, phenol-chloroform extracted

and ethanol precipitated. Samples were run on a denaturing formaldehyde-containing gel and transferred to a nylon membrane.

Riboprobes for detection of U1, U2, U4, U5, U6 snRNAs were T7-transcribed from plasmids containing the cloned snRNAs and using the following oligos: for U1, JW1187 (5' cctaatacgactcactataGGGAACGAGCAAAGTTG-3') and JW1188 (5'-GAGGAGATCAAGAAGTCCTAC-3'), for U2, JW1189 (5'-cctaatacgactcactatagGGCGTCAACCATCAAGTC-3') and JW1190 (5'-GGTGGCGCTGCAAGAGG-3'), for U4, JW1191 (5'-cctaatacgactcactataggGACACTCGAGTCTCATTC-3') and JW1192 (5'-GTCCTAAAGTACTAATCCACC-3'), for U5, JW1193 (5'-cctaatacgactcactataggGCCCTCCTTACTCATTGAG-3') and JW1194 (5'-CAAGCAGCTTTACAGATCAATG), for U6, JW1195 (5'-cctaatacgactcactataggGAAATAAATCTCTTTGTAAAACG-3') and JW1196 (5'-GTTCGCGAAGTAACCCTTCGTG-3')

U4snRNA hybridizations are not shown because addition of MBP:L30 to the reaction coimmunoprecipitates U4, regardless of either pre-mRNA presence or splicing conditions. U4 snRNA contains a kink-turn structure with similarities to the L30 binding site, and L30 shows general affinity to kink-turns.

Riboprobe for detection of *ACT1* substrate was made from a PCR with oligos JV1137 (5'-taatacgactcactatagggCACATACCAGAACCGTTATC-3') and JV1138 (5'-GTATGTTCTAGCGCTTGCAC-3'). Riboprobe for the detection of the *RPL30* pre-mRNA was transcribed from a PCR with JW453 (5'- GCCCCAGTTAAATCCCAAGAATC-3') and JV197 (5'- taatacgactcactataGGTGGTCAAGATATCAGA-3').

Riboprobes were transcribed by T7 polymerase in reactions containing: 30mM DTT, 0.5µl of RNAse Inhibitor, ~100ng PCR template, 12.5µM of ATP, CTP and GTP and supplemented with 3µl of α-^{32}P-UTP (Amersham, PB40383).

2'O-methyl oligonucleotide (against U2) was added into standard splicing reactions (10µM) prior to pre-mRNA addition. Reactions were incubated at 30ºC for 15 min

Mobility shift-assay

RNA-protein binding reactions were performed in 20µl of 30mM Tris-HCl (pH 7.5), 75mM KCl, 2mM $MgCl_2$, 1mM DTT, 50ng/µl of BSA, 40ng/µl of tRNA, 5-50 fmoles of probe and from 0.05 to 1µg of MBP:L30. Reactions were incubated at 25ºC during 20 min and loaded immediately in a running acrylamide gel (6% acrylamide and 0.5xTBE).

RNA-protein crosslinking

A standard T7 transcription reaction was supplemented with 0.2 mM 4-thio-UTP and 0.2 mM rUTP, 50 µM rCTP and 12.5 µM α-32P-CTP (800Ci/mmol, 40mCi/ml). A 30 µl splicing reaction containing 1 fmol of thio-U labelled RNA was irradiated at 365 nm during 10 min at 4ºC. RNase A was added and reactions were incubated 30 min at 37ºC. Immunoprecipitation of Mud2-TAP was performed using IgG-sepharose beads (Amersham) in IPP-150 buffer
(10mM Tris pH 8.0, 150mM NaCl and 0.1% NP40) during 2 hr at 4ºC. Beads were washed twice with IPP-150, and eluted material was loaded on a 8% SDS-PAGE and analyzed by Phosphorimager (Molecular Dynamics).

TAP purification

TAP purifications were performed from extracts corresponding to 2L of yeast culture (yJV30 strain). Cells were grown in glucose or galactose (to overexpress SaL30). Cultures were grown until OD600 was 2-3. Cells were centrifuged and washed to be frozen in liquid nitrogen. Pellets were broken by using a mortar with liquid nitrogen, grinding for 20 min. Extracts were centrifuged (same conditions as described in (Umen and Guthrie, 1995b) and without dialysis were incubated O/N with IgG beads. The procedure and buffers used for TAP purification was exactly the same as described in (Puig et al., 2001).

Chromatin immunoprecipitation

ChIPs were performed as described previously in (Lacadie and Rosbash, 2005). Cells containing the pSM34 were grown in 2% galactose-containing media to express the L30-TAP fusion protein. ChIPs for U1snRNP recruitment was performed with the DTY-128 (Snu71-HTB) biotin-tagged strain. To follow U2snRNP assembly the DTY-130 (Lea1-HTB) strain was used.

Purification of biotin-tagged strains was performed with streptavidin beads (Amersham Biosciences). Quantitative PCR was analyzed on a LightCycler 480 (Roche Diagnostics). All samples in a single PCR run were assayed in duplicate. All data represent the average of at least two independent experiments with the errors bars displaying the average deviation.

Inputs and IP signals were normalized to a primer pair amplifying the center of the *PMA1* gene, except for the ChIP analyses on *ACT1* and with L30-TAP, which were normalized using the first primer pair. Primers used for the analyses are listed in the following table.

Name	**Sequence 5'→3'**	**Position (ref. to start codon)**
SM1	CCCGTCTATTCTCGTGTCGT	-502 forward in RPL30
SM2	ATGATCCTTACTGCGGTGCT	-452 reverse in RPL30
JV702	TTCCATTTGTTGGAATGTTCA	-120 forward in RPL30
JV693	CATCTCTGCGTATATTGATTAATTG	+2 reverse in RPL30
JV696	ATCGTTTACATTTCAACAGGCCCCAG	+239 forward in RPL30
JV703	TGTCTCAAAGACTTGACAGTGGA	+336 reverse in RPL30
JV704	TCTTGACCACCTTGGCTTAAA	+548 forward in RPL30
JV705	CCCATTTTCGGGTAGAAGGT	+666 reverse in RPL30
JV690	TCACTATAGGGCGAATTGGAG	-721 forward in GPD
JV691	CCTATTTTGGGCATGTACGG	-587 reverse in GPD
JV692	AAAACACCAAGAACTTAGTTTCGAC	-64 forward in GPD
JV697	CGATTAAGTTGGGTAACGCCAGGG	+307 reverse in LACZ
JV547	TCTTCCTGAGGCCGATACTG	+480 forward in LACZ
JV548	AATGGGATAGGTTACGTTGGTG	+551 reverse in LACZ
JV549	CGCTGTACTGGAGGCTGAAG	+960 forward in LACZ
JV550	CACCACGCTCATCGATAATTT	+1080 reverse in LACZ
JV551	TTGAAAATGGTCTGCTGCTG	+1259 forward in LACZ
JV552	CGGCGTTAAAGTTGTTCTGC	+1389 reverse in LACZ
JV698	GCGAATACCTGTTCCGTCAT	+2168 forward in LACZ

JV699	ACATCCAGAGGCACTTCACC	+2245 reverse in LACZ
JV682	GGCTGGTGTCGAAATCTTGT	+1107 forward in PMA1
JV683	CTTTCTGGAAGCAGCCAAAC	+1245 reverse in PMA1
JV705.1	CCGGCCTCTATTTTCCATTT	-383 forward in ACT1
JV706	AGAGGCGAGTTTGGTTTCAA	-270 reverse in ACT1
JV707	TTTTTCTTCCCAAGATCGAAA	-16 forward in ACT1
JV708	GGGACCGTGCAATTCTTCT	+47 reverse in ACT1
JV709	CTCGTGCTGTCTTCCCATCT	+388 forward in ACT1
JV710	TGGATTGAGCTTCATCACCA	+470 reverse in ACT1
JV711	TCAAGATCATTGCTCCTCCA	+1288 forward in ACT1
JV712	AGATGGACCACTTTCGTCGT	+1393 reverse in ACT1

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.

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

Abruzzi, K. C., Lacadie, S., and Rosbash, M. (2004). Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. Embo J *23*, 2620-2631.

Allen, T., Shen, P., Samsel, L., Liu, R., Lindahl, L., and Zengel, J. M. (1999). Phylogenetic analysis of L4-mediated autogenous control of the S10 ribosomal protein operon. J Bacteriol *181*, 6124-6132.

Ansari, A., and Schwer, B. (1995). SLU7 and a novel activity, SSF1, act during the PRP16-dependent step of yeast pre-mRNA splicing. Embo J *14*, 4001-4009.

Ares, M., Jr., Grate, L., and Pauling, M. H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. Rna *5*, 1138-1139.

Ast, G. (2004). How did alternative splicing evolve? Nat Rev Genet *5*, 773-782.

Bartels, C., Klatt, C., Luhrmann, R., and Fabrizio, P. (2002). The ribosomal translocase homologue Snu114p is involved in unwinding U4/U6 RNA during activation of the spliceosome. EMBO Rep *3*, 875-880.

Berglund, J. A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. (1997). The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. Cell *89*, 781-787.

Bleichert, F., and Baserga, S. J. (2007). The long unwinding road of RNA helicases. Mol Cell *27*, 339-352.

Bon, E., Casaregola, S., Blandin, G., Llorente, B., Neuveglise, C., Munsterkotter, M., Guldener, U., Mewes, H. W., Van Helden, J., Dujon, B., and Gaillardin, C. (2003). Molecular evolution of eukaryotic genomes: hemiascomycetous yeast spliceosomal introns. Nucleic Acids Res *31*, 1121-1135.

Boon, K. L., Norman, C. M., Grainger, R. J., Newman, A. J., and Beggs, J. D. (2006). Prp8p dissection reveals domain structure and protein interaction sites. Rna *12*, 198-205.

Bousquet-Antonelli, C., Presutti, C., and Tollervey, D. (2000). Identification of a regulated pathway for nuclear pre-mRNA turnover. Cell *102*, 765-775.

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.

Caspary, F., and Seraphin, B. (1998). The yeast U2A'/U2B complex is required for pre-spliceosome formation. Embo J *17*, 6348-6358.

Chan, S. P., and Cheng, S. C. (2005). The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. J Biol Chem *280*, 31190-31199.

Chan, S. P., Kao, D. I., Tsai, W. Y., and Cheng, S. C. (2003). The Prp19p-associated complex in spliceosome activation. Science *302*, 279-282.

Chao, J. A., and Williamson, J. R. (2004). Joint X-ray and NMR refinement of the yeast L30e-mRNA complex. Structure *12*, 1165-1176.

Chen, J. Y., Stands, L., Staley, J. P., Jackups, R. R., Jr., Latus, L. J., and Chang, T. H. (2001). Specific alterations of U1-C protein or U1 small nuclear RNA can eliminate the requirement of Prp28p, an essential DEAD box splicing factor. Mol Cell *7*, 227-232.

Cowper, A. E., Caceres, J. F., Mayeda, A., and Screaton, G. R. (2001). Serine-arginine (SR) protein-like factors that antagonize authentic SR proteins and regulate alternative splicing. J Biol Chem *276*, 48908-48914.

Dabeva, M. D., and Warner, J. R. (1993). Ribosomal protein L32 of Saccharomyces cerevisiae regulates both splicing and translation of its own transcript. J Biol Chem *268*, 19669-19674.

Das, B., Butler, J. S., and Sherman, F. (2003). Degradation of normal mRNA in the nucleus of Saccharomyces cerevisiae. Mol Cell Biol *23*, 5502-5515.

Datta, B., and Weiner, A. M. (1991). Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. Nature *352*, 821-824.

de Souza, S. J., Long, M., Klein, R. J., Roy, S., Lin, S., and Gilbert, W. (1998). Toward a resolution of the introns early/late debate: only phase zero introns are correlated with the structure of ancient proteins. Proc Natl Acad Sci U S A *95*, 5094-5099.

Dong, S., Li, C., Zenklusen, D., Singer, R. H., Jacobson, A., and He, F. (2007). YRA1 autoregulation requires nuclear export and cytoplasmic Edc3p-mediated degradation of its pre-mRNA. Mol Cell *25*, 559-573.

Du, H., and Rosbash, M. (2001). Yeast U1 snRNP-pre-mRNA complex formation without U1snRNA-pre-mRNA base pairing. Rna *7*, 133-142.

Du, H., and Rosbash, M. (2002). The U1 snRNP protein U1C recognizes the 5' splice site in the absence of base pairing. Nature *419*, 86-90.

Eng, F. J., and Warner, J. R. (1991). Structural basis for the regulation of splicing of a yeast messenger RNA. Cell *65*, 797-804.

Fewell, S. W., and Woolford, J. L., Jr. (1999). Ribosomal protein S14 of Saccharomyces cerevisiae regulates its expression by binding to RPS14B pre-mRNA and to 18S rRNA. Mol Cell Biol *19*, 826-834.

Fink, G. R. (1987). Pseudogenes in yeast? Cell *49*, 5-6.

Flaherty, S. M., Fortes, P., Izaurralde, E., Mattaj, I. W., and Gilmartin, G. M. (1997). Participation of the nuclear cap binding complex in pre-mRNA 3' processing. Proc Natl Acad Sci U S A *94*, 11893-11898.

Fong, Y. W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. Nature *414*, 929-933.

Fortes, P., Inada, T., Preiss, T., Hentze, M. W., Mattaj, I. W., and Sachs, A. B. (2000). The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. Mol Cell *6*, 191-196.

Fortes, P., Kufel, J., Fornerod, M., Polycarpou-Schwarz, M., Lafontaine, D., Tollervey, D., and Mattaj, I. W. (1999). Genetic and physical interactions involving the yeast nuclear cap-binding complex. Mol Cell Biol *19*, 6543-6553.

Gietz, D., St Jean, A., Woods, R. A., and Schiestl, R. H. (1992). Improved method for high efficiency transformation of intact yeast cells. Nucleic Acids Res *20*, 1425.

Gorenstein, C., and Warner, J. R. (1976). Coordinate regulation of the synthesis of eukaryotic ribosomal proteins. Proc Natl Acad Sci U S A *73*, 1547-1551.

Gornemann, J., Kotovic, K. M., Hujer, K., and Neugebauer, K. M. (2005). Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. Mol Cell *19*, 53-63.

Grainger, R. J., and Beggs, J. D. (2005). Prp8 protein: at the heart of the spliceosome. Rna *11*, 533-557.

Halic, M., Becker, T., Frank, J., Spahn, C. M., and Beckmann, R. (2005). Localization and dynamic behavior of ribosomal protein L30e. Nat Struct Mol Biol *12*, 467-468.

Hausner, T. P., Giglio, L. M., and Weiner, A. M. (1990). Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. Genes Dev *4*, 2146-2156.

He, F., Li, X., Spatrick, P., Casillo, R., Dong, S., and Jacobson, A. (2003). Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast. Mol Cell *12*, 1439-1452.

Hilliker, A. K., Mefford, M. A., and Staley, J. P. (2007). U2 toggles iteratively between the stem IIa and stem IIc conformations to promote pre-mRNA splicing. Genes Dev *21*, 821-834.

Hirose, Y., and Manley, J. L. (2000). RNA polymerase II and the integration of nuclear events. Genes Dev *14*, 1415-1429.

Holstege, F. C., Jennings, E. G., Wyrick, J. J., Lee, T. I., Hengartner, C. J., Green, M. R., Golub, T. R., Lander, E. S., and Young, R. A. (1998). Dissecting the regulatory circuitry of a eukaryotic genome. Cell *95*, 717-728.

Horowitz, D. S., and Abelson, J. (1993). A U5 small nuclear ribonucleoprotein particle protein involved only in the second step of pre-mRNA splicing in Saccharomyces cerevisiae. Mol Cell Biol *13*, 2959-2970.

Ivanov, A. V., Malygin, A. A., and Karpova, G. G. (2005). Human ribosomal protein S26 suppresses the splicing of its pre-mRNA. Biochim Biophys Acta *1727*, 134-140.

Izaurralde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I. W. (1995). A cap-binding protein complex mediating U snRNA export. Nature *376*, 709-712.

Jones, M. H., Frank, D. N., and Guthrie, C. (1995). Characterization and functional ordering of Slu7p and Prp17p during the second step of pre-mRNA splicing in yeast. Proc Natl Acad Sci U S A *92*, 9687-9691.

Kim, S. H., and Lin, R. J. (1996). Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. Mol Cell Biol *16*, 6810-6819.

Kistler, A. L., and Guthrie, C. (2001). Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. Genes Dev *15*, 42-49.

Konarska, M. M., and Sharp, P. A. (1987). Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. Cell *49*, 763-774.

Konarska, M. M., Vilardell, J., and Query, C. C. (2006). Repositioning of the reaction intermediate within the catalytic center of the spliceosome. Mol Cell *21*, 543-553.

Lacadie, S. A., and Rosbash, M. (2005). Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. Mol Cell *19*, 65-75.

Lamond, A. I., Konarska, M. M., Grabowski, P. J., and Sharp, P. A. (1988). Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. Proc Natl Acad Sci U S A *85*, 411-415.

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

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

Li, B., Vilardell, J., and Warner, J. R. (1996). An RNA structure involved in feedback regulation of splicing and of translation is critical for biological fitness. Proc Natl Acad Sci U S A *93*, 1596-1600.

Li, Z., Paulovich, A. G., and Woolford, J. L., Jr. (1995). Feedback inhibition of the yeast ribosomal protein gene CRY2 is mediated by the nucleotide sequence and secondary structure of CRY2 pre-mRNA. Mol Cell Biol *15*, 6454-6464.

Libri, D., Graziani, N., Saguez, C., and Boulay, J. (2001). Multiple roles for the yeast SUB2/yUAP56 gene in splicing. Genes Dev *15*, 36-41.

Liu, L., Query, C. C., and Konarska, M. M. (2007). Opposing classes of prp8 alleles modulate the transition between the catalytic steps of pre-mRNA splicing. Nat Struct Mol Biol *14*, 519-526.

Madhani, H. D., and Guthrie, C. (1992). A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. Cell *71*, 803-817.

Malygin, A. A., Parakhnevitch, N. M., Ivanov, A. V., Eperon, I. C., and Karpova, G. G. (2007). Human ribosomal protein S13 regulates expression of its own gene at the splicing step by a feedback mechanism. Nucleic Acids Res *35*, 6414-6423.

Maniatis, T., and Tasic, B. (2002). Alternative pre-mRNA splicing and proteome expansion in metazoans. Nature *418*, 236-243.

Mayas, R. M., Maita, H., and Staley, J. P. (2006). Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nat Struct Mol Biol *13*, 482-490.

McPheeters, D. S., and Muhlenkamp, P. (2003). Spatial organization of protein-RNA interactions in the branch site-3' splice site region during pre-mRNA splicing in yeast. Mol Cell Biol *23*, 4174-4186.

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

Moore, M. J., Schwartzfarb, E. M., Silver, P. A., and Yu, M. C. (2006). Differential recruitment of the splicing machinery during transcription predicts genome-wide patterns of mRNA splicing. Mol Cell *24*, 903-915.

Newman, A. J., and Norman, C. (1992). U5 snRNA interacts with exon sequences at 5' and 3' splice sites. Cell *68*, 743-754.

Nilsen, T. W. (2005). Spliceosome assembly in yeast: one ChIP at a time? Nat Struct Mol Biol *12*, 571-573.

Nomura, M., Gourse, R., and Baughman, G. (1984). Regulation of the synthesis of ribosomes and ribosomal components. Annu Rev Biochem *53*, 75-117.

Orphanides, G., and Reinberg, D. (2002). A unified theory of gene expression. Cell *108*, 439-451.

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.

Perriman, R., and Ares, M., Jr. (2000). ATP can be dispensable for prespliceosome formation in yeast. Genes Dev *14*, 97-107.

Perriman, R., Barta, I., Voeltz, G. K., Abelson, J., and Ares, M., Jr. (2003). ATP requirement for Prp5p function is determined by Cus2p and the structure of U2 small nuclear RNA. Proc Natl Acad Sci U S A *100*, 13857-13862.

Perriman, R. J., and Ares, M., Jr. (2007). Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. Genes Dev *21*, 811-820.

Preker, P. J., and Guthrie, C. (2006). Autoregulation of the mRNA export factor Yra1p requires inefficient splicing of its pre-mRNA. Rna *12*, 994-1006.

Preker, P. J., Kim, K. S., and Guthrie, C. (2002). Expression of the essential mRNA export factor Yra1p is autoregulated by a splicing-dependent mechanism. Rna *8*, 969-980.

Proudfoot, N. J., Furger, A., and Dye, M. J. (2002). Integrating mRNA processing with transcription. Cell *108*, 501-512.

Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., Wilm, M., and Seraphin, B. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods *24*, 218-229.

Query, C. C., and Konarska, M. M. (2004). Suppression of multiple substrate mutations by spliceosomal prp8 alleles suggests functional correlations with ribosomal ambiguity mutants. Mol Cell *14*, 343-354.

Query, C. C., and Konarska, M. M. (2006). Splicing fidelity revisited. Nat Struct Mol Biol *13*, 472-474.

Raghunathan, P. L., and Guthrie, C. (1998). RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. Curr Biol *8*, 847-855.

Rain, J. C., and 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.

Rappsilber, J., Ryder, U., Lamond, A. I., and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. Genome Res *12*, 1231-1245.

Reed, R. (2003). Coupling transcription, splicing and mRNA export. Curr Opin Cell Biol *15*, 326-331.

Reed, R., and Hurt, E. (2002). A conserved mRNA export machinery coupled to pre-mRNA splicing. Cell *108*, 523-531.

Roy, J., Kim, K., Maddock, J. R., Anthony, J. G., and Woolford, J. L., Jr. (1995). The final stages of spliceosome maturation require Spp2p that can interact with the DEAH box protein Prp2p and promote step 1 of splicing. Rna *1*, 375-390.

Ruskin, B., and Green, M. R. (1985). Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing. Nature *317*, 732-734.

Ruskin, B., Zamore, P. D., and Green, M. R. (1988). A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell *52*, 207-219.

Rutz, B., and Seraphin, B. (1999). Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly. Rna *5*, 819-831.

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.

Schwer, B., and Gross, C. H. (1998). Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. Embo J *17*, 2086-2094.

Schwer, B., and Guthrie, C. (1991). PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. Nature *349*, 494-499.

Schwer, B., and Guthrie, C. (1992). A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. Embo J *11*, 5033-5039.

Seraphin, B., Kretzner, L., and Rosbash, M. (1988). A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. Embo J *7*, 2533-2538.

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

Shatkin, A. J., and Manley, J. L. (2000). The ends of the affair: capping and polyadenylation. Nat Struct Biol *7*, 838-842.

Siliciano, P. G., and Guthrie, C. (1988). 5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. Genes Dev *2*, 1258-1267.

Small, E. C., Leggett, S. R., Winans, A. A., and Staley, J. P. (2006). The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase. Mol Cell *23*, 389-399.

Sontheimer, E. J., and Steitz, J. A. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. Science *262*, 1989-1996.

Spingola, M., and Ares, M., Jr. (2000). A yeast intronic splicing enhancer and Nam8p are required for Mer1p-activated splicing. Mol Cell *6*, 329-338.

Spingola, M., Armisen, J., and Ares, M., Jr. (2004). Mer1p is a modular splicing factor whose function depends on the conserved U2 snRNP protein Snu17p. Nucleic Acids Res *32*, 1242-1250.

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

Staley, J. P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell *92*, 315-326.

Staley, J. P., and Guthrie, C. (1999). An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. Mol Cell *3*, 55-64.

Stevens, S. W., Ryan, D. E., Ge, H. Y., Moore, R. E., Young, M. K., Lee, T. D., and Abelson, J. (2002). Composition and functional characterization of the yeast spliceosomal penta-snRNP. Mol Cell *9*, 31-44.

Strasser, K., Masuda, S., Mason, P., Pfannstiel, J., Oppizzi, M., Rodriguez-Navarro, S., Rondon, A. G., Aguilera, A., Struhl, K., Reed, R., and Hurt, E. (2002). TREX is a conserved complex coupling transcription with messenger RNA export. Nature *417*, 304-308.

Strauss, E. J., and Guthrie, C. (1991). A cold-sensitive mRNA splicing mutant is a member of the RNA helicase gene family. Genes Dev *5*, 629-641.

Tagwerker, C., Zhang, H., Wang, X., Larsen, L. S., Lathrop, R. H., Hatfield, G. W., Auer, B., Huang, L., and Kaiser, P. (2006). HB tag modules for PCR-based gene tagging and tandem affinity purification in Saccharomyces cerevisiae. Yeast *23*, 623-632.

Tardiff, D. F., Lacadie, S. A., and Rosbash, M. (2006). A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. Mol Cell *24*, 917-929.

Tardiff, D. F., and Rosbash, M. (2006). Arrested yeast splicing complexes indicate stepwise snRNP recruitment during in vivo spliceosome assembly. Rna *12*, 968-979.

Tarn, W. Y., Hsu, C. H., Huang, K. T., Chen, H. R., Kao, H. Y., Lee, K. R., and Cheng, S. C. (1994). Functional association of essential splicing factor(s) with PRP19 in a protein complex. Embo J *13*, 2421-2431.

Tarn, W. Y., Lee, K. R., and Cheng, S. C. (1993). Yeast precursor mRNA processing protein PRP19 associates with the spliceosome concomitant with or just after dissociation of U4 small nuclear RNA. Proc Natl Acad Sci U S A *90*, 10821-10825.

Torchet, C., Bousquet-Antonelli, C., Milligan, L., Thompson, E., Kufel, J., and Tollervey, D. (2002). Processing of 3'-extended read-through transcripts by the exosome can generate functional mRNAs. Mol Cell *9*, 1285-1296.

Tsai, R. T., Tseng, C. K., Lee, P. J., Chen, H. C., Fu, R. H., Chang, K. J., Yeh, F. L., and Cheng, S. C. (2007). Dynamic interactions of Ntr1-Ntr2 with Prp43 and with U5 govern the recruitment of Prp43 to mediate spliceosome disassembly. Mol Cell Biol *27*, 8027-8037.

Udem, S. A., and Warner, J. R. (1972). Ribosomal RNA synthesis in Saccharomyces cerevisiae. J Mol Biol *65*, 227-242.

Umen, J. G., and Guthrie, C. (1995a). Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. Rna *1*, 584-597.

Umen, J. G., and Guthrie, C. (1995b). The second catalytic step of pre-mRNA splicing. Rna *1*, 869-885.

Umen, J. G., and Guthrie, C. (1996). Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3' splice site selection. Genetics *143*, 723-739.

Valcarcel, J., Gaur, R. K., Singh, R., and Green, M. R. (1996). Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. Science *273*, 1706-1709.

Vilardell, J., Chartrand, P., Singer, R. H., and Warner, J. R. (2000a). The odyssey of a regulated transcript. Rna *6*, 1773-1780.

Vilardell, J., and Warner, J. R. (1994). Regulation of splicing at an intermediate step in the formation of the spliceosome. Genes Dev *8*, 211-220.

Vilardell, J., and Warner, J. R. (1997). Ribosomal protein L32 of Saccharomyces cerevisiae influences both the splicing of its own transcript and the processing of rRNA. Mol Cell Biol *17*, 1959-1965.

Vilardell, J., Yu, S. J., and Warner, J. R. (2000b). Multiple functions of an evolutionarily conserved RNA binding domain. Mol Cell *5*, 761-766.

Villa, T., and Guthrie, C. (2005). The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. Genes Dev *19*, 1894-1904.

Vinogradov, A. E. (2001). Intron length and codon usage. J Mol Evol *52*, 2-5.

Wang, Q., Zhang, L., Lynn, B., and Rymond, B. C. (2008). A BBP-Mud2p heterodimer mediates branchpoint recognition and influences splicing substrate abundance in budding yeast. Nucleic Acids Res.

Warner, J. R., Mitra, G., Schwindinger, W. F., Studeny, M., and Fried, H. M. (1985). Saccharomyces cerevisiae coordinates accumulation of yeast ribosomal proteins by modulating mRNA splicing, translational initiation, and protein turnover. Mol Cell Biol *5*, 1512-1521.

Wassarman, D. A., and Steitz, J. A. (1992). Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. Science *257*, 1918-1925.

White, S. A., Hoeger, M., Schweppe, J. J., Shillingford, A., Shipilov, V., and Zarutskie, J. (2004). Internal loop mutations in the ribosomal protein L30 binding site of the yeast L30 RNA transcript. Rna *10*, 369-377.

Wu, J. A., and Manley, J. L. (1991). Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. Nature *352*, 818-821.

Wu, S., Romfo, C. M., Nilsen, T. W., and Green, M. R. (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature *402*, 832-835.

Wyatt, J. R., Sontheimer, E. J., and Steitz, J. A. (1992). Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. Genes Dev *6*, 2542-2553.

Xu, Y. Z., and Query, C. C. (2007). Competition between the ATPase Prp5 and Branch Region-U2 snRNA Pairing Modulates the Fidelity of Spliceosome Assembly. Mol Cell *28*, 838-849.

Zamore, P. D., and Green, M. R. (1989). Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. Proc Natl Acad Sci U S A *86*, 9243-9247.

Zamore, P. D., Patton, J. G., and Green, M. R. (1992). Cloning and domain structure of the mammalian splicing factor U2AF. Nature *355*, 609-614.

Zengel, J. M., and Lindahl, L. (1994). Diverse mechanisms for regulating ribosomal protein synthesis in Escherichia coli. Prog Nucleic Acid Res Mol Biol *47*, 331-370.

Zhang, M., and Green, M. R. (2001). Identification and characterization of yUAP/Sub2p, a yeast homolog of the essential human pre-mRNA splicing factor hUAP56. Genes Dev *15*, 30-35.

Zhuang, Y., and Weiner, A. M. (1986). A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell *46*, 827-835.