Partial purification of the yeast U2 snRNP reveals a novel yeast pre-mRNA splicing factor required for pre-spliceosome assembly

Friederike Caspary, Anna Shevchenko, Matthias Wilm and Bertrand Séraphin[1]

EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

[1]Corresponding author
e-mail: Seraphin@EMBL-Heidelberg.de

We have partially purified the U2 snRNP of *Saccharomyces cerevisiae*. Identification of some proteins consistently found in the purified fractions by nanoelectrospray mass spectrometry indicated the presence of a novel splicing factor named Rse1p. The *RSE1* gene is essential and codes for a 148.2 kDa protein. We demonstrated that Rse1p associates specifically with U2 snRNA at low salt concentrations. In addition, we showed that Rse1p is a component of the pre-spliceosome. Depletion of Rse1p and analysis of a conditional mutant indicated that Rse1p was required for efficient splicing *in vivo*. *In vitro* Rse1p is required for the formation of pre-spliceosomes. Database searches revealed that Rse1p is conserved in humans and that it belongs to a large protein family that includes polyadenylation factors and DNA repair proteins. The characteristics of Rse1p suggest that its human homologue could be a subunit of the SF3 splicing factor.

Keywords: DNA repair/polyadenylation/*RSE1*/
Saccharomyces cerevisiae/SF3

Introduction

Removal of introns from nuclear primary transcripts (pre-mRNA) occurs by two transesterification reactions that take place in a multicomponent complex termed the spliceosome (Moore *et al.*, 1993). The spliceosome is built by the ordered association of the U1, U2, U4, U5 and U6 small nuclear ribonucleoprotein particles (snRNPs) and extrinsic protein factors with the pre-mRNA (reviewed by Moore *et al.*, 1993; Krämer, 1996; Will and Lührmann, 1997; Staley and Guthrie, 1998). This process is initiated by interaction of U1 snRNP with the pre-mRNA 5′ splice site (5′SS) leading to the formation of commitment complex 1 in yeast (CC1; Séraphin and Rosbash, 1989a, 1991; Rosbash and Séraphin, 1991). In the subsequent step, the non-snRNP splicing factors, BBP/SF1 and Mud2p/U2AF^{65} (Krämer, 1992; Zamore *et al.*, 1992; Abovich *et al.*, 1994; Abovich and Rosbash, 1997; Berglund *et al.*, 1998), recognize downstream regions of the intron, allowing for formation of yeast commitment complex 2 (CC2). The spliceosome assembly pathway in mammals appears to be very similar. E complexes identified in this system (Michaud and Reed, 1991) are thought to be functionally equivalent to commitment complexes. The U2

snRNP, which interacts with the pre-mRNA branchpoint region (BP), is added to these early complexes in an ATP-dependent step, resulting in the formation of the pre-spliceosome (reviewed by Moore *et al.*, 1993). Finally, U4, U5 and U6 snRNPs, pre-assembled in a single particle, join in and form the mature spliceosome, in which the pre-mRNA splicing reaction occurs.

Numerous proteins, which are either components of snRNPs or free splicing factors, are involved in spliceosome assembly. Some of these proteins were identified genetically in yeast while others were characterized biochemically in various systems. Using these tools, the yeast U1 snRNP, involved in the earliest step of spliceosome formation, recently has been characterized extensively (Neubauer *et al.*, 1997; Gottschalk *et al.*, 1998). This progress resulted in part from important improvements in protein identification by mass spectrometry (Neubauer *et al.*, 1997). In contrast, our knowledge of the yeast U2 snRNP is much more modest. Besides the common Sm proteins (Salgado-Garrido *et al.*, 1999), two U2 snRNP-specific proteins have been characterized in yeast. The Yib9 protein (Yib9p) was identified by similarity with its human homologue U2B″ (Polycarpou-Schwarz *et al.*, 1996) and by genetic means (Tang *et al.*, 1996). More recently, the Lea1 protein (Lea1p) was shown to be the homologue of metazoan U2A′ (Caspary and Séraphin, 1998). Both yeast proteins are required specifically for the formation of the pre-spliceosome (Caspary and Séraphin, 1998). Biochemical fractionation of human cell extracts has demonstrated that the multisubunit SF3a and SF3b factors are required, in addition to U2 snRNP, for the formation of the pre-spliceosome (reviewed in Krämer, 1996; Will and Lührmann, 1997). Simultaneously, characterization of the human U2 snRNP revealed the presence of distinct forms in HeLa nuclear extracts. The 12S U2 snRNP contains the Sm, U2A′ and U2B″ proteins associated with the U2 snRNA, whereas the salt-labile 17S U2 snRNP harbours additional proteins (Behrens *et al.*, 1993b). These additional factors were found to be identical to the SF3a and SF3b subunits (Brosi *et al.*, 1993a; reviewed in Krämer, 1996; Will and Lührmann, 1997). Both complexes are thought to interact with the highly conserved 5′ end of U2 snRNA (Behrens *et al.*, 1993b). The same factors were also identified during affinity purification of the mammalian pre-spliceosome (Bennett *et al.*, 1992; reviewed by Hodges and Beggs, 1994). It has been shown that they enter the spliceosome at the same time as U2 snRNA and remain associated with the spliceosome until the splicing reaction is completed.

The SF3a complex has been characterized in detail and has been shown to consist of three proteins named SAP61/SF3a^{60}, SAP62/SF3a^{66} and SAP114/SF3a^{120} (Bennett and Reed, 1993; Brosi *et al.*, 1993b; reviewed by Krämer, 1996; Reed, 1996). These proteins are similar to the yeast

PRP9, *PRP11* and *PRP21* gene products which have also been shown to associate in a multisubunit complex (Legrain and Chapon, 1993; Ruby *et al.*, 1993; Rain *et al.*, 1996; for reviews see Hodges and Beggs, 1994; Krämer, 1996). SF3b consists of at least four subunits, SAP49, SAP130, SAP145 and SAP155. Except for SAP130, these proteins can be cross-linked to the pre-mRNA region flanking the branchpoint (Gozani *et al.*, 1996, 1998). Therefore, it has been proposed that these proteins would anchor the U2 snRNP to the pre-mRNA (Gozani *et al.*, 1996, 1998; Krämer, 1996; Will and Lührmann, 1997). Analysis of the SAP49, SAP145 and SAP155 sequences revealed that these proteins have homologues in yeast which have been named HSH49 (Igel *et al.*, 1998), CUS1 (Gozani *et al.*, 1996; Wells *et al.*, 1996) and ySAP155 (Wang *et al.*, 1998), respectively.

We have initiated a biochemical characterization of the U2 snRNP. This led to the identification of a novel essential splicing factor, Rse1p. Rse1p is required for efficient splicing *in vivo* and pre-spliceosome formation *in vitro*. Surprisingly, Rse1p belongs to a large protein family including polyadenylation factors and DNA repair proteins. Putative Rse1p homologues exist in human and *Caenorhabditis elegans*. Our results suggest that the human Rse1p homologue could be a component of the multi-subunit SF3 splicing factor.

Results

Identification of Rse1p as a U2 snRNP-associated protein

We partially purified the yeast U2 snRNP using the TAP (tandem affinity purification) procedure (G.Rigaut, A.Shevchenko, B.Rutz, M.Wilm, M.Mann and B.Séraphin, submitted). For this purpose, we first introduced the TAP tag at the C-terminus of Lea1p that encodes the yeast homologue of the human U2A′ protein (Caspary and Séraphin, 1998). Starting with 5 l of cell culture, proteins and RNAs associated with Lea1p were purified following the two-step affinity purification of the TAP method. Primer extension analysis of the RNAs present in the input and purified fractions revealed that the U2 spliceosomal snRNA was selected specifically and that ~5% of the starting U2 snRNA was recovered in the partially purified fraction (data not shown). The proteins present in these fractions were concentrated by trichloroacetic acid (TCA) precipitation and loaded on a 7–25% exponential gradient SDS–polyacrylamide gel. Proteins were visualized by silver staining (Shevchenko *et al.*, 1996). Several proteins present at low levels could be detected in the second and third fraction of the final elution (Figure 1, lanes 1 and 2).

To verify reproducibility, we used a slightly different purification approach. The two different units of the TAP tag, namely TAP-A and TAP-C (G.Rigaut, A.Shevchenko, B.Rutz, M.Wilm, M.Mann and B.Séraphin, submitted), were fused to two different protein subunits of the U2 snRNP. In this case, the TAP-A tag was fused to the C-terminus of the common snRNP core protein SmB and the TAP-C tag was fused to the C-terminus of Lea1p (Gottschalk *et al.*, 1998). The first step selects for the U1, U2, U4 and U5 snRNPs while the second step is specific for U2 snRNP. Analysis of RNA purified under these new conditions indicated that U2 snRNA was again selected
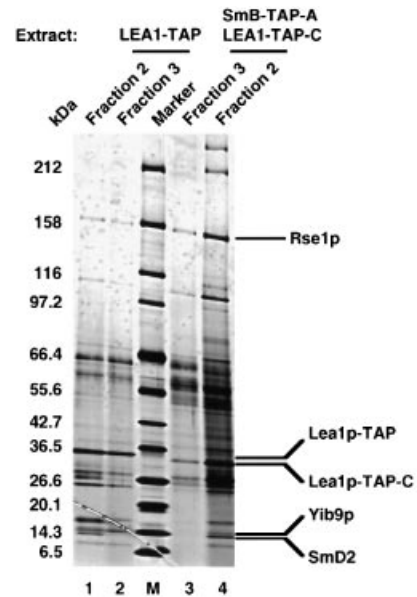

Fig. 1. Identification of proteins associated with U2 snRNP. U2 snRNP was partially purified from extracts harbouring either the Lea1p–TAP or the Lea1p–TAP-C and SmB–TAP-A fusions. Proteins of the second and the third final elution fractions were concentrated by TCA precipitation and loaded on a 7–25% gradient SDS–polyacrylamide gel. After silver staining, proteins were identified by nanoelectrospray mass spectrometry. Lea1p and Yib9p were also detected by Western blotting. Identified proteins are shown on the right. In addition to the proteins indicated, we also identified the faint protein band migrating with an apparent molecular weight of 37 kDa (just above Lea1p) as glyceraldehyde-3-phosphate dehydrogenase. Unfortunately, we could not identify the additional proteins consistently present in the different preparations (e.g. the band migrating with an apparent mol. wt of 100 kDa) due to the low amount of material.

specifically after the second step and that the efficiency was quite similar.

Comparison of the proteins recovered by the two approaches (Figure 1) revealed that the main bands were essentially identical. As expected, Lea1p that had been fused to tags of different sizes migrated with slightly different apparent molecular weight in the two purifications (Figure 1). Western blot experiments using antibodies specific for Lea1p confirmed this conclusion (data not shown). Yib9p could also be identified in the U2 snRNP fractions using specific antibodies (data not shown). Although the protein levels were very low, four bands that were present consistently in several partially purified fractions could be analysed by nanoelectrospray mass spectrometry (Shevchenko *et al.*, 1996; Wilm and Mann, 1996). This confirmed that the 32 kDa protein was the Lea1p–TAP fusion protein (four peptides sequenced). The two bands running at 14 kDa contained the U2 snRNP-specific Yib9p (upper band; two peptides sequenced; Polycarpou-Schwarz *et al.*, 1996; Tang *et al.*, 1996) and the common snRNP core protein, SmD2 (lower band; two peptides sequenced; Salgado-Garrido *et al.*, 1999). These results indicated that the partially purified fraction contained significant amounts of U2 snRNP-associated proteins.

One peptide sequence from the ~155 kDa protein band unambiguously matched the yeast open reading frame (ORF) YML049C encoded on chromosome XIII which recently has been named *RSE1*. *RSE1* stands for RNA

splicing and ER to Golgi transport factor 1, because Rse1p recently was reported to affect protein transport to the Golgi apparatus by preventing splicing of the SAR1 pre-mRNA (Chen *et al.*, 1998). The predicted size of the protein encoded by this ORF (153.8 kDa) was very similar to the apparent size of the protein on SDS–PAGE gels, supporting our identification.

The RSE1 gene is essential

To determine whether the *RSE1* gene is essential, we replaced a copy of the corresponding ORF with the *Kluyveromyces lactis URA3* marker in a diploid strain. After sporulation, tetrads were dissected. Ten out of 13 tetrads gave rise to two viable spores, while the two other spores germinated but stopped growing after a few divisions. For the three other tetrads, only one spore was growing. In this case, some but not all of the remaining spores germinated. All growing spores were *ura3⁻*, indicating that they carried the wild-type *RSE1* gene (data not shown). Therefore, we conclude that *RSE1* is an essential gene.

Rse1p is associated specifically with the U2 snRNA

To test whether Rse1p is associated specifically with U2 snRNA, we performed immunoprecipitation experiments with extracts containing a tagged version of Rse1p. We first constructed a tagged version of Rse1p by fusing a cassette containing the *GAL* promoter and two IgG-binding domains of the *Staphylococcus aureus* protein A (ProtA; Lafontaine and Tollervey, 1996) to the predicted translation start site of the YML049C ORF. This fusion did not generate any phenotype, indicating that it was fully functional. However, when we attempted to deplete the tagged protein by incubating the corresponding strain in glucose-containing medium, growth (monitored for several days; data not shown) and splicing of endogenous pre-mRNAs were not affected even though the tagged protein level analysed by Western blot was reduced far below 1% of the initial level (data not shown). Comparison of the Rse1 protein sequence with the sequences of related proteins (see also below) indicated that the first 50 N-terminal residues of the YML049C ORF were not present in the related proteins (see Figure 6), suggesting strongly that they were also unlikely to be present in the endogenous Rse1p. This fusion was therefore likely to contain 150 bp of genomic DNA upstream of the natural Rse1p translational start site, and this sequence was possibly sufficient to drive expression of untagged Rse1p which would explain our result.

To avoid this problem, we generated a new strain in which the *GAL* promoter–protein A cassette was fused to residue 51 of the YML049C ORF. A haploid strain carrying this new fusion protein (ProtA–Rse1p) had a growth rate comparable with that of an isogenic wild-type strain in galactose medium, indicating that the tagged protein was fully functional. Extracts prepared from these cells (ProtA–RSE1) and from an isogenic wild-type strain (negative control) were incubated with IgG–agarose beads. The bound (pellets) and unbound (supernatants) fractions, as well as the starting extracts (input) were analysed for their protein and RNA content (Figure 2, and data not shown). Western blot analysis confirmed that the ProtA–Rse1p was expressed and precipitated efficiently (data not
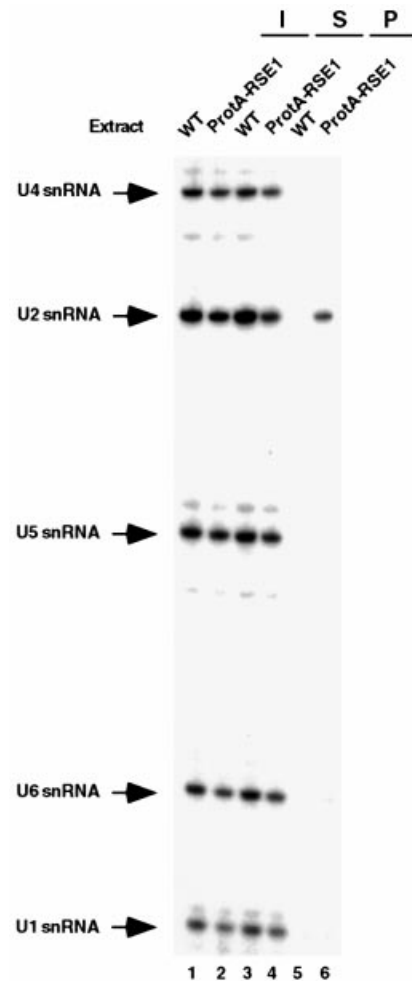

Fig. 2. Co-immunoprecipitation of U2 snRNA by ProtA–Rse1p. Extracts from a control wild-type yeast strain (WT) and a strain harbouring protein A-tagged Rse1p (ProtA–Rse1p) were immunoprecipitated with IgG–agarose beads. The extracted RNAs of input (lanes 1 and 2), supernatant (lanes 3 and 4) and pellet (lanes 5 and 6) fractions were analysed by primer extension with specific primers for U1, U2, U4, U5 and U6 snRNAs. The positions of the corresponding signals are shown on the left. RNA from 10 times more extract was used for the pellets relative to the input and supernatant fractions.

shown). Primer extension analysis revealed that the U2 snRNA was the only spliceosomal snRNA co-precipitated with ProtA–Rse1p (Figure 2, lane 6). This co-precipitation was specific since no signal could be detected with the control extract derived from a wild-type strain (Figure 2, lane 5) and since only background levels of U1, U4, U5 and U6 snRNAs were detected in the pellet (Figure 2, lane 6). Quantification of the signals present in the input and pellet fraction indicated that only ~10% of the total U2 snRNA was co-precipitated with ProtA–Rse1p (Figure 2, lanes 2 and 6; note that 10 times more material was used in the pellet fraction). For comparison, ~50% of the U2 snRNA can be co-precipitated with a Lea1p–ProtA fusion (Caspary and Séraphin, 1998). This suggested that Rse1p might be only partially associated with the U2 snRNA and/or that the 150 mM NaCl present during the immunoprecipitation prevented efficient interaction. To test the latter possibility, we repeated the immunoprecipitation experiment at various salt concentrations. At 50 mM
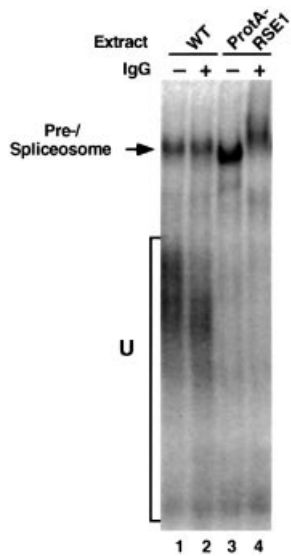
Fig. 3. Rse1p is a component of the pre-spliceosome. Splicing complexes were assembled on radioactively labelled wild-type pre-mRNA in the presence of ATP using extracts from either wild-type (WT) or ProtA–Rse1p (ProtA–RSE1)-harbouring strains. Complexes were incubated with (lanes 2 and 4) or without (lanes 1 and 3) IgG and fractionated by native gel electrophoresis. Pre-/spliceosome indicates the position of the non-shifted pre-spliceosome and spliceosome in this system; U indicates the free pre-mRNA and unspecific protein–pre-mRNA complexes (Séraphin and Rosbash, 1989b).

salt concentration the specificity was lost while at 300 mM salt concentration co-precipitation of U2 snRNA was reduced to nearly background level (data not shown). Similar results were obtained with the construct carrying the 50 extra amino acids between ProtA and Rse1p (data not shown). These experiments demonstrate that Rse1p associates specifically with the U2 snRNA. However, this interaction is salt sensitive and the association is probably only partial and/or transient. This suggests that Rse1p is not a core U2 snRNP protein.

Rse1p is a component of the pre-spliceosome

The experiments described above identified Rse1p as a U2 snRNP-associated protein. It was, however, unclear whether Rse1p was involved in pre-mRNA splicing or rather in some steps of snRNP metabolism. To address this question, we tested whether Rse1p is present in splicing complexes. Splicing extracts prepared from either a wild-type strain or a strain carrying the ProtA–Rse1p construct (see above) were incubated with a radioactively labelled wild-type pre-mRNA to allow spliceosome assembly. IgG was added to these reactions and splicing complexes were detected following native gel electrophoresis (Figure 3). Pre-spliceosomes and spliceosomes (pre-/spliceosomes) formed in the wild-type extract migrated with the same mobility in the presence or absence of IgG (Figure 3, lanes 1 and 2). In contrast, pre-/spliceosomes formed in the extract containing the ProtA–Rse1p were clearly shifted upwards in the presence of IgG (Figure 3, compare lanes 3 and 4). This demonstrates that Rse1p is a component of the pre-/spliceosome. However, as the pre-spliceosome and mature spliceosome cannot be resolved on this native gel system (Séraphin and Rosbash, 1989a), we could not conclude whether Rse1p was present in pre-

spliceosomes. To test this possibility, we repeated this experiment using extracts that had been pre-incubated in the presence of a DNA oligonucleotide complementary to U6 snRNA. This treatment leads to RNase H-mediated U6 snRNA degradation (the level of U6 snRNA was determined by primer extension, data not shown) and prevents mature spliceosome assembly (Fabrizio *et al.*, 1989). Pre-spliceosomes formed under these conditions were shifted in the presence of IgG (data not shown). These data indicated that Rse1p is at least a component of the pre-spliceosome.

Rse1p is required for efficient splicing in vivo

The data presented above indicate that Rse1p is likely to be a splicing factor. To test this possibility, we first used a strain containing a C-terminal TAP- (G.Rigaut, A.Shevchenko, B.Rutz, M.Wilm, M.Mann and B.Séraphin, submitted) tagged Rse1p (Rse1p–TAP). A haploid strain carrying this construct was viable but thermosensitive, suggesting that this particular protein fusion was only partially functional (data not shown). This strain, as well as an isogenic wild-type strain for comparison, was transformed with a set of reporter plasmids containing *RP51A* intron derivatives inserted into the *lacZ* coding sequence (Teem and Rosbash, 1983). Four reporter constructs were selected. They contained either the wild-type *RP51A* intron or a mutated *RP51A* intron with a strong mutation in the 5′ splice site (5′SS; GUAU**a**U, mutated base in lower case and bold) and/or a weak mutation in the branchpoint region (BP; UA**u**UAAC). The mutated reporter constructs made the assay more sensitive, allowing also for the detection of subtle defects in splicing. An empty vector was used as a negative control. Transformants were shifted to 37°C (overnight) and RNA splicing was assayed by following β-galactosidase production (data not shown) as well as through direct RNA analysis by primer extension (Figure 4A). The *RSE1*–TAP allele only mildly affected production of the endogenous *RP51A* mRNAs as well as wild-type and branchpoint mutant reporters (Figure 4A, lanes 1, 2, 4, 6, 8 and 10). Splicing of the 5′ splice site mutant pre-mRNAs was, however, affected in the tagged strain, as shown by the accumulation of reporter pre-mRNA and the decreased levels of reporter mRNA (Figure 4A, compare lanes 3 and 5 with 7 and 9, respectively). Quantitative analysis of splicing efficiencies (ratio of mRNA to pre-mRNA, M/P; Pikielny and Rosbash, 1985; Figure 4A, bottom) confirmed that processing of the 5′ splice site mutant pre-mRNA was affected most by the tagged Rse1p. The other reporters were also affected, although to a lesser extent (Figure 4A, bottom). Similar results were obtained by measuring β-galactosidase activity (data not shown). At 30°C, the overall effect was quantitatively weaker, consistent with the thermosensitive phenotype of the strain (data not shown).

We also depleted the ProtA–Rse1p that was under the control of the *GAL* promoter by incubating the corresponding strain in glucose-containing medium. The strain stopped growing after ~30 h with concomitant reduction of the ProtA–Rse1p level (see below). ProtA–RSE1 and control wild-type cells were therefore collected at three different time points after transfer to glucose medium, and splicing of the endogenous U3 precursor snoRNAs was analysed by primer extension. At all three time points,
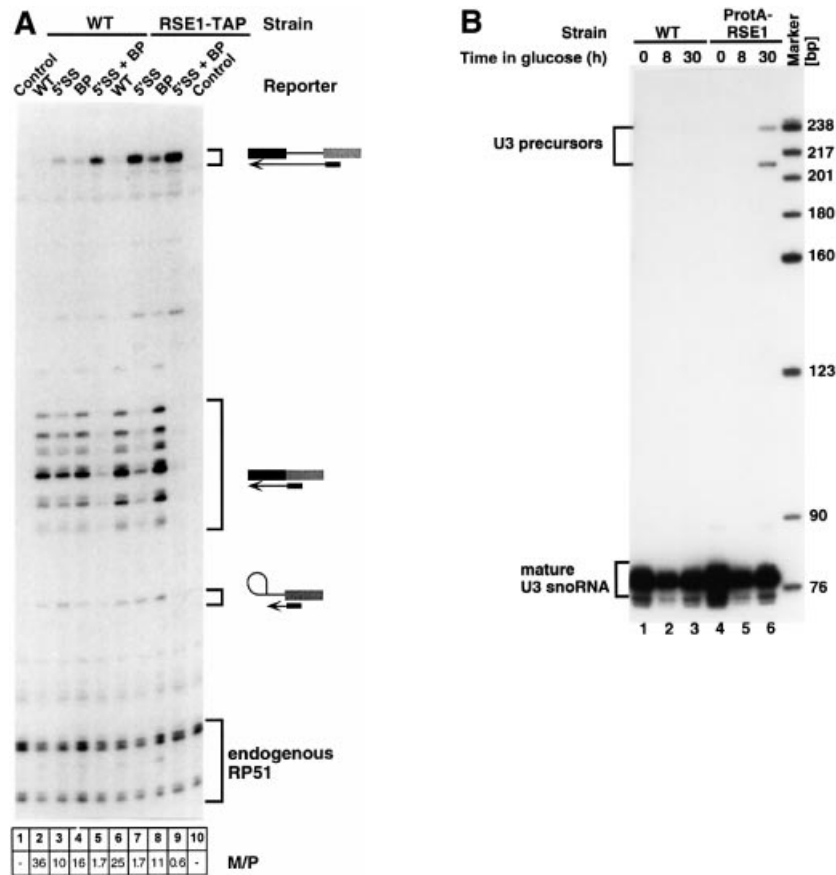
Fig. 4. Rse1p is involved in splicing *in vivo*. (**A**) Reporter constructs were introduced into a wild-type strain (WT) and a strain harbouring the C-terminal tagged Rse1p (RSE1–TAP). These reporters were: control, vector without reporter gene (lanes 1 and 10); WT, wild-type *RP51A* intron (lanes 2 and 6); 5′SS, 5′ splice site mutation (GUAUaU) in the *RP51A* intron (lanes 3 and 7); BP, branch point mutation (UAuUAAC) in the *RP51A* intron (lanes 4 and 8); and 5′SS + BP, combination of the two mutations (lanes 5 and 9). The different extension products are labelled on the right. Signals were quantified using a phosphoimager (Fuji). The ratio of mRNA to pre-mRNA was calculated and is indicated below the figure (M/P). Background signals in control lanes were not analysed. (**B**) U3 precursor RNA accumulates in genetically ProtA–Rse1p-depleted strains. A ProtA–Rse1p-containing strain (ProtA–RSE1) and an isogenic wild-type strain (WT) were grown in glucose medium. Total RNA was extracted at the indicated time points, and U3 precursor and mature RNA were analysed by primer extension.

only the mature U3 snoRNA was detected in the wild-type strain (Figure 4B, lanes 1–3). The same situation was observed for the ProtA–RSE1 strain grown in glucose medium for 0 or 8 h (Figure 4B, lanes 4 and 5). However, after 30 h of growth in glucose medium, we observed accumulation of the U3 precursor snoRNAs (Figure 4B, lane 6). As both a conditional *RSE1* allele and depletion of Rse1p lead to *in vivo* splicing defects, we conclude that Rse1p is a novel general splicing factor.

Rse1p depletion prevents pre-spliceosome formation

To establish the role of Rse1p in splicing, we prepared extracts depleted of Rse1p. For this purpose, a splicing extract containing the ProtA–Rse1p fusion was incubated with IgG–agarose beads. To control for non-specific depletion and inactivation of the extract during the depletion procedure, the same extract was incubated in parallel with glutathione–agarose beads and a wild-type extract was incubated with IgG–agarose beads. Proteins and RNAs remaining in the extracts after the depletion procedure were analysed by Western blotting and primer extension, respectively (Figure 5A; data not shown). The level of ProtA–Rse1p was not affected by the mock depletion on

glutathione–agarose beads (Figure 5A, lanes 3 and 5) while far more than 99% of the protein was removed by incubation with IgG–agarose beads (Figure 5A, lanes 3 and 4; protein levels were determined by comparing the signal of the depleted extract with signals from serial dilutions of the complete extract; data not shown). ProtA–Rse1p removal was relatively specific since the levels of Lea1p and U2 snRNA in these extracts were only weakly reduced (Figure 5A, bottom; data not shown). These results indicated that U2 snRNP was not co-depleted quantitatively with ProtA–Rse1p. Splicing complex formation was assayed in these various extracts. Pre-/spliceosomes assembled efficiently in both the untreated wild-type and ProtA–RSE1 extracts (Figure 5B, lanes 1 and 3). CC2 and to a lesser extent also CC1 accumulated in the ProtA–Rse1p-depleted extract, whereas no pre-/spliceosomes could be detected (Figure 5B, lane 4). This effect was dependent on ProtA–Rse1p depletion as it was not observed either when the wild-type extract had been incubated with IgG–agarose beads (Figure 5B, lane 2) or when the ProtA–Rse1p extract was mock depleted with glutathione–agarose beads (Figure 5B, lane 5). This suggested a role for Rse1p in pre-spliceosome formation.

To confirm the function of Rse1p in pre-/spliceosome
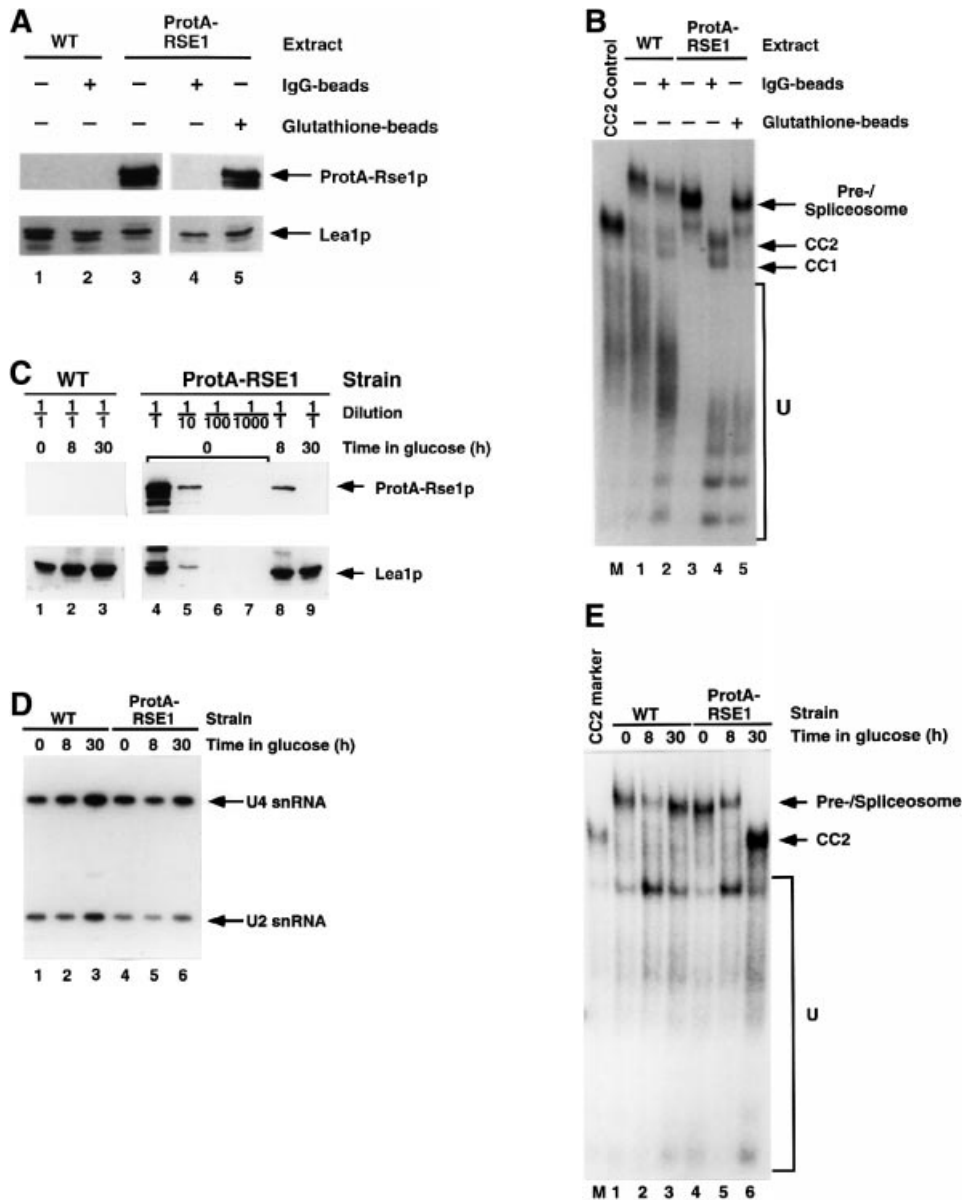
Fig. 5. Rse1p is required for pre-spliceosome formation. (**A**) *In vitro* depletion of Rse1p. ProtA–Rse1p and Lea1p levels were analysed by Western blotting in wild-type and ProtA–Rse1p-containing extracts before (lanes 1 and 3) and after treatment with either IgG– or glutathione–agarose beads (lanes 2, 4 and 5). Proteins were detected using either IgG or Lea1p-specific antibodies. (**B**) Extracts treated as described in (A) were tested for splicing complex formation. Complexes were assembled on a radioactively labelled wild-type pre-mRNA and fractionated by native gel electrophoresis. Lanes are equivalent to the lanes in (A). The complexes in lanes 3 and 5 running just above CC2 do not correspond to a characterized splicing complex. In lane M, ATP was depleted by pre-incubation of the wild-type splicing extract with glucose (CC2 marker). Complexes are labelled as in Figure 3; CC2 indicates the commitment complex 2. (**C**) Western blot analysis of ProtA–Rse1p and Lea1p levels during genetic depletion. An extract derived from the ProtA–RSE1 strain grown in galactose medium (initial conditions) was diluted 10-, 100- and 1000-fold (lanes 4–7) to compare with the ProtA–Rse1p levels after 8 and 30 h. Signals corresponding to ProtA–Rse1p and Lea1p are indicated on the right. (**D**) Primer extension analysis of U2 snRNA levels during genetic depletion of ProtA–Rse1p. Total RNA was extracted from strains at the same time points as in Figure 4B and levels of U4 and U2 snRNAs were monitored. (**E**) Analysis of splicing complex formation during genetic depletion of ProtA–Rse1p. Extracts prepared at the indicated time points (Figure 4B) were incubated with radioactively labelled wild-type pre-mRNA and fractionated by native gel electrophoresis. CC2 marker was prepared as described in (B). Complexes are labelled as in (B).

formation, we analysed splicing complex formation following *in vivo* depletion of Rse1p. For this purpose, extracts were prepared from cells expressing ProtA–Rse1p at several time points after transfer to glucose-containing medium (see Figure 4B above). The protein and RNA contents of these extracts were determined by Western blot and primer extension analysis. The level of ProtA–Rse1p was decreased ~10-fold following a 8 h incubation in glucose medium and >100-fold following 30 h incuba-

tion in glucose medium (Figure 5C, top row). This depletion did not affect the level of Lea1p (Figure 5C, bottom row) and U2 snRNA (Figure 5D; compare lanes 4–6). Pre-/spliceosomes formed efficiently in extracts prepared from a wild-type strain grown in galactose or glucose (Figure 5E, lanes 1–3), the ProtA–RSE1 strain grown in galactose (Figure 5E, lane 4) and the ProtA–RSE1 strain grown for 8 h in glucose medium (Figure 5E, lane 5). However, in Rse1p-depleted extracts, obtained

from cells grown for 30 h in glucose medium, pre-spliceosome formation was blocked and CC2 accumulated (Figure 5E, lane 6).

The results of these *in vitro* and *in vivo* depletion experiments strongly suggest that the splicing defect in the absence of Rse1p is caused by a block in the spliceosome assembly pathway at the transition step from CC2 to pre-spliceosome. This conclusion was supported further by our observation that the partially functional *RSE1*–TAP allele also hindered pre-spliceosome formation (data not shown). Depletion of U2 snRNA and of Lea1p have been reported to prevent pre-spliceosome formation (Séraphin and Rosbash, 1989a; Caspary and Séraphin, 1998). However, the block observed in Rse1p-depleted extracts cannot be attributed to decreased U2 snRNA or Lea1p levels, as these factors are not or only slightly affected by the Rse1p depletion (Figure 5A, C and D). At this stage, we cannot rule out that other factors essential for pre-spliceosome formation are either co-depleted with Rse1p or destabilized in the absence of Rse1p and responsible for the observed phenotype. However, this is unlikely since the same defect in pre-spliceosome assembly was detected following *in vivo* and *in vitro* depletion (Figure 5B and C) and through the analysis of the conditional *RSE1*–TAP allele (data not shown). Furthermore, Rse1p is a component of the pre-spliceosome (Figure 3). Therefore, it is not unexpected that the formation of this complex could not proceed in the absence of Rse1p.

Rse1p is conserved in eukaryotes and is similar to polyadenylation factors and DNA repair proteins

To determine whether Rse1p is a conserved splicing factor, we searched protein and nucleic acid sequence databases using the amino acid sequence of Rse1p as a probe. This revealed several proteins or putative translation products with significant similarity to Rse1p. Among them, two proteins of unknown function from *C.elegans* and human cells displayed 24 and 25% sequence identity, respectively, with the yeast Rse1p at the amino acid level. These two proteins were 61% identical to each other. The three proteins were similar over their entire length even though identities were more concentrated in the C-terminal region. Ten other proteins also showed significant similarity to Rse1p. However, the level of identity was much lower. Multiple sequence alignment revealed that sequence similarity was concentrated in several blocks, which in some cases contain absolutely conserved residues (a subset of these blocks is depicted in Figure 6). Interestingly, all these proteins (except for the *Dictyostelium discoideum* sequence which is probably a partial sequence) have a closely located translation initiation codon (Figure 6). The proposed translation initiation codon of the YML049C ORF was located >55 amino acids upstream of the translation initiation codon of the Rse1p-related proteins. Since a Rse1 protein lacking the first 50 residues encoded by the YML049C ORF is fully functional (see above), this suggests that the corresponding translation initiation site had been mis-assigned. Consistent with this possibility, the construct carrying the 50 extra amino acids between ProtA and Rse1p was associated with U2 snRNA and pre-/spliceosome but, unlike ProtA–Rse1p, its depletion did not block growth and led only to a partial block of pre-/spliceosome formation (data not shown). This supports the idea that the methionine at position 51 is the natural translation initiation site.

Among the proteins related to Rse1p is the 127 kDa subunit of the human UV-damaged DNA-binding protein (UV-DDB) which is absent in cells derived from individuals with xeroderma pigmentosum group E DNA repair defect. Two functional UV-DDB homologues from monkey (*Cercopithecus aethiops*) and *C.elegans* are also similar to Rse1p. Other related proteins have been shown to function as the 160 kDa subunit of the cleavage–polyadenylation-specific factor (CPSF) which is involved in 3′ end processing of pre-mRNA. Sequence similarities between these two protein subfamilies had already been noticed (Jenny and Keller, 1995), but the overall similarity to the Rse1p splicing factor had not been reported. A neighbour-joining tree reflecting sequence similarities demonstrated that these proteins clearly fall into three subfamilies (data not shown). One subfamily contains the yeast Rse1p and its putative human and *C.elegans* homologues. Another subfamily contains the human, monkey and *D.discoideum* UV-DDB (Takao *et al.*, 1993; Dualan *et al.*, 1995; Alexander *et al.*, 1996) and related proteins from *C.elegans*, *Arabidopsis thaliana* and *Schizosaccharomyces pombe*. Finally, the third subfamily contains known CPSF subunits (*Homo sapiens*, *Bos taurus*, *Saccharomyces cerevisiae*) and a putative homologue from *S.pombe* (Keller *et al.*, 1991; Jenny and Keller, 1995; Stumpf and Domdey, 1996). The tree strongly suggests that the human and *C.elegans* proteins which are most highly related to Rse1p are functional homologues and, therefore, could be also involved in pre-mRNA splicing. Furthermore, this shows that proteins involved in processes as different as DNA repair, polyadenylation and pre-mRNA splicing may have a common origin. This suggests that these proteins might have related functions that have not yet been discovered (see Discussion).

Discussion

Partial purification of the yeast U2 snRNP followed by identification of proteins by mass spectrometry revealed a novel factor named Rse1p. The essential Rse1p interacts at low salt concentrations with U2 snRNA and is a component of the pre-spliceosome. Rse1p is required for splicing *in vivo* and appears to be required at the transition step from CC2 to pre-spliceosome during splicing complex formation. Rse1p has putative homologues in human and *C.elegans*, suggesting that Rse1p is a widely distributed general splicing factor. Rse1p also belongs to a new protein family containing DNA repair proteins and polyadenylation factors.

Rse1p is a new U2 snRNP-associated splicing factor with properties similar to those of the human SF3 factor

We have partially purified the yeast U2 snRNP. Nano-electrospray mass spectrometry analysis revealed the presence of three U2 snRNP proteins (Lea1p, Yib9p and SmD2) in the purified fractions. We also identified a new factor for which the name Rse1p had been assigned. Using three different tagged versions of Rse1p (ProtA–Rse1p, Rse1p–TAP and the construct carrying the 50 extra amino acids between ProtA and Rse1p), we demonstrate that it

```
UV-C.aethiops       -----------MSYNMVVTAQKPTAVNGCVTAHFT        NLLIAKNTRLEIYVV            LRPVKEVG-MYGKIAVMELFR
UV-H.sapiens        -----------MSYNMVVTAQKPTAVNGCVTGHFT        NLLIAKNTRLEIYVV            LRPVKEVG-MYGKIAVMELFR
UV-A.thaliana       ----------MSVWNMAVTAQKPTCVTHSCVGNFT        NLIVAKSTRIEIHLL           LQTILDVP-LYGRIATMELFR
UV-D.discoideum     -----------------------------------       ---------------M          LKPMFDVN-IYGRISVLKLFS
UV-C.elegans        -----------MPISVCVSAKKASVVVESVVGNFT        NLIVARGNRIDVQLV           LKNVCEIP-IYGQVLTIALVK
UV-S.pombe          -----------MTYVVTYLHKPSSIRNAVFCKFV         NVIVAKVNCLEVLSY           LCLITSAN-IFAKIVNVKAFK
RSE1-H.sapiens      -----------MFLYNLTLQRATGISFAIHGNFS   (0-11) EIVVSRGKILELLRP   (4-48)  VHTLLTVE-VFGVIRSLMAFR
RSE1-C.elegans      -----------MHLVNLTLQGQSAINQAIQGNFS         EIVWRGGSALELLTL           IKVMCHQD-IFGIVRSLLAFR
RSE1-S.cerevisiae   (50-) MVAKDDELYLVHLTLKKQTNFVHSCIGHFV       QLCVATETHLELMDT           LKLIAKFQNLFATITSMKSLD
CPSF-H.sapiens      -----------MYAVYKQAHPPTGLEFSMYCNFF         NLVVAGTSQLYVVYRL          LELVASFS-FFGNVMSMASVQ
CPSF-B.taurus       -----------MYAVYKQAHPPTGLEFSMYCNFF         NLVVAGTSQLYVVYRL          LELVASFS-FFGNVMSMASVQ
CPSF-S.pombe        -----------MSTIFQDLVDSTVIKNAVQGQFT         NLVVSKVNSLHLFEI           LRLVSQVK-VFGTITEISALK
CPSF-S.cerevisiae   -----------MNVYDDVLDATVVSHSLATHFT          ELLVVRTNILSVYRP           LYLTDEFK-FHGLITDIGLIP
consensus                                                     . . . . .   . .           . . . . . .   . . .
```

```
VIDVKFLY-GCQAPTICFVY         VVAMETFTNLGPIVDMCVVDL         QLVTCSGAFKEGSLRIIRNGIGIHEHASIDLPG-IKGLWPL
VIDVKFLY-GCQAPTICFVY         VVAMETFTNLGPIVDMCVVDL         QLVTCSGAFKEGSLRIIRNGIGIHEHASIDLPG-IKGLWPL
VLDIKFLY-GCTKPTIAVLY         VEILEKYVNLGPIVDFCVVDL         QVVTCSGAYKDGSLRIVRNGIGINEQASVELQG-IKGMWSL
VLDMTFLY-GCKVPTIAVLF         VTYLEAFTNIGPVVDFCVVDA         QIVTCSGTYRDGSLRIIRNGIGTABQASIELEG-IKGIFPI
VSDFKFVDTGEDVYRVAFIY   (85-  SVILETYSNIGPIRDMVMVES         QLVTCTGADKDGSLRVIRNGIGIDELASVDLAG-VVGIFPI
VVDIAMLY-NSSRPSIAVLY    146) LVLVDELDSLSPILFCQIADL  (1-18)  SIITCSCAYKDGTLRIIRNSINIENVALIEMEG-IKDFFSV
GVDVGFEN-----PMFACLE         LIPIDSMDSLSPLTDAVIGDI         QIYSLVCRGARSSLRVLRHGLEVSEMAVSELPGNPNAVWTV
GIDVGFEN-----PTFACLE         LSILSQQLNLNPSIKSQIVSD         QTYSLVCGRGARSSLRVLRNGLEISEMAVSDLPGNPNAVWTV
VVDVNFNN-----PCFVTLE         FEVCDSILNIGPCANAAMGEP         PLSIATKHFTNNKIITLTNAVNYSNLISTSLPPNATKLWLI
IVDLQFLH-GYYEPTLLILF         FEVCDSILNIGPCANAAMGEP         EIVVCSGYGKNGALSVLQKSIRPQVVTTFELPG-CYDMWTV
IVDLQFLH-GYYEPTLLILF         LEICDVLTNIGPITDFAVGKA         EIVVCSGYGKNGALSVLQKSIRPQVVTTFELPG-CYDMWTV
ILDVKFLY-CYREPTLAILY         LELLSSLRNVGPITSLTVGKV         ELVGTACADGAGGLVVPRRNIFPLIAGEFQFDG-CEALWTV
IIDIQFLK-NFTKPTIALLY                                       SLVATSGNGSGSHLTVIQTSVQPEIELALKFIS-ITQIWNL
   . . *  .   . . . . . .                . . . . *  .       . . . . . .   . . . .   . . . .   . . . .
```

```
TLVLSFVGQTRVLML         LIQITSASVRLV         MVEFNGKLLASINSTVRLYEW         FILVGDLMRSVLLLAY
TLVLSFVGQTRVLML         LIQITSASVRLV         MVEFNGKLLASINSTVRLYEW         FILVGDLMRSVLLLAY
FLVVSFISETRILAM         LVQVTSNSVRLV         LNAFNGKLLASINQKIQLYKW         FIAVGDLMKSISLLIY
YLITSFIECTKVLSF         LIQITNVSINLI         LLSFNGRLTAAVHKRLFSIRY         FILVGDMMKSMSLLVE
YVIVSLSDETHVLQI         ILQATEKQIRLM         IRILNGRLVAAINSSIRLFEW         EVAVADVMRSVSLLSY
YIFLSLICETRAIIV         ILQITTKEIRLF         LVLYKHLIVAGINASVCIFEY         ELIAADLMKSITVLQF
YIIVSFVNATLVLSI   (28-  LVQVYPDGIRHI  (408-  IAPFQGRVLIGVGKILLYDL   (25-   RVIVSDVQESFIWVRY
YIVVSFVNATLALTI    35)  LVQIYSEGIRHI   545)  IHDFRGMAIVGFGRFLRMYDI    31)  RHIVSDSQESVHFLRY
LLFITFPKKTMILQI         IIQVCTAEIRHI         MLKFKNFLLTAMGSTIVLYGL         RLAVGDIHESVTLFIW
FLILSREDSTMILQT         IVQVSPLGIRLL         LCHCNGHLVSAIGQKIFLWSL         FILAADVMKSISLLRY
FLILSREDSTMILQT         IVQVSPLGIRLL         LCHCNGHLVSAIGQKIFLWSL         FILAADVMKSISLLRY
YLVLSKEKESFIFLA         MVQICPTSLRVY         VCEVDGYLLSGQGQKVIVRAL         LLLFGDVRQNVTFVGF
ITTDSTKSRSDIYES         IIQVLTNHLYLY         VCEVSGRFMISQSQKVLVRDI         LLIIGDAMQGFQFIGF
   . . . . . .   .           . . . . . *  .      . . . . . .            .   . . . . . *  .   .
```

```
FLGAENAFNLFVC         VLFGTVNGMIGLVTSL         PATGFIDGDLIESFLDIS         (1140)
FLGAENAFNLFVC         VLFGTVNGMIGLVTSL         PATGFIDGDLIESFLDIS         (1140)
YLGTDNCFNIFTV         VIFGTVSCMIGVIASL         EAKGYLDGDLIESFLDLS         (1102)
FIGAEASNNFIVV         ILYASVNCSIGVVASI         DSKNFIDGDLIETFLDLK         (1139)
ILGGEAHLNLFTV   (26-  IMFGINQCTIGMIVQI   (48-  PPSGFVDGDLVESIDDMD         (1129)
YFVTEADGNAVIL    32)  LLCATVDCSLMIVGDA    74)  P-SDLIDGSLIESILGLR  (34-   (1072)
VAGADKFGNICVV         LVYTTLSCGIGILVPF         PVKNVIDGDLCEQRNSME   81)   (1217)
VAVADKFGNLAVV         LVYTTIGCAIGCLVSF         PCKSVIDGDICEQFSLMD         (1220)
VIGADRYGNAWTL         CIYMGLQCTVGCFIPL         PVRKVIDGDLCENFLRLS         (1361)
FLVSDRDRNLMVY         TWFATLDCGIGLLLPM         AVRNVLDGELLNRYLYLS   (29-   (1444)
FLVSDRDRNLMVY         TWFATLDCGIGLLLPM         AVRNVLDGELLNRYLYLS    39)   (1444)
FVVADTSGNLRLL         VMVNS-DCGLQMLVPI         PTRRILDGMLTDYFTYMS         (1441)
FAATDADRNVHVL         VGGQV-DCSVFKIVPL         SMRPMLDFNVLRRFCGLA         (1357)
   . . . . .   .           . . . . . *  . . . .       . . . *  . . . . .   . .
```

Fig. 6. Sequence alignment of a subset of conserved regions among the members of a new protein family containing yeast Rse1p. UV = UV-DDB protein homologues. Hyphens indicate gaps. The numbers between the alignment blocks indicate the minimal and maximal number of removed residues. The specific N-terminal 50 amino acid residues of the Rse1p from *S.cerevisiae* are indicated by '50–'. Identical amino acids are highlighted with dark grey, and light grey illustrates conservation of physicochemical properties at that position. Absolutely conserved residues are indicated by a star below the alignment. The numbers at the end indicate the total number of residues of each protein.

is associated with U2 snRNP, albeit only at low salt concentrations (<300 mM). Furthermore, only a fraction of the U2 snRNA (≤10%) is associated with Rse1p. In contrast, the U2 snRNP-specific protein Lea1p co-precipitates ~50% of U2 snRNA at the same salt concentrations (Caspary and Séraphin, 1998; data not shown). This suggests that Rse1p interacts only weakly or transiently with U2 snRNA. In contrast, pre-spliceosomes seem to be associated quantitatively with Rse1p, and depletion experiments suggest that Rse1p is required for pre-spliceosome assembly. Consistent with these results, Rse1p is required for efficient pre-mRNA splicing *in vivo*.

Several features of Rse1p appear to be very similar to SF3 subunits of the human splicing machinery. The human SF3 splicing factor is associated loosely with the U2 snRNA. Its components are found in the salt-sensitive 17S U2 snRNP but not in the more stable 12S U2 snRNP (Behrens *et al.*, 1993b; Brosi *et al.*, 1993a). Similarly, association of Rse1p with U2 snRNP is lost when the salt concentration is increased to 300 mM. Like Rse1p, the human SF3 proteins were purified and identified as essential factors required for pre-spliceosome formation (Krämer, 1988; Krämer and Utans, 1991; Brosi *et al.*, 1993b) and have been shown to be part of the pre-spliceosome (Bennett *et al.*, 1992; Staknis and Reed, 1994). Biochemical fractionation of the SF3 factor indicated that it consists of seven subunits associated in two subcomplexes named SF3a and SF3b (Brosi *et al.*, 1993b; Krämer, 1996). The three subunits of the SF3a complex (SAP61/SF3a[60], SAP62/SF3a[66] and SAP112/SF3a[120]) were found to be homologous to the three yeast splicing factors Prp9, Prp11 and Prp21, respectively (Behrens *et al.*, 1993a; Bennett and Reed, 1993; Brosi *et al.*, 1993a; Chiara *et al.*, 1994; Krämer *et al.*, 1995). The human and yeast proteins are not only similar in sequence but form related complexes involving conserved protein–protein interactions (Legrain

and Chapon, 1993; Ruby *et al.*, 1993; Rain *et al.*, 1996; reviewed by Hodges and Beggs, 1994; Krämer, 1996). Furthermore, in both species, this complex has been proposed to interact with the 5′ end of the U2 snRNA and shown to be required for pre-spliceosome formation (reviewed by Krämer, 1996). Only three of the four subunits of the SF3b splicing factor have been characterized, namely SAP49/SF3b[50], SAP145/SF3B[145] and SAP155/SF3b[155] (Champion-Arnaud and Reed, 1994; Gozani *et al.*, 1996; Schmidt-Zachmann *et al.*, 1998; Wang *et al.*, 1998). Homologues of these proteins have been identified in yeast (HSH49, CUS1 and ySAP155; Gozani *et al.*, 1996; Wells *et al.*, 1996; Igel *et al.*, 1998; Wang *et al.*, 1998). Again, conservation goes beyond sequence similarities. SAP145 and Cus1p are required for pre-spliceosome formation (Gozani *et al.*, 1996; Wells *et al.*, 1996) and they interact with homologous proteins, SAP49 and HSH49, respectively (Champion-Arnaud and Reed, 1994; Igel *et al.*, 1998). Functional conservation of SAP155 and ySAP155 has not been established yet, but is very likely given that with 50% identity it currently is the most conserved SF3b subunit at the amino acid level (Wang *et al.*, 1998). All SF3a components as well as the SAP49, SAP145 and SAP155 SF3b subunits have been shown to cross-link with the pre-mRNA close to the branch point (Gozani *et al.*, 1996). The only uncharacterized subunit of the SF3b splicing complex has an approximate mol. wt of 130 kDa (Krämer, 1996) and an acidic pI (Bennett *et al.*, 1992; Champion-Arnaud and Reed, 1994). Interestingly, the putative human Rse1p homologue that we identified has a predicted mol. wt of 135.6 kDa and a calculated pI of 5.13. This strongly suggested that Rse1p could be the yeast homologue of SAP130/SF3b[130]. This hypothesis has now been confirmed (A.Krämer, personal communication). Further support for the presence of Rse1p in the yeast SF3-like complex comes from saturating two-hybrid screens in which Rse1p was found among a dozen proteins interacting with Prp9 (Fromont-Racine *et al.*, 1997). Although the significance of this interaction was not clear (Fromont-Racine *et al.*, 1997), it suggests that Rse1p and its human homologue could be located at the interface between the two SF3 subcomplexes.

Rse1p belongs to a new protein family containing polyadenylation factors and UV repair proteins

Database searches revealed that Rse1p is related to several proteins. Even though the overall level of similarity between these various proteins was low, multiple sequence alignment indicated the presence of conserved blocks of sequences with significant similarities. Nevertheless, the various related proteins are quite large (120–160 kDa) and only very few residues are absolutely conserved (Figure 6). While the sequence similarity is distributed over the complete length of these proteins, it is significantly stronger close to the C-terminus. Consistent with this observation, the introduction of a tag at the C-terminus of Rse1p generated a partially functional allele displaying a thermosensitve phenotype.

An unrooted neighbour-joining tree, built from the multiple sequence alignment, revealed that this protein family is composed of three distinct subfamilies. One of the subfamilies contains a subunit of a dimer-binding UV-DDB (Takao *et al.*, 1993; Dualan *et al.*, 1995;

Alexander *et al.*, 1996). This subunit has been shown to interact with the damaged DNA (Reardon *et al.*, 1993). A leucine zipper motif and an adjacent basic region serving as a putative DNA-binding site (Figure 6, between block 8 and 9) were proposed (Alexander *et al.*, 1996). This region is highly conserved between all members of this subfamily, but this high degree of conservation is not extended to the other two subfamilies. The second subfamily contains the human CPSF160 subunit of the human polyadenylation machinery and related homologues from other species. In most cases, these homologues have also been shown to be implicated functionally in polyadenylation (Jenny and Keller, 1995; Stumpf and Domdey, 1996). The CPSF complex contains three or four subunits and is implicated in specific recognition of the AAUAAA polyadenylation signal (Murthy and Manley, 1995). CPSF160 has been shown to cross-link to the pre-mRNA substrate (Keller *et al.*, 1991) and, therefore, has been proposed to be involved in RNA recognition. The presence of a degenerate RNA recognition motif (RNP; Burd and Dreyfuss, 1994) and of a bipartite nuclear localization signal (Jenny and Keller, 1995) in this protein has been suggested. However, the corresponding sequences are not conserved in all members of this subfamily (Figure 6; data not shown). The third subfamily contains Rse1p and related proteins from *C.elegans* and human. Interestingly, the two first branches of the tree contain only proteins that have been shown to have identical functions in different species or putative homologues that have not been characterized functionally yet. It is likely, therefore, that the human and *C.elegans* proteins, which together with the yeast Rse1p belong to the third branch, are also functional homologues in these species.

The observation of sequence conservation between these various proteins suggests that they may have related functions. However, given that one group of proteins is involved in DNA repair while the other groups are involved in RNA metabolism, it is unclear what this function could be. On the one hand, these various proteins could be conserved because they interact with nucleic acids. This is indeed the case for UV-DDB and CPSF160. In contrast, SAP130/SF3b[130], the putative human Rse1p homologue, could not be cross-linked to the pre-mRNA (Gozani *et al.*, 1996). However, it remains possible that it interacts with another RNA (e.g. U2 snRNA). On the other hand, all characterized proteins of the family have been shown to be present in multimeric complexes. Similarly, Rse1p is likely to interact with other protein splicing factors (see above). The sequence conservation might therefore reflect the conservation of protein–protein interaction interface(s). However, the proteins that have been found to interact with UV-DDB and CPSF160 and the proteins of the SF3 complex have not been reported to share any obvious similarity. It is, nevertheless, possible that proteins of the Rse1p family all interact with other proteins commonly involved in nucleic acid metabolism such as DNA/RNA helicases. Further work will be required to establish the significance of the sequence similarity observed between the various members of the Rse1 protein family.

Materials and methods

Strains
For the C-terminal tagging of Rse1p, Lea1p and SmB and the disruption of the *RSE1* gene, the haploid wild-type strain MGD453-13D (Séraphin

et al., 1988) and the isogenic diploid strain BSY320 were used. For the N-terminal tagging of Rse1p, the *GAL*-deficient strain YDL401 (Lafontaine and Tollervey, 1996) was used. All yeast transformations were performed as described previously (Soni and Carmichael, 1993).

To tag Lea1p and Rse1p at the C-terminus, a construct containing TAP (G.Rigaut, A.Shevchenko, B.Rutz, M.Wilm, M.Mann and B.Séraphin, submitted) together with the *K.lactis TRP1* marker (E.Bragado-Nilsson and B.Séraphin, unpublished) was inserted in the genome downstream of, and in-frame with, the *LEA1* and *RSE1* ORFs by transformation with PCR fragments (Puig *et al.*, 1998). For *LEA1*, this fragment was generated by amplification from plasmid pBS1479 with oligonucleotides FC9-1 (5′-TTAGAAGAGATTGCCAGGCTGGAAAAACTACTCTCTGGTG-GTGTTAAGAGAAGATGGAAAAAGAA-3′) and FC1-2 (described in Caspary and Séraphin, 1998). The PCR product was transformed into the strain MGD453-13D, generating strain BSY678. For *RSE1*, the oligonucleotides FC15 (5′-AGTGGAGGATATAATTCAAACAAT-CAACGAAGTCAGAACAAATTACATGCCATGGAAAAGAGAAGA-3′) and FC14-2 (5′-AGAAAATACATTAACACGCCTTTATACTT-GGTATTATTACTGTAAAATCATACGACTCACTATAGGG-3′) were used, generating strain BSY737.

BSY677, the strain containing C-terminal TAP-A (G.Rigaut, A.Shevchenko, B.Rutz, M.Wilm, M.Mann and B.Séraphin, submitted) tagged SmB was constructed as described above using plasmid pBS1470 (*KL-URA3* marker) and the oligonucleotides FC8-1 (5′-CCTCAAACA-AGGAAGTTTCAGCCCCCACCAGGTTTTAAAAGAAAAGAGAA-TTTGTATTTTCAGG-3′) and JS61 (5′-AAGTATACGGAAAC-TATATTAGACTACACTACATCAACCTACGACTCACTATAGGG-3′). BSY677 was used additionally to tag Lea1p with TAP-C at its C-terminus (G.Rigaut A.Shevchenko, B.Rutz, M.Wilm, M.Mann and B.Séraphin, submitted) using the plasmid pBS1512 (*KL-TRP1*) and the oligonucleotides FC9-1 and FC1-2, resulting in the new strain BSY723.

N-terminal tagging of Rse1p was according to Lafontaine and Tollervey (1996). The oligonucleotides used were FC18-1 (5′-GGCACCATTA-GCTTCTTTTCTCGCTTTTTTATAATATTTTCCATCATTAACTCTT-GGCCTCCTCTAGT-3′) and FC18-2 (5′-GCCGTGTGGGGCGACAA-GCTTACTACTGCCATTTTGCCACCTCCCCACATATTCGCGTCTA-CTTTCGG-3′) for the *YML049C* ORF, resulting in BSY749, and oligonucleotides FC18-1 and FC39-2 (5′-GTTTCTTTAAAGTAA-GATGGTACAAGTATAGCTCATCATCTTTAGCAACCATATTCGCG-TCTACTTTCGG-3′) for ProtA–Rse1p, resulting in BSY786.

RSE1 was disrupted according to Puig *et al.* (1998). The fragment was generated by amplification from plasmid pBS1365 with oligonucleotides FC14-1 (5′-GGCACCATTAGCTTCTTTTCTCGCTTTTTTATAATAT-TTTCCATCATTAAAAGCTGGAGCTCAAAAC-3′) and FC14-2 (5′-AGAAAATACATTAACACGCCTTTATACTTGGTATTATTACTG-GTAAAATCATACGACTCACTATAGGG-3′). This PCR fragment was transformed into the diploid strain BSY320, resulting in BSY736. Disruption was confirmed by PCR using the primer FC1-2 annealing inside *KL-URA3* (Caspary and Séraphin, 1998) and a second primer FC17 (5′-TCTGCTGCCCCTCTCTA-3′) annealing upstream of *RSE1*. One disruptant was sporulated and dissected.

U2 snRNP purification and protein characterization

Two-step native affinity purification was performed according to G.Rigaut, A.Shevchenko, B.Rutz, M.Wilm, M.Mann and B.Séraphin (submitted) with the strains BSY723 and BSY737. Purification was started from 5 l of cell culture grown to OD = 3.0 for BSY723 and to OD = 2.7 for BSY737. Purified proteins were concentrated by TCA precipitation and separated on a 7–25% SDS–polyacrylamide gradient gel. Proteins were silver stained and in-gel digested as described in Shevchenko *et al.* (1996).

The dried down peptide mixture was taken up in 10% formic acid, desalted over a self-assembled 100 nl Poros R2 column (PerSeptive Biosystems, Cambridge, MA) and eluted directly into a nanoelectrospray capillary using 1 µl of a 5% formic acid, 50% methanol solution. Peptides were analysed on a triple quadrupole mass spectrometer (API III from PE-Sciex, Ontario, Canada). Peptides from the purified protein were identified in the spectrum by careful comparison with a spectrum from a blank, and their fragment spectra were acquired. Proteins were identified in the non-redundant sequence database nrdb obtained from the EMBL-EBI institute (Hinxton, UK) using the sequence-tag algorithm and the PeptideSearch program (Mann, 1996).

Immunoprecipitation and primer extension

Extract preparation from the strains described above, immunoprecipitation and primer extension were as described previously (Séraphin, 1995).

Native gel mobility shift assay

Yeast splicing extracts were prepared as described previously (Séraphin and Rosbash, 1989a). Pre-mRNA was generated by *in vitro* transcription of the *RP51A* intron derivative contained in plasmid pBS195 (Séraphin and Rosbash, 1991) after digestion with *Dde*I. Native gels were according to Séraphin and Rosbash (1989a) but using 3% (37.5:1) acrylamide mix (Protogel; National Diagnostic). To deplete ATP, the reactions were incubated for 10 min at room temperature with 2 mM glucose prior to addition of labelled pre-mRNA (Liao *et al.*, 1992). For the shift of protein A-containing complexes, the splicing extracts were incubated for an additional 10 min at room temperature with 1 µg of IgG antibody.

In vitro depletion and Western blotting

Extracts were made according to Newman (1994). A 200 µl aliquot of extract was adjusted to 100 mM KCl and 0.1% NP-40 prior to incubation with 110 µl of either pre-washed (100 mM KCl, 0.1% NP-40, 10 mM Tris–HCl, pH 8.0) IgG–Sepharose or glutathione–Sepharose beads (Pharmacia) for 2 h at 4°C. Supernatants were used for further analysis. For Western blot analysis, proteins were resolved on 10% SDS–polyacrylamide gels and detected using the peroxidase–anti-peroxidase complex (Sigma) and the ECL detection system (Amersham). Antibodies were raised against recombinant Lea1p and Yib9p (Caspary and Séraphin, 1998).

RNA analysis and genetic depletion

Total RNA extraction and primer extension derived from the strains MGD453-13D and BSY678 were as described previously (Caspary and Séraphin, 1998). For genetic depletion of Rse1p, the strain BSY786 was incubated in YPD medium containing an additional 2% glucose. The OD_{600} of the culture was measured and the concentration of the cells in the culture was adjusted to remain below $OD_{600} = 2.0$. Primer extension of total RNA was performed as described previously (Pikielny and Rosbash, 1985) using a labelled oligonucleotide complementary to the second exon of the *U3* gene (Zavanelli and Ares, 1991).

Sequence alignment

For the database searches, WU-BLAST2 (W.Gish, unpublished) and the GCG package (Devereux *et al.*, 1984) were used. The multiple sequence alignment was done with Clustal_X (Thompson *et al.*, 1997) and presented with BOXSHADE 3.21 software. The neighbour-joining tree was computed, excluding positions with gaps and correcting for multiple substitutions. The SWISSPROT and SPTREMBL accession numbers of the sequences are as follows: RSE1: *H.sapiens* (D1033627), *C.elegans* (O44985) and *S.cerevisiae* (Q04693); CPSF: *H.sapiens* (Q10570), *B.taurus* (Q10569), *S.cerevisiae* (Q06632) and *S.pombe* (E1331571); UV-DDB: *H.sapiens* (Q16531), *Cercopithecus aethiops* (P33194), *C.elegans* (Q21554), *A.thaliana* (O49552), *D.discoideum* (Q23865) and *S.pombe* (O13807)

Note, however, that the predicted exon–intron structure of the UV-DBB gene from *C.elegans* was modified to increase the similarity of the encoded protein to its human and plant homologues. The sequence presented in Figure 6 therefore differs slightly from the original entry.

Acknowledgements

References

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

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.

Alexander,H., Lee,S.K., Yu,S.L. and Alexander,S. (1996) *repE*—the *Dictyostelium* homolog of the human xeroderma pigmentosum group E gene is developmentally regulated and contains a leucine zipper motif. *Nucleic Acids Res.*, **24**, 2295–2301.

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

Behrens,S.E., Tyc,K., Kastner,B., Reichelt,J. and Lührmann,R. (1993b) Small nuclear ribonucleoprotein (RNP) U2 contains numerous additional proteins and has a bipartite RNP structure under splicing conditions. *Mol. Cell. Biol.*, **13**, 307–319.

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

Bennett,M., Michaud,S., Kingston,J. and Reed,R. (1992) Protein components specifically associated with prespliceosome and spliceosome complexes. *Genes Dev.*, **6**, 1986–2000.

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

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

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

Burd,C.G. and Dreyfuss,G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.

Caspary,F. and Séraphin,B. (1998) The yeast U2A′/U2B″ complex is required for pre-spliceosome formation. *EMBO J.*, **17**, 6348–6358.

Champion-Arnaud,P. and Reed,R. (1994) The prespliceosome components SAP 49 and SAP 145 interact in a complex implicated in tethering U2 snRNP to the branch site. *Genes Dev.*, **16**, 1974–1983.

Chen,E.J., Frand,A.R., Chitouras,E. and Kaiser,C.A. (1998) A link between secretion and pre-mRNA processing defects in *Saccharomyces cerevisiae* and the identification of a novel splicing gene, *RSE1*. *Mol. Cell. Biol.*, **18**, 7139–7146.

Chiara,M.D., Champion Arnaud,P., Buvoli,M., Nadal Ginard,B. and Reed,R. (1994) Specific protein–protein interactions between the essential mammalian spliceosome-associated proteins SAP 61 and SAP 114. *Proc. Natl Acad. Sci. USA*, **91**, 6403–6407.

Devereux,J., Haeberli,P. and Smithies,O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.*, **12**, 387–395.

Dualan,R., Brody,T., Keeney,S., Nichols,A.F., Admon,A. and Linn,S. (1995) Chromosomal localization and cDNA cloning of the genes (*DDB1* and *DDB2*) for the p127 and p48 subunits of a human damage-specific DNA binding protein. *Genomics*, **29**, 62–69.

Fabrizio,P., McPheeters,D.S. and Abelson,J. (1989) *In vitro* assembly of yeast U6 snRNP: a functional assay. *Genes Dev.*, **3**, 2137–2150.

Fromont-Racine,M., Rain,J.-C. and Legrain,P. (1997) Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nature Genet.*, **16**, 277–282.

Gottschalk,A. *et al*. (1998) A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. *RNA*, **4**, 374–393.

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

Gozani,O., Potashkin,J. and Reed,R. (1998) A potential role for U2AF–SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol. Cell. Biol.*, **18**, 4752–4760.

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

Igel,H., Wells,S., Perriman,R. and Ares,M.,Jr (1998) Conservation of structure and subunit interactions in yeast homologues of splicing factor 3b (SF3b) subunits. *RNA*, **4**, 1–10.

Jenny,A. and Keller,W. (1995) Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage and polyadenylation specificity factor. *Nucleic Acids Res.*, **23**, 2629–2635.

Keller,W., Bienroth,S., Lang,K.M. and Christofori,G. (1991) Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3′ processing signal AAUAAA. *EMBO J.*, **10**, 4241–4249.

Krämer,A. (1988) Presplicing complex formation requires two proteins and U2 snRNP. *Genes Dev.*, **2**, 1155–1167.

Krämer,A. (1992) Purification of splicing factor SF1, a heat stable protein that functions in the assembly of a presplicing complex. *Mol. Cell. Biol.*, **12**, 4545–4552.

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

Krämer,A. and Utans,U. (1991) Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. *EMBO J.*, **10**, 1503–1509.

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

Lafontaine,D. and Tollervey,D. (1996) One-step PCR mediated strategy for the construction of conditionally expressed and epitope tagged yeast proteins. *Nucleic Acids Res.*, **24**, 3469–3471.

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

Liao,X.C., Colot,H.V., Wang,Y. and Rosbash,M. (1992) Requirements for U2 snRNP addition to yeast pre-mRNA. *Nucleic Acids Res.*, **20**, 4237–4245.

Mann,M. (1996) A shortcut to interesting human genes: peptide sequence tags, expressed-sequence tags and computers. *Trends Biol. Sci.*, **21**, 494–495.

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., Query,C. and Sharp,P. (1993) Splicing of precursors to mRNA by the spliceosome. In Gesteland,R. and Atkins,J. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.

Murthy,K.G. and Manley,J.L. (1995) The 160-kD subunit of human cleavage–polyadenylation specificity factor coordinates pre-mRNA 3′-end formation. *Genes Dev.*, **9**, 2672–2683.

Neubauer,G., Gottschalk,A., Fabrizio,P., Séraphin,B., Lührmann,R. and Mann,M. (1997) Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proc. Natl Acad. Sci. USA*, **94**, 385–390.

Newman,A. (1994) Analysis of pre-mRNA splicing in yeast. In Higgins,S.J. and Hames,B.D. (eds), *RNA Processing*. Vol. I. Oxford University Press, Oxford, UK, pp. 179–196.

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

Polycarpou-Schwarz,M., Gunderson,S., Kandels-Lewis,S., Séraphin,B. and Mattaj,I. (1996) *Drosophila* SNF/D25 combines the function of the two snRNP proteins U1A and U2B″ that are encoded separately in human, potato and yeast. *RNA*, **2**, 11–23.

Puig,O., Rutz,B., Luukkonen,B.G.M., Kandels-Lewis,S., Bragado-Nilsson,E. and Séraphin,B. (1998) New constructs and strategies for efficient PCR-based gene manipulations in yeast. *Yeast*, **14**, 1139–1146.

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

Reardon,J.T., Nichols,A.F., Keeney,S., Smith,C.A., Taylor,J.S., Linn,S. and Sancar,A. (1993) Comparative analysis of binding of human damaged DNA-binding protein (XPE) and *Escherichia coli* damage recognition protein (UvrA) to the major ultraviolet photoproducts: T[c,s]T, T[t,s]T, T[6-4]T and T[Dewar]T. *J. Biol. Chem.*, **268**, 21301–21308.

Reed,R. (1996) Initial splice-site recognition and pairing during pre-mRNA splicing. *Curr. Opin. Genet. Dev.*, **6**, 215–220.

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

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

Salgado-Garrido,J., Bragado-Nilsson,E., Kandels-Lewis,S. and Séraphin,B. (1999) Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *EMBO J.*, **18**, 3451–3462.

Schmidt-Zachmann,M.S., Knecht,S. and Kramer,A. (1998) Molecular characterization of a novel, widespread nuclear protein that colocalizes with spliceosome components. *Mol. Biol. Cell*, **9**, 143–160.

Séraphin,B. (1995) Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *EMBO J.*, **14**, 2089–2098.

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

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

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

Séraphin,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.

Shevchenko,A., Wilm,M., Vorm,O. and Mann,M. (1996) Mass spectrometric sequencing of proteins on silver-stained polyacrylamide gels. *Anal. Chem.*, **68**, 850–858.

Soni,R. and Carmichael,J.P. (1993) Parameters affecting lithium acetate-mediated transformation of *S.cerevisiae* and development of a rapid and simplified procedure. *Curr. Genet.*, **24**, 455–459.

Staknis,D. and Reed,R. (1994) Direct interactions between pre-mRNA and six U2 small nuclear ribonucleoproteins during spliceosome assembly. *Mol. Cell. Biol.*, **14**, 2994–3005.

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

Stumpf,G. and Domdey,H. (1996) Dependence of yeast pre-mRNA 3′-end processing on CFT1: a sequence homolog of the mammalian AAUAAA binding factor. *Science*, **274**, 1517–1520.

Takao,M., Abramic,M., Moos,M.,Jr, Otrin,V.R., Wootton,J.C., McLenigan,M., Levine,A.S. and Protic,M. (1993) A 127 kDa component of a UV-damaged DNA-binding complex, which is defective in some xeroderma pigmentosum group E patients, is homologous to a slime mold protein. *Nucleic Acids Res.*, **21**, 4111–4118.

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

Teem,J.L. and Rosbash,M. (1983) Expression of a β-galactosidase gene containing the ribosomal protein 51 intron is sensitive to the *rna2* mutation of yeast. *Proc. Natl Acad. Sci. USA*, **80**, 4403–4407.

Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876–4882.

Wang,C., Chua,K., Seghezzi,W., Lees,E., Gozani,O. and Reed,R. (1998) Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. *Genes Dev.*, **12**, 1409–1414.

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

Will,C. and Lührmann,R. (1997) Protein functions in pre-mRNA splicing. *Curr. Opin. Cell Biol.*, **9**, 320–328.

Wilm,M. and Mann,M. (1996) Analytical properties of the nanoelectrospray ion source. *Anal. Chem.*, **68**, 1–8.

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.

Zavanelli,M.I. and Ares,M.,Jr (1991) Efficient association of U2 snRNPs with pre-mRNA requires an essential U2 RNA structural element. *Genes Dev.*, **5**, 2521–2533.