The EMBO Journal Vol.18 No.16 pp.4535–4548, 1999

Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6·U5] tri-snRNP

Alexander Gottschalk[1], Gitte Neubauer[2], Josette Banroques[3], Matthias Mann[2,4], Reinhard Lührmann[1,5,6] and Patrizia Fabrizio[1,6]

[1]Institut für Molekularbiologie und Tumorforschung (IMT), Philipps-Universität Marburg, Emil-Mannkopff-Strasse 2, D-35037 Marburg, [2]European Molecular Biology Laboratory (EMBL), Protein and Peptide Group, Meyerhofstr. 1, D-69117 Heidelberg and [5]Department of Cellular Biochemistry, Max-Planck-Institute of Biophysical Chemistry, Am Fassberg 11, D-37077 Göttingen, Germany and [3]Centre de Génétique Moléculaire du CNRS, Laboratoire Propre Associé à l'Université Pierre et Marie Curie, 91198 Gif-sur-Yvette, France

[4]Present address: Protein Interaction Laboratory, Department of Molecular Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark

[6]Corresponding authors
e-mail: fabrizio@imt.uni-marburg.de and luehrmann@imt.uni-marburg.de

The 25S [U4/U6·U5] tri-snRNP (small nuclear ribonucleoprotein) is a central unit of the nuclear pre-mRNA splicing machinery. The U4, U5 and U6 snRNAs undergo numerous rearrangements in the spliceosome, and knowledge of all of the tri-snRNP proteins is crucial to the detailed investigation of the RNA dynamics during the spliceosomal cycle. Here we characterize by mass spectrometric methods the proteins of the purified [U4/U6·U5] tri-snRNP from the yeast *Saccharomyces cerevisiae*. In addition to the known tri-snRNP proteins (only one, Lsm3p, eluded detection), we identified eight previously uncharacterized proteins. These include four Sm-like proteins (Lsm2p, Lsm5p, Lsm6p and Lsm7p) and four specific proteins named Snu13p, Dib1p, Snu23p and Snu66p. Snu13p comprises a putative RNA-binding domain. Interestingly, the *Schizosaccharomyces pombe* orthologue of Dib1p, Dim1p, was previously assigned a role in cell cycle progression. The role of Snu23p, Snu66p and, additionally, Spp381p in pre-mRNA splicing was investigated *in vitro* and/or *in vivo*. Finally, we show that both tri-snRNPs and the U2 snRNP are co-precipitated with protein A-tagged versions of Snu23p, Snu66p and Spp381p from extracts fractionated by glycerol gradient centrifugation. This suggests that these proteins, at least in part, are also present in a [U2·U4/U6·U5] tetra-snRNP complex.

Keywords: Lsm proteins/mass spectrometry/pre-mRNA splicing/RSE1/snRNP proteins

Introduction

Nuclear pre-mRNA splicing occurs via a two-step transesterification reaction that is catalysed by a large ribonucleoprotein complex, the spliceosome. The spliceosome is formed by the ordered interaction of the U1, U2, U5 and U4/U6 small nuclear ribonucleoproteins (snRNPs) and several splicing factors with the pre-mRNA (for reviews see Rymond and Rosbash, 1992; Krämer, 1996; Staley and Guthrie, 1998; Burge *et al*., 1999). Splicing complex formation involves the initial interaction of the U1 snRNP with the 5′ splice site, followed by U2 snRNP recognizing the branch site to form the pre-spliceosome. Spliceosome assembly is completed by the subsequent association of the U4/U6 and U5 snRNPs in the form of a 25S [U4/U6·U5] tri-snRNP complex. Activation of the spliceosome into a catalytically active machine requires numerous RNA conformational rearrangements, in particular of the RNAs of the tri-snRNP. Thus, the two helices of the U4/U6 interaction domain dissociate and a new base-pairing interaction is formed between U2 and U6. Concomitant with these events, the U1 snRNP dissociates from the 5′ splice site which is then recognized by the conserved ACAGAG sequence of the U6 snRNA. The conserved loop I of U5 snRNA contacts exonic sequences at the 5′ and 3′ splice sites, probably in a sequential manner, while the splicing reaction proceeds. The first reaction creates two intermediates, a cleaved upstream exon and an intron–downstream exon in a branched lariat configuration, while the second reaction results in the ligation of the exons and excision of the lariat–intron. The spliceosome then disassembles and the products of the reaction are released. After dissociation of the spliceosome, the U4/U6 snRNA duplex is re-formed and the [U4/U6·U5] tri-snRNP complex is regenerated and thought to take part in new rounds of splicing (Moore *et al*., 1993; Staley and Guthrie, 1998).

Interestingly, the formation and dynamics of the spliceosomal RNA interactions are evolutionarily conserved (Nilsen, 1998). A solution to understanding the mechanisms of the several RNA structural rearrangements of the U4, U6 and U5 RNAs may be the knowledge of the protein moiety of the spliceosome, in particular of the tri-snRNP complex itself.

Biochemical purification, protein sequencing and cDNA cloning have demonstrated that the human [U4/U6·U5] tri-snRNP is associated with at least 15 specific proteins, in addition to the seven Sm proteins (B/B′, D1, D2, D3, E, F and G) common to all snRNPs except U6 (for reviews see Will *et al*., 1995; Teigelkamp *et al*., 1998). The analysis of isolated and fractionated yeast snRNPs indicated that the yeast tri-snRNP particle may have a similar protein composition (Fabrizio *et al*., 1994). However, only 10 yeast tri-snRNP proteins have been characterized so far. Eight of these proteins were identified by genetic approaches and their association with the tri-snRNP was demonstrated biochemically; these are Prp8p, Prp4p, Prp6p, Prp38p, Prp18p, Prp3p, Prp31p and Spp381p

(Lossky *et al*., 1987; Banroques and Abelson, 1989; Bjorn *et al*., 1989; Abovich *et al*., 1990; Blanton *et al*., 1992; Horowitz and Abelson, 1993; Anthony *et al*., 1997; Weidenhammer *et al*., 1997; Lybarger *et al*., 1999). The remaining two proteins, Snu114p and Brr2p/Snu246p/Rss1p/Slt22p, were identified via their homology to the human tri-snRNP proteins U5-200 kDa and U5-116 kDa, respectively, and they were also shown to be bona fide yeast tri-snRNP proteins (Lauber *et al*., 1996; Fabrizio *et al*., 1997). The latter protein was identified independently also by genetic means (Lin and Rossi, 1996; Noble and Guthrie, 1996; Xu *et al*., 1996). The association of Sm proteins with the yeast tri-snRNP has also been demonstrated (Rymond, 1993; Roy *et al*., 1995; Séraphin, 1995; Bordonné and Tarassov, 1996). In addition, three Sm-like proteins, Lsm4p/Uss1p, Lsm3p (formerly denoted SmX4p) and Lsm8p, were shown by immunoprecipitation to be associated with the U6 and the U4/U6 snRNP and, to some extent, also with the [U4/U6·U5] tri-snRNP (Cooper *et al*., 1995; Séraphin, 1995; Pannone *et al*., 1998). Six additional Sm-like proteins can be identified in the yeast genome (Fromont-Racine *et al*., 1997). The question of whether all of the Sm-like proteins or only a certain subset of these are integrated into the yeast tri-snRNP remains to be answered.

In this work, we describe our efforts to characterize the protein composition of the yeast [U4/U6·U5] tri-snRNP in more detail using biochemical methods. We therefore exploited the recent advances made in both the development of efficient yeast snRNP purification protocols and the rapid identification of isolated proteins by mass spectrometric methods (Fabrizio *et al*., 1994; Neubauer *et al*., 1997; Gottschalk *et al*., 1998). Specifically, we have isolated the [U4/U6·U5] tri-snRNP from yeast in preparative amounts using a strain that expresses the tri-snRNP protein Prp4p with a histidine tag on its C-terminus. The isolation procedure involves: (i) an initial affinity purification of all m_3G-capped snRNPs from crude extracts using anti-m_3G-cap antibodies; (ii) selection of the [U4/U6·U5] tri-snRNP from this mixture by nickel-nitrilotriacetic acid (NTA) chromatography; and (iii) a final purification of the tri-snRNP by glycerol gradient centrifugation. The tri-snRNP proteins were partially sequenced by a combination of MALDI and nanoelectrospray mass spectrometry and their genes were identified in the database via multiple peptide sequence tags (Shevchenko *et al*., 1996; Wilm *et al*., 1996; Neubauer *et al*., 1997). We identified eight novel proteins in the tri-snRNP. These include four Sm-like proteins, Lsm2p, Lsm5p, Lsm6p and Lsm7p, and four specific proteins termed Snu13p, Dib1p, Snu23p and Snu66p. The role in pre-mRNA splicing of Snu23p and Snu66p has been investigated employing *in vitro* and/or *in vivo* approaches. These studies additionally include the tri-snRNP protein Spp381p which was identified independently by genetic methods while our work was in progress (Lybarger *et al*., 1999). Interestingly, protein A (protA)-tagged versions of the latter three proteins allowed, in addition to tri-snRNPs, the precipitation also of U2 snRNPs, most likely as a preformed [U2·U4/U6·U5] tetra-snRNP complex. Moreover, we show that the splicing factor Rse1p (Chen *et al*., 1998) which was co-isolated here is an U2 snRNP-associated protein.

Results and discussion

Purification of the yeast [U4/U6·U5] tri-snRNP

We have used our previously established protocol (Fabrizio *et al*., 1994; Neubauer *et al*., 1997) to purify the [U4/U6·U5] tri-snRNP at 100 mM salt from a strain expressing 8×His-tagged Prp4p, an U4/U6 snRNP-specific protein (Banroques and Abelson, 1989; Bjorn *et al*., 1989). From the total snRNP mixture obtained after anti-m_3G-cap affinity chromatography, the [U4/U6·U5] tri-snRNP as well as low amounts of U1 and U2 snRNPs were isolated by Ni^{2+}-NTA chromatography. We observed previously that an interaction of U1 and U5 snRNPs led to co-purification of the two particles (Neubauer *et al*., 1997), which could explain the presence of U1 snRNP here. However, co-isolation of U2 snRNPs together with tri-snRNPs was unexpected (see below). To separate the [U4/U6·U5] tri-snRNP from U1 and U2 snRNPs, as well as from proteins bound non-specifically to either of the two columns or to the tri-snRNP, the isolated material was subjected to 10–30% glycerol gradient centrifugation at 200 mM KCl. The sedimentation profile of snRNPs in this gradient is shown in Figure 1A as an analysis of snRNAs extracted from the gradient fractions. The [U4/U6·U5] tri-snRNP peaked in fractions corresponding to a sedimentation coefficient of 25S (fractions 14–16), while most of the U1 and U2 snRNPs sedimented at ~18S (fractions 11–13). Fractions 15 and 16 contained pure [U4/U6·U5] tri-snRNP, essentially free of U1 and U2 snRNPs.

The protein profile of the snRNP sedimentation in the 16–25S region of the gradient is shown in Figure 1B. At least 34 protein bands with mol. wts ranging from ~10 to 200 kDa sedimented in this region. Peaks of U1 (lanes 12 and 13) and tri-snRNP proteins (lanes 14 and 15) that correspond to the respective RNA peaks (Figure 1A) can be distinguished. This is exemplified by the comparison of the sedimentation of the U1-specific proteins in bands 7 and 8 (identified as Snu71p and a mixture of Prp39p and Prp40p, respectively; see below) with the tri-snRNP proteins in bands 1 and 2 (Prp8p and Brr2p/Snu246p/Rss1p/Slt22p, respectively; Figure 1B and C). A number of proteins remained at the top of the gradient (data not shown).

Mass spectrometric analysis of snRNP proteins

In order to obtain enough material to identify the isolated proteins by mass spectrometry, proteins from the fractions of four separate gradients containing the majority of the [U4/U6·U5] tri-snRNP (corresponding to fractions 13–16, Figure 1B) were pooled together (henceforth termed tri-snRNP pool). Although we omitted the peak fractions of the U1 and U2 snRNPs, low amounts of these particles were also present in the tri-snRNP pool (Figure 1A). We reasoned that by combining information about the identity of the proteins and their co-sedimentation with snRNAs in gradients as well as by performing tagging and immunoprecipitation experiments, we would be able to assign each protein unequivocally to either the [U4/U6·U5] tri-snRNP, U1 or U2 snRNP. The proteins of the tri-snRNP pool were separated by one-dimensional SDS–PAGE (Figure 1C). A total of 34 bands were excised, and peptides generated proteolytically from these bands were

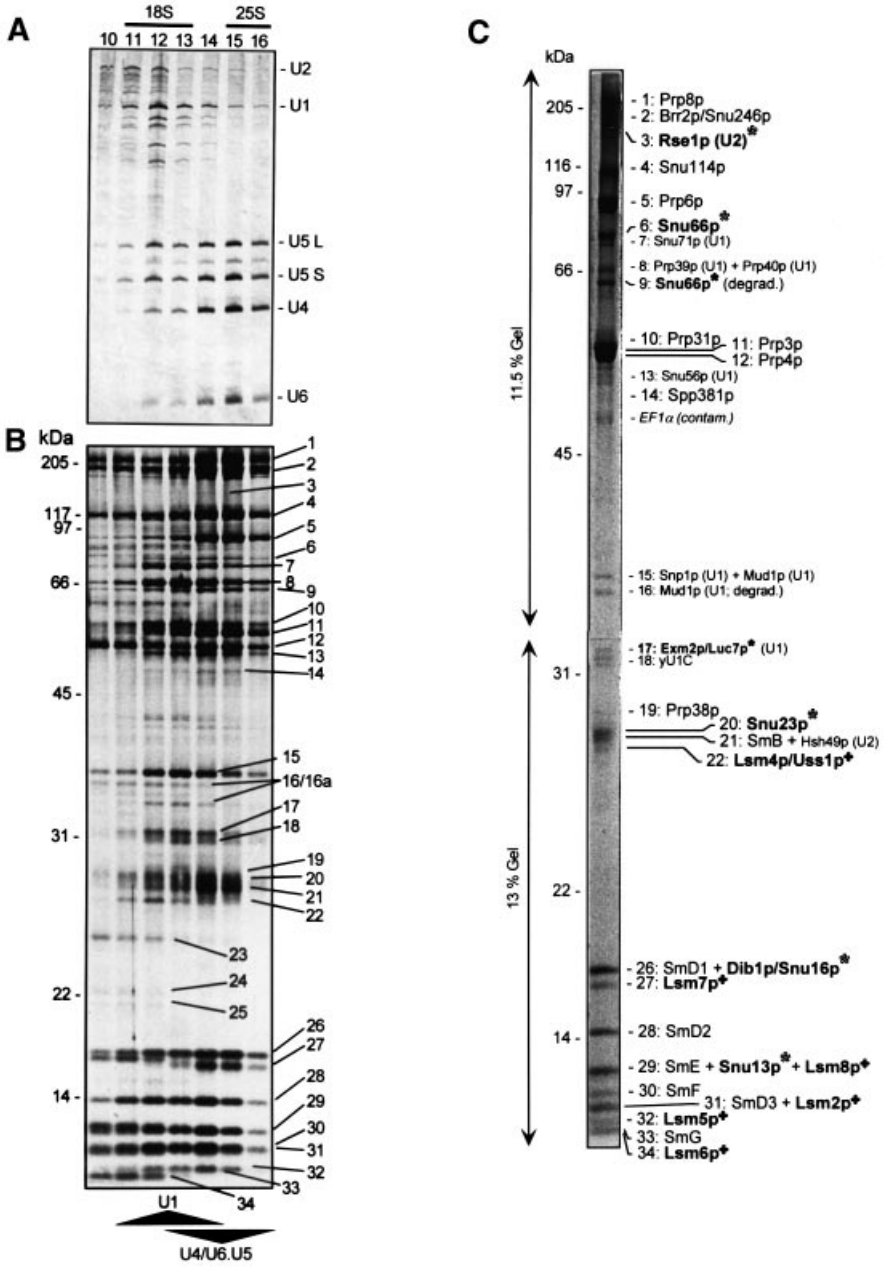

Fig. 1. Biochemical purification of the yeast [U4/U6·U5] tri-snRNP and identification of its protein components by mass spectrometry. The tri-snRNP was purified from whole-cell extract by anti-m_3G-cap immunoaffinity and Ni^{2+}-NTA chromatography followed by 10–30% glycerol gradient centrifugation. Gradient fractions were extracted by phenol/chloroform/isoamylalcohol and analysed for their RNA (**A**) and protein content (**B**). Gels in (A) and (B) were stained with silver. The identity of the snRNAs is indicated. Yeast rRNAs from a parallel gradient were analysed as sedimentation markers (not shown). (**C**) Proteins from fractions 13–16 of four different gradients [that were technically identical to that shown in (B)], containing the majority of the tri-snRNP were pooled, separated on a 60 cm SDS–gel consisting of 11.5% (upper part) and 13% acrylamide (lower part) and Coomassie Blue stained. The protein bands were excised, in-gel digested, and peptides were then eluted from the gel and analysed by a combination of MALDI and nanoelectrospray tandem mass spectrometry. The bands that were excised from the gel shown in (C) are numbered [also in (B)] and their identity is indicated on the right. Several bands contained mixtures. Novel proteins are in bold and marked by an asterisk, Sm-like proteins are in bold and labelled with a cross, and the association with U1 or U2 snRNPs is indicated in parentheses (all other proteins are tri-snRNP associated). One contaminating protein (EF1α) smeared into the tri-snRNP peak from the top of the gradient. Note that (A) and (B) do not originate from the same gradient, so that snRNA and protein peaks do not correspond exactly.

analysed by mass spectrometry. Since several of the bands contained mixtures of two or more proteins, both MALDI peptide mapping and nanoelectrospray tandem mass spectrometry were applied in order to ensure the identification of all the proteins present in one band (as exemplified by the tandem mass spectrometric analysis of band 14; Figure 2). The proteins identified by mass spectrometry and database searching are listed in Table I, and the respective bands are indicated in Figure 1B and C.

We additionally pooled the proteins from the top of the gradient and treated these as described above for the tri-snRNP pool (data not shown). Eight bands were obtained, from which nine proteins could be identified. Among these was Npl3p, which we previously had co-isolated

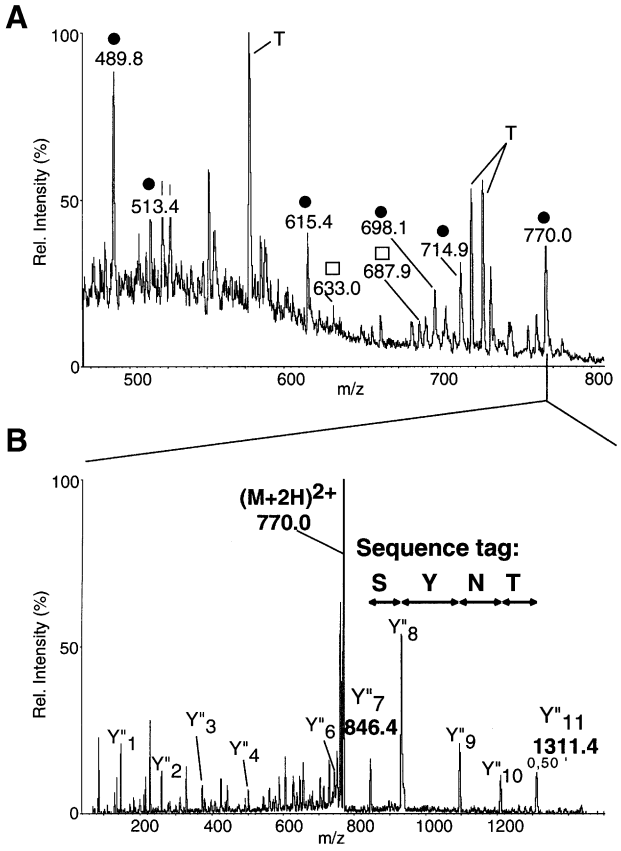

Fig. 2. Mass spectrometric analysis of the protein mixture found in band 14. (**A**) Part of the Q1 scan (normal mass spectrum) of the peptide mixture obtained after tryptic digestion of the protein band and micropurification of the resulting peptide mixture. The labelled peaks were selected in the first part of the mass spectrometer in turn and fragmented in its collision chamber. Mass spectra of the resulting fragments (tandem mass spectra) contain peptide ions which were subjected to tandem mass spectrometry, and the partial sequences derived can be used to identify the proteins in a sequence database using peptide sequence tags (Mann and Wilm, 1994). Two different proteins were identified by tandem mass spectrometry: marked with bullets are the peptides derived from Spp381p, the squares indicate peptides originating from a degradation product of Prp31p, and the peaks labelled with T are autolysis products of trypsin. (**B**) Product ion scan of the doubly charged peptide ion at m/z 770.0. From the fragment spectrum, a partial sequence of this peptide was determined, the sequence tag (846.4)SYNT (1311.4) was constructed and a database search identified uniquely Spp381p in the database. The C-terminal or Y″ ion series confirms the sequence of this peptide (LNTNYSTNEELIK).

with the U1 snRNP and which has a low affinity for U1, U2 and [U4/U6·U5] tri-snRNPs (Gottschalk *et al.*, 1998). The remaining eight proteins are either enzymes of metabolic pathways or 'housekeeping' proteins and thus are likely to be contaminants (data not shown).

From the tri-snRNP pool, a total of 36 distinct proteins could be identified (Figure 1B and C; Table I). Seven of these were the canonical Sm proteins B, D1, D2, D3, E, F and G, which associate with the U1, U2, U4 and U5 snRNAs. Due to the presence of U1 and some U2 snRNPs in the tri-snRNP pool, several U1- and U2-specific proteins were also identified. These include seven previously characterized U1 snRNP proteins (Figure 1C; Table I) as well as one novel U1 protein of 32 kDa (YDL087c), termed Luc7p/Exm2p, which recently has been identified in a genetic screen (P.Fortes, D.Bibao-Cortes, M.Fornerod, G.Rigaut, W.Raymond and I.W.Mattaj, submitted). Reinspection of our previous U1 protein sequencing data (Neubauer *et al.*, 1997) revealed that Luc7p/Exm2p was also present in our original U1 snRNP preparation. Since its sequence was not available in the yeast genome database at that time, we were unable to ascribe the identified peptides to any yeast protein. This protein was not studied further here.

Two proteins of the tri-snRNP pool were revealed to be U2-specific, one of which was the previously described Hsh49p (Igel *et al.*, 1998; Table I). The second protein was Rse1p (Figure 1B and C; Table I) which recently has been shown to be involved in pre-mRNA splicing (Chen *et al.*, 1998). Immunoprecipitations from whole-cell extracts harbouring protA-tagged Rse1p clearly demonstrated that this protein is associated primarily with U2 snRNPs (Figure 3, lanes 16–18). In addition to U2, low amounts of [U4/U6·U5] tri-snRNPs were co-precipitated reproducibly with Rse1-protAp. Recently, Rse1p was identified independently in partially purified U2 snRNPs, and its association with U2 was also demonstrated (Caspary *et al.*, 1999). Rse1p is the likely yeast orthologue of the human protein SAP130, which, together with Hsh49p (SAP49), is a subunit of the heteromeric U2 snRNP-associated splicing factor SF3b (R.R.Reed, personal communication). The remaining 19 proteins identified in the tri-snRNP pool were revealed to be either known or novel [U4/U6·U5] tri-snRNP-associated proteins and are described in detail below.

Sm-like proteins in the purified [U4/U6·U5] tri-snRNP

The U6 and the U4/U6 snRNPs are associated with at least three Sm-like proteins, namely Lsm4p/Uss1p, Lsm3p and Lsm8p (Cooper *et al.*, 1995; Séraphin, 1995; Pannone *et al.*, 1998). Additionally, some U5 snRNA was co-precipitated with all three proteins, indicating that they are also associated with the [U4/U6·U5] tri-snRNP, but this co-precipitation was relatively weak and salt sensitive. Two of the known Lsm proteins, namely Lsm4p/Uss1p and Lsm8p, and four novel proteins bearing Sm-like domains, Lsm2p (YBL026w), Lsm5p (YER146w), Lsm6p (YDR378c) and Lsm7p (YNL147w), were identified in the tri-snRNP pool. However, we failed to detect Lsm3p. Since the Sm and Sm-like proteins were found in mixtures containing up to four proteins or protein degradation products, it is possible that Lsm3p eluded detection. Alternatively, this protein may have dissociated from the tri-snRNP during glycerol gradient sedimentation of the isolated snRNPs at 200 mM salt. Consistent with the latter possibility is our observation that some other Lsm proteins dissociated, at least in part, from the tri-snRNP during glycerol gradient centrifugation (Figure 1B).

We would like to point out that the Sm-like proteins were also co-isolated with [U4/U6·U5] tri-snRNPs purified from a strain containing the His-tagged U5-specific protein Snu114p (data not shown). We are therefore confident that the Sm-like proteins characterized in this work are present in the [U4/U6·U5] tri-snRNP complex and are not co-isolated as part of a free U4/U6 snRNP complex. It is also unlikely that we have co-isolated the free U6 snRNP, since its snRNA lacks an m_3G-cap and should therefore not be retained in the first biochemical (anti-m_3G-cap)

Table I. Protein composition of the *S.cerevisiae* [U4/U6·U5] tri-snRNP

[U4/U6·U5] tri-snRNP proteins	M_r(kDa)	Band No.	Motifs	References
Prp8p	280	1	–	Lossky *et al*. (1987)
Brr2p/Snu246p/ Slt22p/Rss1p	246	2	DEXH RNA unwindase	Noble and Guthrie (1996); Lauber *et al*. (1996); Xu *et al*. (1996); Lin and Rossi (1996)
Snu114p	114	4	GTPase domain, EF2-like	Fabrizio *et al*. (1997)
Prp6p	104	5	Contains TPR repeats	Abovich *et al*. (1990)
Snu66p (YOR308c)	66	6, 9 (degr.)	–	this study
Prp31p	56	10	–	Weidenhammer *et al*. (1997)
Prp3p	56	11	–	Anthony *et al*. (1997)
Prp4p	52	12	WD-repeats	Banroques and Abelson (1989); Bjorn *et al*. (1989)
Spp381p	34	14	PEST proteolysis motif	this study; Lybarger *et al*. (1999)
Prp38p	28	19	Serine rich, acidic	Blanton *et al*. (1992)
Prp18p	25	–	–	identified only by Western analysis in this study; Horowitz and Abelson (1993)
Snu23p (YDL098c)	23	20	C_2H_2 type zinc finger	this study
Dib1p/Snu16p	16	26	–	this study
Snu13p (YEL026w)	13	29	Putative RNA binding	this study
Sm proteins				
SmB	22	21, 23–24 (degr.)	Sm motifs 1 and 2	Gottschalk *et al*. (1998)
SmD1	16	26	Sm motifs 1 and 2	Rymond (1993)
SmD2	13	28	Sm motifs 1 and 2	Neubauer *et al*. (1997)
SmD3	11	31	Sm motifs 1 and 2	Roy *et al*. (1995)
SmE	10	29	Sm motifs 1 and 2	Bordonné and Tarassov (1996)
SmF	10	30	Sm motifs 1 and 2	Séraphin (1995)
SmG	9	33	Sm motifs 1 and 2	Séraphin (1995)
Lsm proteins				
Lsm2p (YBL026w)	11	31	Sm motifs 1 and 2	this study
Lsm3p	10	–	Sm motifs 1 and 2	not identified in this study; Séraphin (1995)
Lsm4p/Uss1p	21	22, 23–25 (degr.)	Sm motifs 1 and 2	Cooper *et al*. (1995)
Lsm5p (YER146w)	10	32	Sm motifs 1 and 2	this study
Lsm6p (YDR378c)	14	34	Sm motifs 1 and 2	this study
Lsm7p (YNL147w)	12	27	Sm motifs 1 and 2	this study
Lsm8p	15	29	Sm motifs 1 and 2	Pannone *et al*. (1998)

The tri-snRNP proteins are grouped and their calculated molecular weights, sequence motifs in the respective protein and references are shown. The numbering of protein bands indicates that these proteins were identified by mass spectrometry in this study and corresponds to Figure 1B. Some bands that contained only degradation products are indicated ('degr.'). The U6-specific protein Lsm3p was not identified in this study, but was previously shown to be part of the tri-snRNP (Séraphin, 1995). Prp18p was identified here only by Western blot analysis. U2 snRNP proteins identified in this study: Rse1p (154 kDa; band 3; this study and Chen *et al*., 1998); Hsh49p (band 21). U1 snRNP proteins also identified in this study: Snu71p (band 7), Prp39p (band 8), Prp40p (band 8), Snu56p (band 13), Mud1p (band 15 and 16/16a; degr.), Snp1p (band 15), Exm2p/Luc7p (band 17; P.Fortes, submitted), yU1C (band 18).

purification step. The association of the additional Lsm proteins with the yeast [U4/U6·U5] tri-snRNP was also demonstrated independently by other workers (Mayes *et al*., 1999; Salgado-Garrido *et al.*, 1999). Although it is very likely that similar to Lsm4p and Lsm8p, the four additional Lsm proteins identified here associate directly with U6 snRNA in the tri-snRNP particle, we have as yet no direct proof for this. We also do not know whether the Lsm proteins interact specifically with each other in the tri-snRNP in a way that is similar to the situation in the canonical Sm core structure (Raker *et al*., 1996; Fury *et al*., 1997; Camasses *et al*., 1998; Kambach *et al*., 1999). Clearly, more work is needed to clarify the structural organization of the Lsm proteins in the yeast tri-snRNP.

In addition to the seven Sm-like proteins Lsm2p–Lsm8p, open reading frames encoding two other Lsm proteins, Lsm1p (YJL124c) and Lsm9p (YCR020c-a), were observed in the yeast genome (Fromont-Racine *et al*., 1997). Lsm1p and Lsm9p were not identified in the isolated tri-snRNP nor has a co-precipitation of U6 or U4/U6 snRNPs been reported as yet for these two proteins. Interestingly, Lsm1p recently has been shown to interact genetically with Dcp1p, the enzyme responsible for the decapping of mRNAs in the cytoplasm (Boeck *et al*., 1998), indicating that Lsm proteins are also involved in processes other than pre-mRNA splicing.

Identification of novel tri-snRNP-specific proteins

Nine proteins from the tri-snRNP pool could be assigned unambigously to previously described tri-snRNP components, namely Prp8p, Brr2p/Snu246p/Slt22p/Rss1p, Snu114p, Prp6p, Prp31p, Prp3p, Prp4p, Spp381p and Prp38p (Table I). Prp18p, a further known [U4/U6·U5] tri-snRNP component, was not present in the tri-snRNP pool; however, we identified Prp18p by Western blot analysis at the top of the glycerol gradient (data not shown). Prp18p was most probably co-purified as part of the tri-snRNP at 100 mM KCl, but dissociated from the tri-snRNP during glycerol gradient centrifugation at 200 mM KCl (see also Horowitz and Abelson, 1993). The

remaining four proteins are novel tri-snRNP-associated components, termed Snu13p (YEL026w), Dib1p (YPR082c), Snu23p (YDL098c) and Snu66p (YOR308c).

Snu13p encodes a protein of 125 amino acid residues and contains a domain with significant similarity to the yeast H/ACA box snoRNP protein Nhp2p (Henras *et al*., 1998; Watkins *et al*., 1998). Part of this domain has been proposed to function as a possible RNA recognition motif (Koonin *et al*., 1994). We recently have characterized the human orthologue of Snu13p, the tri-snRNP-specific 15.5 kDa protein, and have shown that it binds directly to U4 snRNA (S.Nottrott, K.Hartmuth, P.Fabrizio, H.Urlaub, I.Vidovic, R.Ficner and R.Lührmann, submitted).

Dib1p (YPR082c) comprises 142 amino acids and is an orthologue of the *Schizosaccharomyces pombe* Dim1p, which previously has been assigned a role in cell cycle progression (Berry and Gould, 1997; Berry *et al*., 1999). Both proteins are essential for cell viability (Berry and Gould, 1997). A mutant of Dim1p was shown to exert a cell cycle arrest before entry into mitosis in *S.pombe* cells; however, the biochemical activity of Dim1p was not identified. Dib1p is clearly an orthologue of the human U5-specific 15 kDa protein. A detailed characterization of the structure of the human U5-15 kDa protein and the essential role of Dib1p in splicing is described in a separate manuscript (K.Reuter, S.Notrott, P.Fabrizio, R.Lürhmann and R.Ficner, submitted). We suggest to name Dib1p also Snu16p, since the name Dib1p does not relate to the splicing process at all.

The Snu23p and Snu66p proteins comprise 194 and 587 amino acids, respectively. While Snu23p contains a putative C_2H_2-type zinc-finger motif between amino acids 82 and 104, Snu66p does not exhibit known sequence motifs, so that no putative functions can be predicted for these proteins. Therefore, the possible role of Snu66p and Snu23p in splicing was examined here in more detail. In our studies, we additionally have included Spp381p, which was also identified by genetic methods and demonstrated to be a tri-snRNP protein while our work was in progress (Lybarger *et al*., 1999).

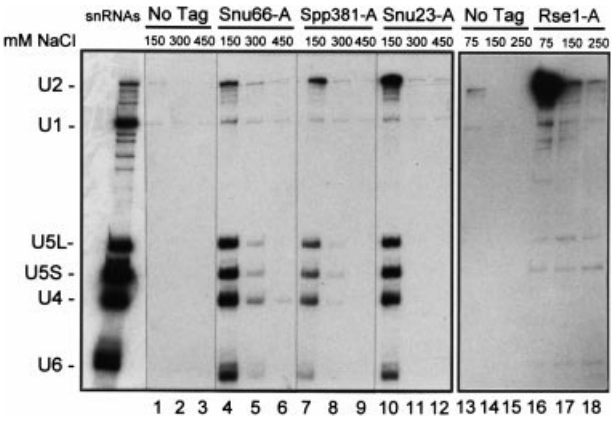

Fig. 3. Association of Snu66-protAp, Spp381-protAp, Snu23-protAp and Rse1-protAp with snRNPs. DNA coding for a tandem repeat of the IgG-binding domain of the *S.aureus* protein A was fused downstream of the coding regions of either *SNU66*, *SPP381*, *SNU23* or *RSE1*. Extracts from the resulting strains, as well as from the parental wild-type strain (lanes 1–3), were immunoprecipitated at various salt concentrations, indicated above the lanes, using IgG–agarose. Precipitation reactions for Rse1-protAp (lanes 16–18) and the corresponding negative controls (lanes 13–15) originate from a different, though technically identical experiment. The snRNA content of the precipitates was assayed by Northern blot analysis (snRNAs are indicated on the left). In the leftmost lane, snRNAs prepared from the tri-snRNP pool were loaded as a control for the labelled probes used. The U4 probe used in the Northern blot analysis of the Rse1-protAp precipitation was unusually weak. After longer exposure, the U4 snRNA signal was clearly visible.

Immunoprecipitation studies reveal an association of Snu66p, Spp381p and Snu23p with both [U4/U6·U5] tri-snRNPs and U2 snRNPs

To corroborate our findings that Snu66p and Snu23p associate with snRNPs, we have used proteins tagged *in vivo* with the protA tag for immunoprecipitation assays (Gottschalk *et al*., 1998; Puig *et al*., 1998). These were performed with the respective extracts at different salt concentrations, and the co-precipitated snRNAs were identified by Northern blot analysis (Figure 3).

At 150 mM salt, the U4, U5 and U6 snRNAs were co-precipitated efficiently with Snu66p and Snu23p, verifying that these two proteins associate with the [U4/U6·U5] tri-snRNP in whole-cell extracts (Figure 3, lanes 4 and 10). The tri-snRNP was also co-precipitated with Spp381p, but not as efficiently as with the other two proteins (Figure 3, lane 7). Interestingly, all immunoprecipitates contained, in addition to tri-snRNPs, also significant amounts of U2 snRNP, yet only slightly higher than background amounts of U1 snRNP (Figure 3, lanes 4, 7 and 10; see also below). Immunoprecipitation of U2 and [U4/U6·U5] tri-snRNPs was specific, as no snRNPs were precipitated from splicing extracts harbouring untagged Snu66p, Spp381p and Snu23p (Figure 3, lane 1). We note that while Lybarger *et al.* (1999) also precipitated [U4/U6·U5] tri-snRNPs with Spp381p, they did not observe any co-precipitation of U2 snRNPs. This discrepancy could be due to the usage of different affinity tags and/or salt concentrations in the two studies.

Next, we investigated the salt sensitivity of the association of Snu66p, Spp381p and Snu23p with the snRNPs. At 300 mM salt, all three proteins were dissociated from U2 snRNP (Figure 3, lanes 5, 8 and 11). Snu23p was also completely dissociated from the tri-snRNP at this salt concentration. In contrast, significant amounts of [U4/U6·U5] tri-snRNPs were still co-precipitated with Snu66p, and low amounts of tri-snRNP with Spp381p, indicating that the two proteins have a higher affinity for tri-snRNPs than U2 snRNPs (see also Figure 4, as described below).

The co-precipitation of both U2 and [U4/U6·U5] tri-snRNPs with Snu66p, Spp381p and Snu23p from extracts raised the interesting question of whether the proteins associated independently with each of the two snRNP particles and/or whether U2 snRNP was co-precipitated as part of a [U2·U4/U6·U5] tetra-snRNP complex. To address this question, we fractionated whole-cell extracts containing protA-tagged Snu66p, Spp381p or Snu23p by glycerol gradient centrifugation prior to immunoprecipitation, and subsequently analysed the snRNPs that were co-precipitated with each of the three proteins across the glycerol gradient.

The distribution of snRNPs in the glycerol gradient prior to immunoprecipitation is shown in Figure 4A, as analysed by Northern blotting of the snRNAs. The bulk of U1 and U2 snRNPs sedimented in fractions 6–10, with

a peak at 18S, but significant amounts of both snRNPs are also observed in fractions 11–15 (corresponding to 25–35S). Free U5 snRNPs sedimented in fractions 6–8, while the 25S [U4/U6·U5] tri-snRNP peaked in fractions 9–11 (Figure 4A). The analyses of the snRNAs co-immunoprecipitated with Snu66p, Spp381p and Snu23p are shown in Figure 4B–D, respectively. Consistent with our immunoprecipitation studies performed with unfractionated splicing extracts, [U4/U6·U5] tri-snRNPs and U2 snRNPs were co-precipitated with all three proteins, whereby the relative ratio of precipitated U2 to tri-snRNP varies for the three proteins (Figure 4B–D). Significantly, no U1 snRNP was precipitated with any of the proteins, demonstrating the specificity of the interaction of the proteins with U2 and tri-snRNP particles. Of particular importance is the comparison of the gradient fractions from which U2 and tri-snRNPs are co-precipitated preferentially (Figure 4B–D) with respect to the distribution of the bulk of these snRNPs in the gradient (Figure 4A). This is most striking for U2 snRNP: although the majority of U2 snRNP sediments in the 16–20S region (fractions 6–10), U2 snRNPs are precipitated exclusively from the 25–35S region (fractions 9–15). Moreover, the peaks of immunoprecipitated U2 and tri-snRNPs coincide for all three proteins (Figure 4B–D). Significantly, the amounts of [U4/U6·U5] tri-snRNPs and U2 snRNPs precipitated from fractions 13–15 of the gradient (corresponding to 30–35S) are clearly enriched when compared with the overall distribution of the respective snRNPs. This enrichment is particularly pronounced in the case of Snu23p (Figure 4D). Our results support the idea that the U2 snRNP, at least in part, is co-precipitated with Spp381p, Snu23p and, to some extent, also with Snu66p as an integral part of a 30–35S [U2·U4/U6·U5] tetra-snRNP. However, U2 and tri-snRNPs were also co-precipitated with the three proteins from the 25S region of the gradient (fractions 9–13). One explanation for this sedimentation behaviour could be that the [U2·U4/U6·U5] tetra-snRNP represents a mixture of complexes that are heterogeneous with respect to their protein composition. Alternatively, it is quite possible that U2 and tri-snRNPs were precipitated additionally with the three proteins independently from each other. Assuming that the proteins were associated with the [U4/U6·U5] tri-snRNP complex at sites located at the interface of the tri-snRNP and the U2 snRNP in the tetra-snRNP complex, it can be envisaged that they retain affinity for both the tri-snRNP and the U2 snRNP particles after dissociation of the tetra-snRNP. This scenario would be reminiscent of the situation in the *Escherichia coli* ribosome: some proteins that are located in the 70S ribosome at the interface between the 30S and 50S subunits are found in both subunits upon dissociation of the 70S ribosomes (Wittmann, 1980).

The idea that Snu66p, Spp381p and Snu23p can associate with [U4/U6·U5] tri-snRNPs independently of U2 snRNPs was corroborated by the following experiment. After digestion of U2 snRNA in yeast extracts by DNA oligonucleotide-targeted RNase H hydrolysis, [U4/U6·U5]

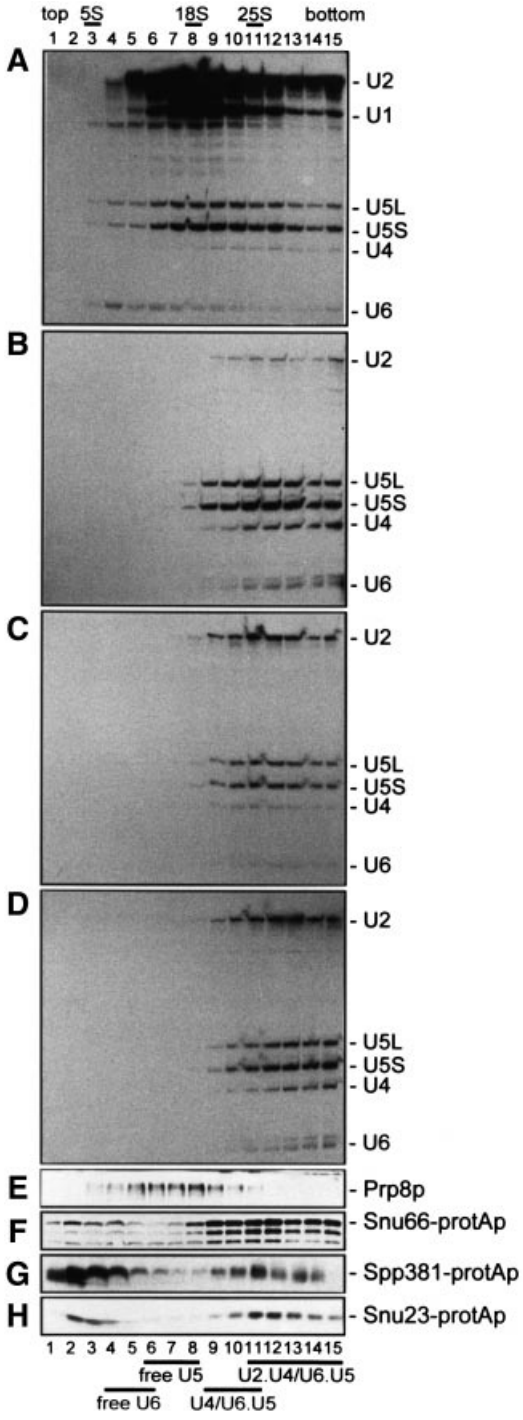

Fig. 4. Snu66p, Spp381p and Snu23p are associated with the [U4/U6·U5] tri-snRNP and a 30–35S [U2·U4/U6·U5] tetra-snRNP. Total cell extracts containing either protA-tagged Snu66p, Spp381p or Snu23p were subjected to 10–30% glycerol gradient centrifugation. rRNAs from yeast were run on a parallel gradient as sedimentation markers (not shown). Each 15 fractions were taken, and a quarter of each fraction was assayed for the total snRNA content by Northern blot (**A**) or for protein content by Western blot analysis for the U5 snRNP-specific Prp8p (**E**), as exemplified using the fractions obtained from the Snu66-protAp extract. The fractions from the three gradients were also assayed for Snu66-protAp (**F**), Spp381-protAp (**G**) and Snu23-protAp (**H**) by Western blot analysis. The remaining three-quarters of the fractions were immunoprecipitated with IgG–agarose. After extensive washing, co-precipitated snRNAs were assayed in the precipitates from the tagged Snu66p (**B**), Spp381p (**C**) and Snu23p (**D**) extract gradient fractions by Northern blotting. Note that longer exposures had to be used in (B–D), as the total snRNA gel shown in (A) contained more material than the precipitates. Upon longer exposure of the film, the signal of Prp8p was also detectable in fractions 12–14 in (E).

tri-snRNPs were co-precipitated efficiently with all three proteins (in the absence of intact U2 snRNPs; data not shown).

To determine what fraction of the soluble cellular amount of each of the Snu66p, Spp381p and Snu23p proteins was actually associated with snRNPs, we investigated the distribution of the three proteins in the glycerol gradient by analysing aliquots from each fraction of the three gradients by Western blotting prior to immunoprecipitation (Figure 4F–H). Significantly, the majority of soluble Snu66p and Snu23p and about half of the soluble Spp381p co-sedimented with the snRNPs in the 25–35S region of the gradient. None of them co-fractionated with the 16–20S U2 or U5 snRNPs, and the remainder of the proteins migrated at the top of the gradients. For comparison, the distribution of the U5-specific protein Prp8p was also analysed across the gradient (Figure 4E). In contrast to the situation described for Snu66p, Spp381p and Snu23p, Prp8p clearly peaks with the bulk of 18S U5 snRNP in the gradient. Upon longer exposure of the immunoblot, Prp8p was also detectable in the 30–35S fractions. Taken together, our combined results indicate that, in cellular extracts, the majority of soluble Snu66p and Snu23p and a significant fraction of Spp381p is associated with [U4/U6·U5] tri-snRNP and [U2·U4/U6·U5] tetra-snRNP complexes, respectively.

Functional analysis of Snu66p, Snu23p and Spp381p

A possible role for Snu66p, Snu23p and Spp381p in pre-mRNA splicing was investigated initially *in vitro*. We thus performed splicing inhibition studies in splicing extracts containing one of the above-mentioned protA-tagged tri-snRNP proteins. For this, we added unspecific IgGs to the splicing assay, reasoning that binding of a bulky IgG molecule to the protA tag could interfere with the tagged protein's proper function. Interestingly, this led to a block of the first step of the splicing reaction with the extract containing Snu66-protAp (Figure 5, lanes 7 and 8). To verify that the splicing block was specific for Snu66p, a purified recombinant fusion protein of Snu66p and GST was added to the Snu66-protAp splicing extract that had been pre-incubated with IgGs to ensure splicing inhibition (Figure 5, lanes 9 and 10). Indeed, splicing was effectively restored by increasing amounts of exogenously added GST–Snu66p (Figure 5, compare lanes 6, 8 and 10). In contrast, GST–Snu66p had no effect on splicing in the absence of IgGs (lanes 5 and 11), nor did it affect splicing in the wild-type extract (lanes 4 and 5). Taken together, our data indicate that Snu66p plays an important role in pre-mRNA splicing *in vitro*. Binding of the antibody could inhibit the function of Snu66p or prevent its interaction with other spliceosomal components. Alternatively, the function of the entire tri-snRNP complex could be inhibited by steric hindrance due to antibody binding.

We also tested Spp381p and Snu23p by this approach, but detected only a minor inhibitory effect in the case of Snu23p (data not shown). Consistent with this, Lybarger *et al.* (1999) showed that Spp381p is dispensable for *in vitro* splicing and for cell viability. As an alternative to the *in vitro* approach, we examined whether genetic depletion of Snu23p and Spp381p would affect pre-mRNA splicing *in vivo*. We constructed inducible alleles of *SNU23*

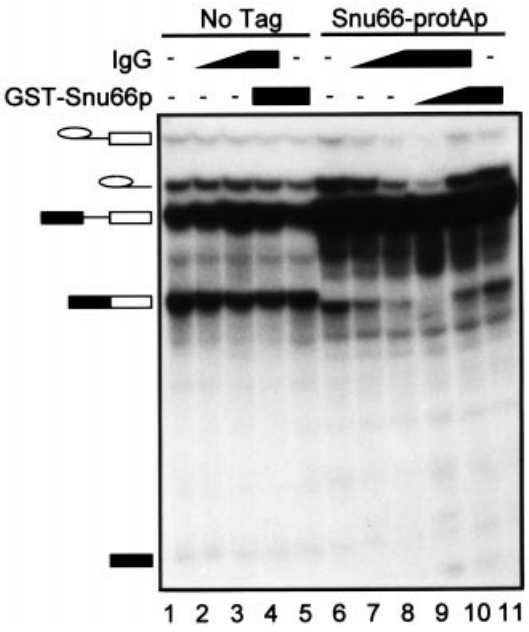

Fig. 5. Snu66p is required for splicing *in vitro*. Splicing reactions were performed using the Snu66-protAp extract (lanes 6–11) or a wild-type extract containing no tagged protein (lanes 1–5) and radiolabelled actin pre-mRNA. Increasing amounts (1.62 and 6.48 μM) of purified rabbit IgGs were added to the reactions (lanes 2–4 and 7–10). To restore splicing, recombinant GST–Snu66p was also added (lanes 4, 5 and 9–11) in two concentrations (0.16 and 0.48 μM). Low and high concentrations of antibodies and protein are indicated by bars above the lanes. Indicated on the left is the identity of the labelled RNA species (from top to bottom): intron–lariat–exon 2 intermediate, excised intron–lariat, pre-mRNA, mature mRNA and cleaved exon 1 intermediate.

and *SPP381* by replacing their natural promoters with the *GAL1* promoter. Transcription from this promoter is repressed when the cells are grown in the presence of glucose and induced in the presence of galactose. When we examined the splicing of pre-U3 RNA in these strains by primer extension, we observed an accumulation of pre-U3A and pre-U3B RNA after 27 and 22 h in glucose for *GAL1::SPP381* and *GAL1::SNU23* (Figure 6A and B), respectively. The accumulation of pre-U3 RNA was not as dramatic as observed with other splicing factors, such as Snu114p (Fabrizio *et al*., 1997), and it was detected only after long depletion times. Nevertheless, the data indicate that Spp381p and Snu23p are required for efficient splicing *in vivo*.

Snu66p and Snu23p are part of activated spliceosomes

The immunoprecipitation studies shown in Figure 3 revealed that Snu66p, Spp381p and, in particular, Snu23p were bound to snRNPs in a salt-sensitive manner. We therefore were interested in investigating whether the three proteins are integral components of the spliceosome or whether they interact only transiently with the splicing machinery. Initially, we performed standard splicing reactions using ^{32}P-labelled actin pre-mRNA with extracts containing protA-tagged Snu66p, Spp381p, Snu23p or, as a control, no tagged protein. After incubation of the reactions, we performed immunoprecipitations with IgG beads and the precipitated RNAs were analysed (Figure 7A). A significant co-precipitation of the unspliced pre-

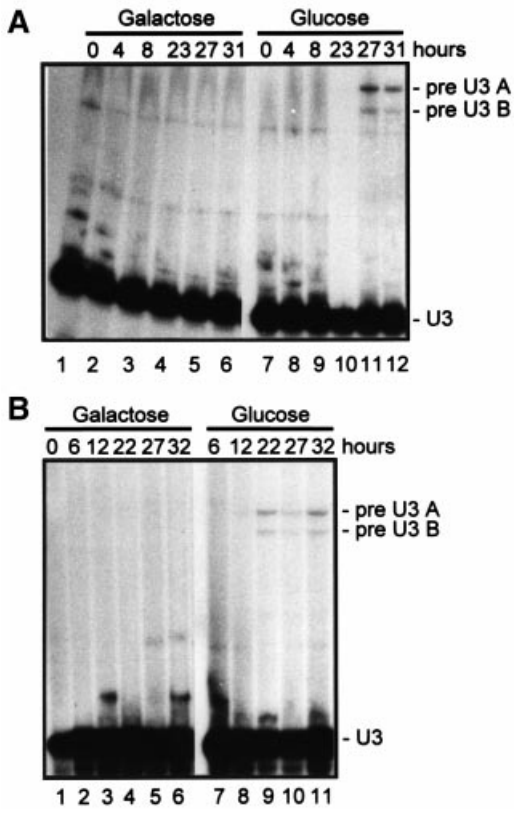

Fig. 6. Spp381p and Snu23p contribute to splicing *in vivo*. Inducible alleles of *SPP381* (**A**) and *SNU23* (**B**) were constructed by integration of a cassette containing the *HIS3* marker gene and the *GAL1* promotor at their 5′ end. The resulting strains were grown in galactose- or glucose-containing medium for the times indicated. Total RNA was prepared from the cells and assayed for the presence of pre-U3A, pre-U3B and mature U3 transcripts by primer extension using a ^{32}P-labelled primer annealing to exon 2 of the U3 RNA (Fabrizio *et al*., 1997). The identity of the reverse transcripts is indicated on the right.

mRNA was observed with Snu66-protAp and Snu23-protAp (compare lanes 4 and 8 with the background control in lane 2). In addition, low amounts of the lariat intermediate and the lariat were also co-precipitated with Snu66-protAp in a reproducible manner (Figure 7A, lane 4). The mature mRNA was not co-precipitated with any of the proteins. In contrast to Snu66-protAp and Snu23-protAp, precipitation of the pre-mRNA is less efficient with Spp381-protAp (lane 6).

In order to investigate the association of the three proteins with spliceosomes in more detail, we have purified spliceosomes which were assembled onto biotinylated pre-mRNA *in vitro* (Grabowski and Sharp, 1986). Splicing reactions were incubated with biotinylated or non-biotinylated substrate, using either wild-type extract or extracts containing the respective tagged protein. Spliceosomal complexes were then co-precipitated using streptavidin beads. The snRNAs and proteins present in the precipitates were assayed by Northern and Western blot analysis, respectively (Figure 7B). From all extracts, U1, U2, U5 and U6 snRNPs, but not U4 snRNP, were co-precipitated with the biotinylated substrate and not with the non-biotinylated pre-mRNA (Figure 7B, lower panel, lanes 2, 6, 8 and 10, compare with lanes 1, 5, 7 and 9 for negative controls). This snRNA composition of the precipitated splicing complexes indicates that all spliceosomes that have integrated the [U4/U6·U5] tri-snRNP have already undergone the activation step, i.e. the dissociation of U4 snRNP. However, as co-precipitation of U1 snRNPs with activated spliceosomes generally has been observed (e.g. Xie *et al*., 1998), we cannot exclude the possibility that some pre-spliceosomes containing U1 and U2 snRNPs are also present in the precipitate.

Significantly, the Western blot analysis of the precipitates demonstrated that Snu66-protAp and Snu23-protAp are part of activated spliceosomes, while Spp381-protAp was absent (Figure 7B, upper panel, compare lanes 6, 8 and 10 with lanes 5, 7 and 9 for negative controls). In sum, these data, together with the results of the immunoprecipitation experiments (Figure 7A), show that, while Snu23p acts at least until the U4 snRNP has left the spliceosome, Snu66p is present in spliceosomes even after the second step of splicing has occurred. In contrast, Spp381p is not present in purified activated spliceosomes (Figure 7B, lane 8). This suggests that the role of Spp381p may be confined to early spliceosomal assembly steps. Alternatively, it is possible that Spp381p is only loosely bound to the spliceosome and is removed from the complex under the experimental conditions used for its affinity purification. This possibility is very unlikely, however, since Spp381-protAp remains associated with the tri-snRNP and the [U2·U4/U6·U5] tetra-snRNP during immunoprecipitations performed at the same salt concentration as used for the spliceosome purification.

Rse1p is an integral component of the spliceosome

The association of Rse1p with the pre-mRNA and spliceosomal complexes was investigated in a similar way to that described above (Figure 7A and B). The results demonstrated that Rse1p is part of the spliceosome even after the first step of splicing has occurred, since it also associated with the lariat–exon 2 intermediate (Figure 7A, lane 10). Rse1p is the probable orthologue of the human SF3b subunit SAP130 (R.R.Reed, personal communication), which can be cross-linked to pre-mRNA and is critically involved in the correct positioning of the U2 snRNP at the branchpoint region of the intron (Gozani *et al*., 1996). Our findings that Rse1p is associated with the U2 snRNP and, further, that it is an integral part of the spliceosome, support the idea that Rse1p may have a similarly important function in the yeast spliceosome.

Conclusions

The recent advances in mass spectrometry combined with the availability of completely sequenced genomes has opened up unprecedented possibilities in the biochemical characterization of multifactorial protein and protein–nucleic acid complexes (Lamond and Mann, 1997). Here we have exploited these advances to characterize further the protein components of the 25S [U4/U6·U5] tri-snRNP from the yeast *Saccharomyces cerevisiae*. We identified all but one (Lsm3p, see Table I) of the known tri-snRNP proteins. Thus, taken together, the yeast [U4/U6·U5] tri-snRNP comprises at least 28 distinct proteins (Table I). These include the seven canonical Sm proteins, seven Sm-like proteins (Lsm2p–Lsm8p) and 14 tri-snRNP-specific proteins. Assuming that each protein contributes

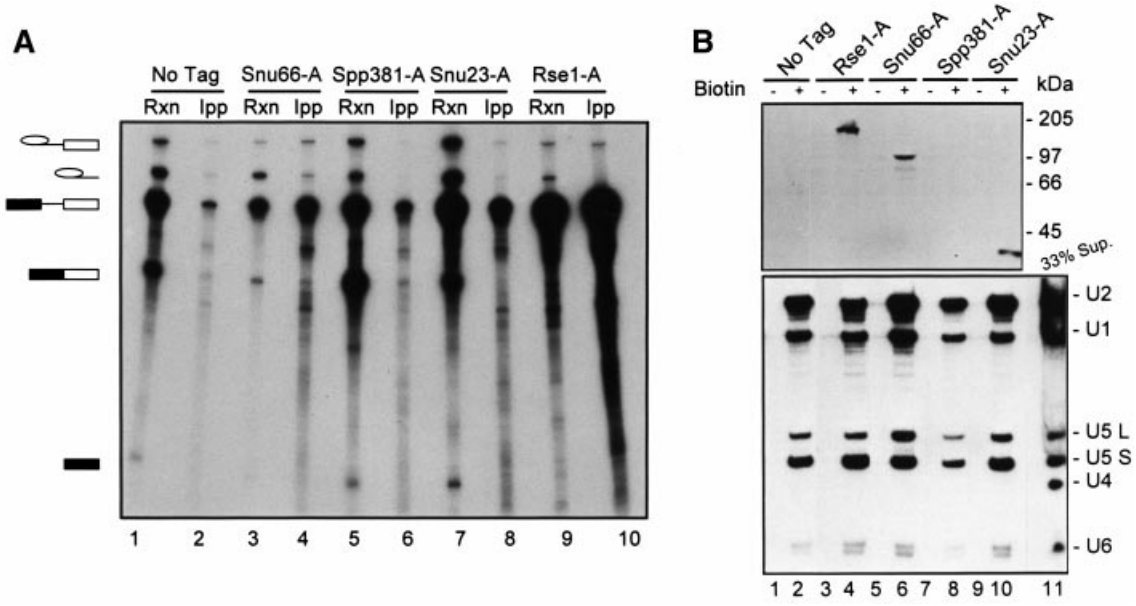

Fig. 7. Association of Snu66p, Spp381p, Snu23p and Rse1p with pre-mRNA during *in vitro* splicing and presence of the proteins in purified spliceosomes. (**A**) Splicing reactions using either extracts containing no tagged protein (lane 1) or protA-tagged Snu66p, Spp381p, Snu23p or Rse1p were incubated with radiolabelled actin pre-mRNA and precipitated with IgG–agarose. Each time, 10% of the total reaction ('Rxn'; odd lanes) and the precipitate from the remaining 90% of the reactions ('Ipp'; even lanes) were assayed for the presence of pre-mRNA, splicing intermediates and products (labelled as described in Figure 5). A part of the sample in lane 3 was lost upon loading the gel. (**B**) Splicing reactions were incubated with actin pre-mRNA that was transcribed either in the absence or presence of biotinylated UTP to incorporate an affinity label. The reactions were stopped on ice, precipitated with streptavidin beads and the washed precipitates were extracted to separate snRNAs and proteins. The snRNAs were assayed by Northern blot (lower panel), the proteins by Western blot analysis for the presence of protein A tags (upper panel). In lane 11, one-third of the supernatant resulting from the precipitation in lane 2 was loaded as a total snRNA control. Since the U6 snRNA probe used in the Northern blot was unusually weak, the U6 signal is poor especially in lane 8, but was clearly visible upon longer exposure of the film.

stoichiometrically, and taking into consideration that two sets of canonical Sm proteins are present (one set each with U4 and U5 snRNA), the predicted molecular weight of the yeast tri-snRNP is at least 1.5 MDa.

The eight novel tri-snRNP proteins identified in this work comprise the Sm-like proteins Lsm2p, Lsm5p, Lsm6p and Lsm7p, as well as four tri-snRNP-specific proteins, termed Snu13p, Dib1p/Snu16p, Snu23p and Snu66p (Table I). Of the nine Sm-like proteins encoded in the yeast genome (see Introduction), only seven (Lsm2p–Lsm8p) appear to be associated with the [U4/U6·U5] tri-snRNP. We recently identified orthologues of the yeast Lsm2p–Lsm8p proteins also in the human tri-snRNP complex (T.Achsel, H.Brahms, B.Kastner, A.Bachi, M.Wilm and R.Lürhmann, manuscript submitted), suggesting that these proteins may function in an evolutionarily conserved manner. Likewise, Snu13p and Dib1p/Snu16p proteins are highly conserved, and their functional counterparts are also present in the human [U4/U6·U5] tri-snRNP complex (S.Nottrott *et al*., submitted; K.Reuter *et al*., submitted).

Snu13p shares a conserved central region with several other RNP-associated proteins, including the yeast snoRNP protein Nhp2p and the ribosomal protein L32. This region previously has been postulated to contain the consensus sequence of a novel RNA recognition motif (Koonin *et al*., 1994). We recently have shown that the human orthologue of Snu13p, the 15.5 kDa tri-snRNP protein, binds directly to the 5′ stem–loop of U4 snRNA (S.Nottrott *et al*., submitted).

Dib1p/Snu16p provides an interesting link to the cell cycle in that a mutant of its *S.pombe* orthologue, Dim1p, has been shown previously to exert a cell cycle arrest before entry into mitosis (Berry and Gould, 1997; Berry *et al*., 1999). Our identification of Dib1p/Snu16p as a component of the yeast [U4/U6·U5] tri-snRNP raises the possibility that the cell cycle phenotype observed for the Dim1p mutant may be indirect, i.e. due to an inhibition of pre-mRNA splicing.

Detailed investigations of the remaining two novel tri-snRNP-specific proteins identified here, Snu66p and Snu23p, suggest possible roles for them in splicing. Thus, immunoinhibition and complementation studies, carried out *in vitro* using an extract from a yeast strain expressing protA-tagged Snu66p, indicate that Snu66p plays an important role in the spliceosome prior to the first step of splicing (Figure 5). In contrast, a yeast extract containing protA-tagged Snu23p was permissive to splicing *in vitro* in the presence of antibodies. However, we provide evidence that Snu23p appears to be required to ensure efficient pre-mRNA splicing *in vivo*, as the genetic depletion of Snu23p leads to the appearance of unspliced pre-U3 RNA (Figure 6B). Both proteins were shown to be stably integrated into spliceosomes assembled onto pre-mRNA substrates *in vitro* and to be retained in spliceosomes from which U4 snRNP had already been dissociated (Figure 7B). Snu66p remained associated even with the lariat–exon 2 intermediate and excised intron (Figure 7A). In contrast, the tri-snRNP protein Spp381p could not be identified in affinity-purified activated spliceosomes (Figure 7B), suggesting that Spp381p is either associated loosely with spliceosomes and thus lost during the purification step, or that it may function only during the early assembly steps of the spliceosome. The

latter idea is consistent with the observations of Lybarger *et al*. (1999) that Spp381p may act in concert with Prp38p, for which a role in the displacement of the U4 snRNP from the spliceosome has been assigned (Xie *et al*., 1998).

The co-immunoprecipitation of U2 and [U4/U6·U5] tri-snRNPs from yeast extracts by each of the tagged Snu23p, Spp381p and Snu66p proteins was unexpected and indicated that they are present in pre-formed [U2·U4/U6·U5] tetra-snRNP complexes. This notion was strongly corroborated by our findings that both U2 and tri-snRNPs were co-precipitated efficiently with Snu23-protAp and Spp381-protAp from the 30–35S region of a glycerol gradient on which yeast extracts had been pre-fractionated by centrifugation. The fact that the U2 snRNP is underrepresented in the Snu66-protAp immunoprecipitate when compared with those obtained with tagged Spp381p or Snu23p (Figure 3) does not contradict our hypothesis. Rather, the binding of an IgG to the tagged protein used for immunoprecipitation may differentially affect the stability of the [U2·U4/U6·U5] tetra-snRNP, depending on the location of this protein within the tetra-snRNP. Along the same lines, we have observed low but significant co-precipitation of U2 with tri-snRNPs using antibodies specific for either of the U5 proteins Prp8p and Snu114p (data not shown). Furthermore, low amounts of tri-snRNPs were also co-precipitated with the tagged U2 protein Rse1p (Figure 3).

Strong evidence for the existence of a yeast [U2·U4/U6·U5] tetra-snRNP complex was first reported by Raghunathan and Guthrie (1998). These authors observed significant co-precipitation of U2 snRNP with [U4/U6·U5] tri-snRNPs when they used the tagged U5 protein Brr2p/Snu246p/Rss1p/Slt22p for pull-down experiments. Evidence for a pre-mRNA-independent formation of multimeric snRNP complexes containing U2, U4, U6 and U5 snRNPs has been reported previously for human nuclear splicing extracts as well (Konarska and Sharp, 1988; Hall and Konarska, 1992; Wassarman and Steitz, 1992).

The question of whether a pre-formed [U2·U4/U6·U5] tetra-snRNP complex plays a role as an intermediate in the assembly pathway of the spliceosome cannot be answered at this point. It is interesting to speculate that the protein–protein and/or protein–RNA contacts mediating the interactions between U2 and the tri-snRNP within the tetra-snRNP may be, at least in part, also responsible for the integration of the tri-snRNP into the pre-spliceosome.

A comparison of the protein components of the yeast and human [U4/U6·U5] tri-snRNPs reveals a striking overlap of evolutionarily conserved proteins in both species. Thus, in addition to the Sm and Sm-like (Lsm) proteins, conserved functional counterparts of at least nine tri-snRNP-specific proteins are present in both tri-snRNP complexes. These proteins include Prp8p, Brr2p/Snu246p/Rss1p/Slt22p, Snu114p, Prp31p, Prp6p, Prp4p, Prp3p, Dib1p/Snu16p and Snu13p (see Will *et al*., 1995; Will and Lührmann, 1997, for a compilation of human tri-snRNP proteins; and our unpublished results). While counterparts of Snu23p and Snu66p in the human [U4/U6·U5] tri-snRNP have not yet been published, expressed sequence tags (ESTs) encoding evolutionarily conserved human homologues of these two yeast proteins exist in the database. It would not be surprising if in the near future these proteins are identified as spliceosomal components which perhaps only loosely interact with the human tri-snRNP particle. Precedents for evolutionarily and functionally conserved spliceosomal proteins exhibiting differential affinities for the tri-snRNP in yeast and man are Prp18p and Prp28p, which are co-isolated with the tri-snRNP in one system but not in the other (Horowitz and Abelson, 1993; Strauss and Guthrie, 1994; Horowitz and Krainer, 1997; Teigelkamp *et al*., 1997).

Nonetheless, a few proteins appear to be unique to a single system. Thus, Spp381p may be considered to be a yeast-specific tri-snRNP protein because neither a counterpart in the human [U4/U6·U5] tri-snRNP nor ESTs encoding human homologues have been identified so far. Along the same lines, the human tri-snRNP contains at least three proteins for which we did not identify clear orthologues in the yeast genome. These are the SnuCyp-20/USA-Cyp cyclophilin (Horowitz *et al*., 1997; Teigelkamp *et al*., 1998), the U5-specific 40 kDa protein (Achsel *et al*., 1998) and the 27 kDa tri-snRNP protein (Fetzer *et al*., 1997). While it is conceivable that the cyclophilin and the 40 kDa protein are replaced by distantly related proteins in the yeast spliceosome, the RS domain-containing 27 kDa protein is most likely a human-specific tri-snRNP protein.

The snRNAs of the yeast and human [U4/U6·U5] tri-snRNPs are evolutionarily conserved and undergo very similar, if not identical, conformational rearrangements during the spliceosomal cycle (see Introduction; Staley and Guthrie, 1998). These RNA conformational transitions are most likely to be facilitated and regulated mainly by snRNP proteins and splicing factors. In this respect, the overall evolutionary conservation of tri-snRNP proteins from yeast and man is significant. This indicates that the nature of the RNA–protein and protein–protein interactions in the two tri-snRNP complexes, as well as their dynamics during the spliceosomal cycle, are very similar. The identification of several novel components of the core of the spliceosome paves the way to elucidating the functions of the many proteins in this dynamic machine.

Materials and methods

Yeast strains, plasmids and oligonucleotides used in this study

The strains used were W303-Prp4-8xHis (this study); YPH499, TR2 (both Sikorski and Hieter, 1989); AGY12 (like TR2, *SNU66-protA-TRP1*; this study); AGY13 (like TR2, *SPP381-protA-TRP1*; this study); AGY14 (like TR2, *SNU23-protA-TRP1*; this study); AGY24 (like TR2, *RSE1-protA-TRP1*; this study); AGY22 (like YPH499, *HIS3/GAL1::SPP381*; this study); and AGY23 (like YPH499, *HIS3/GAL1::SNU23*; this study).

The plasmids employed were yCp-His/Prp4 (this study), pGEX-4T1 (Pharmacia), pGEX-4T1-SNU66 (this study) and pBS1173 (Puig *et al*., 1998). A plasmid containing a *HIS3/GAL1* cassette was a kind gift of Martin Funk.

The DNA oligonucleotides used were: 1, 5′-GGCGGGGATCCA-CCAAAACGGAAAACCTT-3′; 2, 5′-GGCGGCTCGAGTTAATTGT-CATCAAATTC-3′; 3, 5′-AAAACCAGAAGTTCCTACAAGCCTTAG-GCTGACAAAAATTCGAGAAAGCTAGGGCGAATTGGAGCTCCA-C-3′; 4, 5′-TGCTCAGCATACTCCTCTCTATCCCAAGTTCTTCTG-CCAAAATTCGACATCGGGGGATCCACTAGTTCTAGA-3′, 5, 5′-CCTAATCAGCAAATTCTA-3′; 6, 5′-AGGCCGGTATTGTTTCTC-3′; 7, 5′-AGGAAAAATACCCCTGAAAATGGTAGTTTGTTTGAAT-TTGATGACAATAAGCTGGAGCTCAAAAC-3′; 8, 5′-CACCGCAT-CACACACACA-3′; 9, 5′-TGGGGAATTTGCAGCCGG-3′; 10, 5′-CT-CTAGGCAAGAATAGTAGAGATTACGAAGAAACTGAATATTCGG-

TTATAAAGCTGGAGCTCAAAAC-3′; 11, 5′-GAAACCAAGTAAAGATAA-3′; 12, 5′-ACTGAAATTTTTGAATAT-3′; 13, 5′-AAAGTCAGTCAAATGATGGGATTTTCCAATTTTGCCACCTCTAAAAAAAAAGCTGGAGCTCAAAAC-3′; 14, 5′-TGATTTTACAGTAATAAT-3′; 15, 5′-CATGAAGTGCATCATAGA-3′; 16, 5′-GTGGAGGATATAATTχα-CAAACAATCAACGAAGTCAGAACAAATTACATGAAGCTGGAGCTCAAAAC-3′; 17, 5′-CTGGAATTGCT CGAGTAAAGAAATG-3′; 18, 5′-TCGAGTCATCATCATCATCATCACCATGGTTGA C-3′; 19, 5′-TCGAGTCAACCATGGTGATGATGATGATGATGATGAC-3′; 20, 5′- CCACCGCTAGATTAAGAA TTTTTTTTGGTTACTGATTTGAATCACATACGGGGCGATTGGAGCTCCAC-3′; and 21, 5′-TCACTGTCTGAGCTTGTGTCAAGTCTCCTCTTGAAATGTCTAAAACTCATCGGGGGATCCACTAGTTCTAGA-3′.

Construction of a yeast strain expressing His-tagged Prp4p

Plasmid yCp-His/Prp4 was constructed from plasmid yCp-prp4 (Ayadi *et al*., 1997). The stop codon was removed and a new *Xho*I site was introduced by *in vitro* mutagenesis at the 3′ end of the *PRP4* gene using oligo 17. A 8xHis linker (oligos 18 and 19 hybridized together) was cloned, in-frame, into the new *Xho*I site. The resulting plasmid, encoding a C-terminally 8xHis-tagged Prp4p protein, was able to complement a yeast strain derived from W303-1A (R.Rothstein) and carrying a chromosomal copy of the *PRP4* gene, disrupted with the *HIS3* marker gene at a single *Bcl*I restricition site, giving rise to strain W303-Prp4-8xHis.

Purification of tri-snRNPs and mass spectrometric identification of proteins

Purification of snRNPs was as described previously (Neubauer *et al*., 1997). Starting material was 680 g of W303 Prp4-8xHis cells. Total cell extracts were prepared by the liquid nitrogen method (Umen and Guthrie, 1995), using an RM 100 mortar-grinder (Retsch, D-42759 Haan, Germany). Chromatography was in buffer D containing 100 mM KCl, to allow efficient binding of the tri-snRNP to the anti-m_3G-cap column (U1 snRNP binds more efficiently at 200 mM KCl; see Fabrizio *et al*., 1994). A 1.7 mg aliquot of total snRNPs and 80–100 μg of purified tri-snRNP (plus some U1 and U2 snRNPs) were obtained after anti-m_3G-cap and Ni^{2+}-NTA chromatography, respectively. Gradient centrifugation was as described (Neubauer *et al*., 1997). Proteins and RNAs from the gradient fractions were separated by extraction with phenol/chloroform/isoamylalcohol, proteins were precipitated from the organic phase with acetone, and RNAs were precipitated from the aqueous phase with ethanol/sodium acetate. Proteins from the fractions containing the majority of U4, U5 and U6 snRNAs were pooled, run on a 60 cm long SDS–gel consisting of an upper part containing 11.5% and a lower part of 13% polyacrylamide and stained with Coomassie Brilliant Blue. Protein bands were excised, in-gel digested with trypsin and the resulting peptides extracted as described previously (Wilm *et al*., 1996). For most of the protein bands, both MALDI mass mapping and nanoelectrospray tandem mass spectrometry were carried out in order to ensure the identification of all the proteins present as mixtures in single bands. For MALDI mass mapping, the thin film technique was applied for target preparation as described previously (Vorm *et al*., 1994). The mass analysis was performed on a Bruker REFLEX MALDI time-of-flight mass spectrometer (Bruker-Franzen Analytik, Bremen, Germany) equipped with a delayed extraction ion source. The mass accuracy used for database searching was typically better than 70 p.p.m. For nanoelectrospray tandem mass spectrometry, the extracted peptides were desalted over a microcolumn (Neubauer and Mann, 1999) and the peptide mixture was eluted directly into the spraying needle of the nanoelectrospray ion source. The tandem mass measurements were carried out on a Sciex API III triple quadrupole mass spectrometer (Sciex, Perkin-Elmer, Ontario, Canada).

Tagging of proteins, immunoprecipitation of snRNPs from extracts, gradients or splicing reactions, and affinity purification of spliceosomes

Tagging of Snu66p, Spp381p, Snu23p and Rse1p with two copies of the IgG-binding domain of the *Staphylococcus aureus* protA (Lafontaine and Tollervey, 1996), using oligos 5–7, 8–10, 11–13 and 14–16 (as primers 'B1', 'B2' and 'B3'), respectively, was as described (Gottschalk *et al*., 1998; Puig *et al*., 1998). Expression of the protA tag was confirmed by Western blot analysis of cell lysates. Immunoprecipitation and analysis of snRNAs were essentially as described (Gottschalk *et al*., 1998), except that IgG–agarose was used instead of protein A–Sepharose coupled with antisera. Splicing extracts were prepared by the liquid nitrogen method (Umen and Guthrie, 1995). Immunoprecipitation from 10–30% glycerol gradient fractions was performed after separation of each 250 μl extract on 2.4 ml gradients containing 100 mM KCl and run for 3 h at 55 000 r.p.m. in a Beckman TL-100 ultracentrifuge. After fractionation of the gradient into 15 fractions of 160 μl, 120 μl of each fraction were used for immunoprecipitation. Proteins and RNAs were extracted from the precipitates as well as from the remaining 40 μl of each gradient fraction and analysed separately by Western and Northern blot analysis.

Standard splicing reactions containing 20 μl of extract were incubated with 17.5 fmol of radiolabelled actin precursor (100 000 c.p.m.) for 25 min at 23°C (Lin *et al*., 1985). A reaction mixture corresponding to 2 μl of extract was taken for analysis of splicing products. Splicing was then stopped on ice by the addition of 420 μl of buffer NET-2 100 (50 mM Tris–HCl pH 7.5, 100 mM NaCl, 0.05% NP-40) and mixed with 20 μl of IgG–agarose. Precipitation was for 1.5 h at 4°C. After washing, the precipitates as well as the total splicing mixtures were treated with proteinase K for 15 min at 37°C, RNAs were extracted by phenol/chloroform/isoamylacohol, ethanol precipitated and analysed by gel electrophoresis (8% gels) and autoradiography. For precipitation of spliceosomes, actin pre-mRNA was transcribed in the presence of biotin-16-UTP (Boehringer Mannheim) in a ratio of 20:1 UTP:biotin-16-UTP. Standard splicing reactions containing 300 μl of extract and 483 ng of either biotinylated or non-biotinylated actin pre-mRNA were precipitated after a 20 min incubation at 23°C by the addition of 40 μl of streptavidin–agarose in 750 μl of NET-2 containing 100 mM NaCl for 1.5 h at 4°C. Precipitates were washed extensively with NET-2 100, and RNAs and proteins were extracted from the beads and analysed separately by Northern and Western blot analysis. All splicing reactions were pre-incubated at 23°C for 15 min prior to the addition of precursor.

Cloning and expression of GST-Snu66p and splicing inhibition in vitro

The coding sequence of *SNU66* was amplified by PCR from yeast genomic DNA using oligos 1 and 2 and cloned into vector pGEX-4T1 (Pharmacia) after digestion with *Bam*HI and *Xho*I. The resulting plasmid (pGEX4T1-SNU66) was transformed into *E.coli* strain HB101 and the sequence was confirmed by automated DNA sequencing. pGEX4T1-SNU66 was then transformed into *E.coli* strain BL21 and expression and purification of GST–Snu66p was essentially following the supplier's instructions for the GST expression system (Pharmacia). *In vitro* splicing inhibition was assayed with standard splicing reactions that were pre-incubated with purified rabbit IgGs (1.62 and 6.48 μM) and/or purified GST–Snu66p (0.16 and 0.48 μM) for 15 min at 23°C.

Construction of inducible alleles of SPP381 and SNU23 and in vivo splicing inhibition

A plasmid containing a *HIS3/GAL1* cassette was amplified by PCR using two primer pairs (20/21 and 3/4) containing flanking sequences specific to the *SPP381* and *SNU23* genes, respectively. The resulting PCR products were integrated upstream of the coding sequences of each of *SPP381* and *SNU23* in strain YPH499 by homologous recombination, transformants were selected on selective medium lacking histidine and correct integration was assayed by PCR. *In vivo* accumulation of unspliced pre-U3A and pre-U3B RNA transcripts was assayed by primer extension as described previously (Fabrizio *et al*., 1997).

Acknowledgements

We thank Jean D.Beggs for kindly providing antibodies against Prp8p, David S.Horowitz and John Abelson for the Prp18p antiserum, Martin Funk for the *HIS3/GAL1* construct, and Bertrand Séraphin for plasmid pBS1173. We are grateful to Veronica Raker, Henning Urlaub, Klaus Hartmuth, Cindy L.Will and Hans-Peter Vornlocher for critical reading of the manuscript and fruitful discussions, and to Robin R.Reed, Jean D.Beggs, Bertrand Séraphin and Iain W.Mattaj for personal communication of results prior to publication. We also thank Dagmar Meyer for expert technical assistance. This work was supported by the Gottfried Wilhelm Leibniz Programme and grants from the Deutsche Forschungsgemeinschaft (SFB 397, A6/A7).

References

Abovich,N., Legrain,P. and Rosbash,M. (1990) The yeast *PRP6* gene encodes a U4/U6 small nuclear ribonucleoprotein particle (snRNP) protein and the *PRP9* gene encodes a protein required for U2 snRNP binding. *Mol. Cell. Biol.*, **10**, 6417–6425.

Achsel,T., Ahrens,K., Brahms,H., Teigelkamp,S. and Lührmann,R. (1998) The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2 and a novel WD-40 protein. *Mol. Cell. Biol.*, **18**, 6756–6766.

Anthony,J.G., Weidenhammer,E.M. and Woolford,J.L.,Jr (1997) The yeast Prp3 protein is a U4/U6 snRNP protein necessary for integrity of the U4/U6 snRNP and the U4/U6·U5 tri-snRNP. *RNA*, **3**, 1143–1152.

Ayadi,L., Miller,M. and Banroques,J. (1997) Mutations within the yeast U4/U6 snRNP protein Prp4 affect a late stage of spliceosome assembly. *RNA*, **3**, 197–209.

Banroques,J. and Abelson,J.N. (1989) PRP4: a protein of the yeast U4/U6 small nuclear ribonucleoprotein particle. *Mol. Cell. Biol.*, **9**, 3710–3719.

Berry,L.D. and Gould,K.L. (1997) Fission yeast *dim1*(+) encodes a functionally conserved polypeptide essential for mitosis. *J. Cell. Biol.*, **137**, 1337–1354.

Berry,L.D., Feoktistova,A., Wright,M.D. and Gould,K.L. (1999) The *Schizosaccharomyces pombe dim1*(+) gene interacts with the anaphase-promoting complex or cyclosome (APC/C) component lid1(+) and is required for APC/C function. *Mol. Cell. Biol.*, **19**, 2535–2546.

Bjorn,S.P., Soltyk,A., Beggs,J.D. and Friesen,J.D. (1989) *PRP4* (*RNA4*) from *Saccharomyces cerevisiae*: its gene product is associated with the U4/U6 small nuclear ribonucleoprotein particle. *Mol. Cell. Biol.*, **9**, 3698–3709.

Blanton,S., Srinivasan,A. and Rymond,B.C. (1992) *PRP38* encodes a yeast protein required for pre-mRNA splicing and maintenance of stable U6 small nuclear RNA levels. *Mol. Cell. Biol.*, **12**, 3939–3947.

Boeck,R., Lapeyre,B., Brown,C.E. and Sachs,A.B. (1998) Capped mRNA degradation intermediates accumulate in the yeast *spb8-2* mutant. *Mol. Cell. Biol.*, **18**, 5062–5072.

Bordonné,R. and Tarassov,I. (1996) The yeast *SME1* gene encodes the homologue of the human E core protein. *Gene*, **176**, 111–117.

Burge,C.B., Tuschl,T. and Sharp,P.A. (1999) Splicing of precursors to mRNAs by the spliceosome. In Gesteland,R.F., Chech,T.R. and Atkins,J.F. (eds), *The RNA World II*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 525–560.

Camasses,A., Bragado-Nilsson,E., Martin,R., Séraphin,B. and Bordonné,R. (1998) Interactions within the yeast Sm core complex: from proteins to amino acids. *Mol. Cell. Biol.*, **18**, 1956–1966.

Caspary,F., Shevchenko,A., Wilm,M. and Séraphin,B. (1999) Partial purification of the yeast U2 snRNP reveals a novel yeast pre-mRNA splicing factor required for pre-spliceosome assembly. *EMBO J.*, **18**, 3463–3474.

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.

Cooper,M., Johnston,L.H. and Beggs,J.D. (1995) Identification and characterization of Uss1p (Sdb23p): a novel U6 snRNA-associated protein with significant similarity to core proteins of small nuclear ribonucleoproteins. *EMBO J.*, **14**, 2066–2075.

Fabrizio,P., Esser,S., Kastner,B. and Lührmann,R. (1994) Isolation of *S.cerevisiae* snRNPs: comparison of U1 and U4/U6·U5 to their human counterparts. *Science*, **264**, 261–265.

Fabrizio,P., Laggerbauer,B., Lauber,J., Lane,W.S. and Lührmann,R. (1997) An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. *EMBO J.*, **16**, 4092–4106.

Fetzer,S., Lauber,J., Will,C.L. and Lührmann,R. (1997) The [U4/U6·U5] tri-snRNP-specific 27K protein is a novel SR protein that can be phosphorylated by the snRNP-associated protein kinase. *RNA*, **3**, 344–355.

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.

Fury,M.G., Zhang,W., Christodoulopoulos,I. and Zieve,G.W. (1997) Multiple protein:protein interactions between the snRNP common core proteins. *Exp. Cell Res.*, **237**, 63–69.

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.

Grabowski,P.J. and Sharp,P.A. (1986) Affinity chromatography of splicing complexes: U2, U5 and U4 + U6 small nuclear ribonucleoprotein particles in the spliceosome. *Science*, **233**, 1294–1299.

Hall,K.B. and Konarska,M.M. (1992) The 5′ splice site consensus RNA oligonucleotide induces assembly of U2/U4/U5/U6 small nuclear ribonucleoprotein complexes. *Proc. Natl Acad. Sci. USA*, **89**, 10969–10973.

Henras,A., Henry,Y., Bousquet-Antonelli,C., Noaillac-Depeyre,J., Gelugne,J.P. and Caizergues-Ferrer,M. (1998) Nhp2p and Nop10p are essential for the function of H/ACA snoRNPs. *EMBO J.*, **17**, 7078–7090.

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.

Horowitz,D.S. and Krainer,A.R. (1997) A human protein required for the second step of pre-mRNA splicing is functionally related to a yeast splicing factor. *Genes Dev.*, **11**, 139–151.

Horowitz,D.S., Kobayashi,R. and Krainer,A.R. (1997) A new cyclophilin and the human homologues of yeast Prp3 and Prp4 form a complex associated with U4/U6 snRNPs. *RNA*, **3**, 1374–1387.

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.

Kambach,C., Walke,S., Young,R., Avis,J.M., de la Fortelle,E., Raker,V.A., Lührmann,R., Li,J. and Nagai,K. (1999) Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell*, **96**, 375–387.

Konarska,M.M. and Sharp,P.A. (1988) Association of U2, U4, U5 and U6 small nuclear ribonucleoproteins in a spliceosome-type complex in absence of precursor RNA. *Proc. Natl Acad. Sci. USA*, **85**, 5459–5462.

Koonin,E.V., Bork,P. and Sander,C. (1994) A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme? *Nucleic Acids Res.*, **22**, 2166–2167.

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

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.

Lamond,A. and Mann,M. (1997) Cell biology and the genome projects—a concerted strategy for characterizing multiprotein complexes by using mass spectrometry. *Trends Cell Biol.*, **7**, 139–142.

Lauber,J., Fabrizio,P., Teigelkamp,S., Lane,W.S., Hartmann,E. and Lührmann,R. (1996) The HeLa 200 kDa U5 snRNP-specific protein and its homologue in *Saccharomyces cerevisiae* are members of the DEXH-box protein family of putative RNA helicases. *EMBO J.*, **15**, 4001–4015.

Lin,J. and Rossi,J.J. (1996) Identification and characterization of yeast mutants that overcome an experimentally introduced block to splicing at the 3′ splice site. *RNA*, **2**, 835–848.

Lin,R.J., Newman,A.J., Cheng,S.C. and Abelson,J. (1985) Yeast mRNA splicing *in vitro*. *J. Biol. Chem.*, **260**, 14780–14792.

Lossky,M., Anderson,G.J., Jackson,S.P. and Beggs,J. (1987) Identification of a yeast snRNP protein and detection of snRNP–snRNP interactions. *Cell*, **51**, 1019–1026.

Lybarger,S., Beickman,K., Brown,V., Dembla-Rajpal,N., Morey,K., Seipelt,R. and Rymond,B.C. (1999) Elevated levels of a U4/U6·U5 snRNP-associated protein, Spp381p, rescue a mutant defective in spliceosome maturation. *Mol. Cell. Biol.*, **19**, 577–584.

Mann,M. and Wilm,M. (1994) Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal. Chem.*, **66**, 4390–4399.

Mayes,A.E., Verdone,L., Legrain,P. and Beggs,J.D. (1999) Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J.*, **18**, 4321–4331.

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

Neubauer,G. and Mann,M. (1999) Mapping of phosphorylation sites of gel-isolated proteins by nanoelectrospray tandem mass spectrometry: potentials and limitations. *Anal. Chem.*, **71**, 235–242.

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.
Nilsen,T.W. (1998) RNA–RNA interactions in nuclear pre-mRNA splicing. In Simons,R.W. and Grunberg-Manago,M. (eds), *RNA Structure and Function*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 279–307.
Noble,S.M. and Guthrie,C. (1996) Identification of novel genes required for yeast pre-mRNA splicing by means of cold-sensitive mutations. *Genetics*, **143**, 67–80.
Pannone,B.K., Xue,D. and Wolin,S.L. (1998) A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J.*, **17**, 7442–7453.
Puig,O., Rutz,B., Luukkonen,B.G., 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.
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.
Raker,V.A., Plessel,G. and Lührmann,R. (1996) The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle *in vitro*. *EMBO J.*, **15**, 2256–2269.
Roy,J., Zheng,B., Rymond,B.C. and Woolford,J.L.,Jr (1995) Structurally related but functionally distinct yeast Sm D core small nuclear ribonucleoprotein particle proteins. *Mol. Cell. Biol.*, **15**, 445–455.
Rymond,B.C. (1993) Convergent transcripts of the yeast *PRP38-SMD1* locus encode two essential splicing factors, including the D1 core polypeptide of small nuclear ribonucleoprotein particles. *Proc. Natl Acad. Sci. USA*, **90**, 848–852.
Rymond,B.C. and Rosbash,M. (1992) Yeast pre-mRNA splicing. In Jones,E.W., Pringle,J.R. and Broach,J.R. (eds), *The Molecular and Cellular Biology of the Yeast Saccharomyces*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, Vol. 2, pp. 143–192.
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.
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.
Shevchenko,A., Jensen,O.N., Podtelejnikov,A.V., Sagliocco,F., Wilm,M., Vorm,O., Mortensen,P., Boucherie,H. and Mann,M. (1996) Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl Acad. Sci. USA*, **93**, 14440–14445.
Sikorski,R.S. and Hieter,P. (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics*, **122**, 19–27.
Staley,J.P. and Guthrie,C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs and things. *Cell*, **92**, 315–326.
Stevens,S.W. and Abelson,J. (1999) Purification of the yeast U4/U6·U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proc. Natl Acad. Sci. USA*, **96**, 7226–7231.
Strauss,E.J. and Guthrie,C. (1994) PRP28, a 'DEAD-box' protein, is required for the first step of mRNA splicing *in vitro*. *Nucleic Acids Res.*, **22**, 3187–3193.
Teigelkamp,S., Mundt,C., Achsel,T., Will,C.L. and Lührmann,R. (1997) The human U5 snRNP-specific 100-kD protein is an RS domain-containing, putative RNA helicase with significant homology to the yeast splicing factor Prp28p. *RNA*, **3**, 1313–1326.
Teigelkamp,S., Achsel,T., Mundt,C., Göthel,S.F., Cronshagen,U., Lane,W.S., Marahiel,M. and Lührmann,R. (1998) The 20kD protein of human [U4/U6·U5] tri-snRNPs is a novel cyclophilin that forms a complex with the U4/U6-specific 60kD and 90kD proteins. *RNA*, **4**, 127–141.
Umen,J.G. and Guthrie,C. (1995) A novel role for a U5 snRNP protein in 3′ splice site selection. *Genes Dev.*, **9**, 855–868.
Vorm,O., Roepstorff,P. and Mann,M. (1994) Improved resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. *Anal. Chem.*, **66**, 3281–3287.
Wassarman,D.A. and Steitz,J.A. (1992) Interactions of small nuclear RNAs with precursor messenger RNA during *in vitro* splicing. *Science*, **257**, 1918–1925.
Watkins,N.J., Gottschalk,A., Neubauer,G., Kastner,B., Fabrizio,P., Mann,M. and Lührmann,R. (1998) Cbf5p, a potential pseudouridine synthase and Nhp2p, a putative RNA- binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA*, **4**, 1549–1568.
Weidenhammer,E.M., Ruiz-Noriega,M. and Woolford,J.L.,Jr (1997) Prp31p promotes the association of the U4/U6×U5 tri-snRNP with prespliceosomes to form spliceosomes in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **17**, 3580–3588.
Will,C.L. and Lührmann,R. (1997) Protein functions in pre-mRNA splicing. *Curr. Opin. Cell Biol.*, **9**, 320–328.
Will,C.L., Fabrizio,P. and Lührmann,R. (1995) Nuclear pre-mRNA splicing. In Eckstein,F. and Lilley,D.M.J. (eds), *Nucleic Acids and Molecular Biology*. Springer-Verlag, Berlin, Vol. 9, pp. 342–372.
Wilm,M., Shevchenko,A., Houthave,T., Breit,S., Schweigerer,L., Fotsis,T. and Mann,M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano electrospray mass spectrometry. *Nature*, **379**, 466–469.
Wittmann,H.G. (1980) Structure of ribosomal proteins. In Chambliss,G., Craven,G.R., Davies,J., Davis,K., Kahan,L. and Nomura,M. (eds), *Ribosomes. Structure, Function and Genetics*. University Park Press, Baltimore, MD, pp. 51–88.
Xie,J., Beickman,K., Otte,E. and Rymond,B.C. (1998) Progression through the spliceosome cycle requires Prp38p function for U4/U6 snRNA dissociation. *EMBO J.*, **17**, 2938–2946.
Xu,D., Nouraini,S., Field,D., Tang,S.J. and Friesen,J.D. (1996) An RNA-dependent ATPase associated with U2/U6 snRNAs in pre-mRNA splicing. *Nature*, **381**, 709–713.

Received May 17, 1999; revised and accepted June 23, 1999

Noted added in proof

We would also like to mention that while this work was at proof stage, a paper was published by Stevens and Abelson (1999), who have also studied the yeast tri-snRNP by biochemical and mass spectrometric methods and have essentially identified the same proteins.