Characterization of Sm-like proteins in yeast and their association with U6 snRNA

Andrew E. Mayes, Loredana Verdone,
Pierre Legrain¹ and Jean D. Beggs²

Institute of Cell and Molecular Biology, University of Edinburgh,
King's Buildings, Mayfield Road, Edinburgh EH9 3JR, UK and
¹Laboratoire du Métabolisme des ARN, Institut Pasteur, 25–28 rue du
Dr Roux, 75724, Paris Cedex 15, France

²Corresponding author
e-mail: jbeggs@ed.ac.uk

Seven Sm proteins associate with U1, U2, U4 and U5 spliceosomal snRNAs and influence snRNP biogenesis. Here we describe a novel set of Sm-like (Lsm) proteins in *Saccharomyces cerevisiae* that interact with each other and with U6 snRNA. Seven Lsm proteins co-immunoprecipitate with the previously characterized Lsm4p (Uss1p) and interact with each other in two-hybrid analyses. Free U6 and U4/U6 duplexed RNAs co-immunoprecipitate with seven of the Lsm proteins that are essential for the stable accumulation of U6 snRNA. Analyses of U4/U6 di-snRNPs and U4/U6·U5 tri-snRNPs in Lsm-depleted strains suggest that Lsm proteins may play a role in facilitating conformational rearrangements of the U6 snRNP in the association–dissociation cycle of spliceosome complexes. Thus, Lsm proteins form a complex that differs from the canonical Sm complex in its RNA association(s) and function. We discuss the possible existence and functions of alternative Lsm complexes, including the likelihood that they are involved in processes other than pre-mRNA splicing.

Keywords: Lsm/Sm proteins/snRNP/splicing/U6 snRNP

Introduction

Pre-mRNA splicing in eukaryotes occurs in the spliceosome, a multimeric ribonucleoprotein (RNP) complex composed of five snRNAs (U1, U2, U4, U5 and U6) and a large number of proteins. The snRNAs, in the form of RNA–protein complexes (snRNPs), assemble on the substrate pre-mRNA in an ordered manner (Moore *et al.*, 1993). The U1 and U2 snRNPs bind at the 5' splice site and branch point of the intron, respectively, and then the U4, U5 and U6 snRNAs are added in the form of a pre-assembled U4/U6·U5 tri-snRNP complex. The U4 and U6 snRNAs share extensive sequence complementarity and are found base paired in a U4/U6 duplex particle, although an independent U6 particle also exists, as U6 is more abundant than U4. In the spliceosome, a number of important RNA conformational changes occur, the most dramatic of which is destabilization of the U4/U6 duplex, which frees U6 to interact with U2 snRNA prior to initiation of the splicing reaction. Thus, splicing complexes are highly dynamic structures, and proteins are believed

to play critical regulatory roles in splicing complex assembly and turnover (Staley and Guthrie, 1998).

The snRNP proteins are considered to fall into two classes, the Sm (or core) proteins, which are associated with the U1, U2, U4 and U5 snRNAs, and the snRNP-specific proteins each of which associates with only one snRNP species. There are also many non-snRNP proteins which interact with the spliceosome during spliceosome assembly, the splicing reactions or spliceosome disassembly (reviewed in Moore *et al.*, 1993; Krämer, 1996).

The human Sm proteins cross-react with antisera from patients suffering from the autoimmune disorder systemic lupus erythematosus and were named B/B', D1, D2, D3, E, F and G based on their relative mobility in SDS–polyacrylamide gels (Lerner and Steitz, 1979; van Venrooij, 1987). These proteins play important roles in the biogenesis of the snRNP particles, associating in the cytosol with a structurally conserved region known as the Sm site (Branlant *et al.*, 1982) in the newly transcribed and exported U1, U2, U4 and U5 snRNAs (Mattaj and DeRobertis, 1985). This acts as a signal for the hypermethylation of the 5' cap of these snRNAs (Mattaj, 1986) which, together with the Sm proteins, forms a bipartite nuclear localization signal for the snRNP (Fischer and Lührmann, 1990; Hamm *et al.*, 1990). Maturation of the snRNP then continues in the nucleus with the addition of the snRNP-specific proteins (Zieve and Sauterer, 1990).

Saccharomyces cerevisiae possesses a homologous set of seven Sm proteins (Smb, Smd1, Smd2, etc.; Neubauer *et al.*, 1997; Gottschalk *et al.*, 1998). In addition to sharing sequence identities with the human proteins, the yeast Sm proteins have similar properties to their human counterparts; indeed the human SmD1 and SmE proteins can functionally complement null alleles of the respective yeast genes (Rymond *et al.*, 1993; Bordonné and Tarassov, 1996). However, very little is known about the pathway of snRNP biogenesis in yeast.

Sequence comparisons of the Sm proteins from a range of species led to the identification of a conserved motif, the Sm or snRNP core protein motif (Cooper *et al.*, 1995; Hermann *et al.*, 1995; Séraphin, 1995). The motif is composed of two conserved blocks of amino acids (32 and 14 residues) separated by a non-conserved spacer region of variable length. Truncation of the Sm motif of either human SmB' or SmD3 prevents these proteins forming a complex, suggesting that both conserved regions are required for intermolecular interactions (Hermann *et al.*, 1995; Camasses *et al.*, 1998).

In addition to these canonical Sm proteins, other sequences containing the Sm motif have been identified in *S. cerevisiae*. Uss1p, which was identified genetically and was characterized biochemically as a novel splicing factor (Cooper *et al.*, 1995), and Smx4p, which was

Table I. *Saccharomyces cerevisiae* Sm-like proteins and their genes

Gene name ^a	ORF name ^a	Other names	Sm protein sub-group ^b	Mol. wt (kDa) ^c
<i>LSM1</i>	YJL124c	<i>SPB8</i> ^d	SmB	20.3
<i>LSM2</i>	YBL026w	<i>SMX5</i> ^e , <i>SNP3</i>	SmD1	11.2
<i>LSM3</i>	YLR438c-A	<i>USS2</i> , <i>SMX4</i> ^e	SmD2	10.0
<i>LSM4</i>	YER112w	<i>USS1</i> ^f , <i>SDB23</i> ^g	SmD3	21.3
<i>LSM5</i>	YER146w		SmE	10.4
<i>LSM6</i>	YDR378c		SmF	13.8
<i>LSM7</i>	YNL147w		SmG	12.1
<i>LSM8</i>	YJR022w			14.5
	YCR020c-A	<i>SMX1</i> ^e		9.7

^aAs available via the SGD database (<http://genome-www.stanford.edu/Saccharomyces/>).

^bAs determined by Fromont-Racine *et al.* (1997).

^cAs calculated from the amino acid sequence.

^dBoeck *et al.* (1998); ^eSraphin (1995); ^fCooper *et al.* (1995); ^gParkes and Johnston (1992).

recognized by its genomic sequence as containing an Sm motif (Sraphin, 1995), were reported to associate primarily with free U6 and U4/U6 particles. Another of the Sm-like proteins, Spb8p, has been proposed to play a role in the decapping of mRNAs (Boeck *et al.*, 1998).

Searches of the *S.cerevisiae* genome database identified other open reading frames (ORFs) encoding putative Sm-like proteins. Four of these proteins, together with Uss1p, Smx4p and Spb8p, were grouped into seven sub-families with the human and yeast canonical Sm proteins on the basis of sequence similarity (Fromont-Racine *et al.*, 1997). To simplify the nomenclature, it has been decided to name (or rename in some cases) the genes encoding these Sm-like proteins *LSM* (Like *Sm*; Table I). Another two hypothetical proteins, Lsm8p (Yjr022p) and Smx1p (Ycr020-Ap), contain the Sm motif but are not structurally similar to any particular sub-family in the sequence alignment. Nevertheless, two-hybrid screens with Lsm8p as bait identified interactions with several of the Lsm proteins (Fromont-Racine *et al.*, 1997; unpublished data), and Lsm8p was itself identified in a two-hybrid screen with Hsh49p, the yeast homologue of the human splicing factor SAP49 (Fromont-Racine *et al.*, 1997). In contrast, exhaustive two-hybrid screens with Smx1p as bait did not produce any convincing interactions (P.Legrain and M.Fromont-Racine, unpublished data), and a protein A-tagged Smx1p showed no interaction with the spliceosomal snRNAs (Sraphin, 1995).

Thus structural similarity to the Sm proteins together with the two-hybrid data suggest that some of the Lsm proteins might also form a complex, and the inclusion of Lsm3p (Smx4p) and Lsm4p (Uss1p) in this set may indicate a U6 snRNA association. Superficially at least, the intrinsic differences between U6 and the other spliceosomal snRNAs would suggest that a U6-associated complex might have a fundamentally different role from that of the canonical Sm core complex. The sequence of U6 snRNA is the most highly conserved of all the spliceosomal snRNAs (Brow and Guthrie, 1988), but it lacks a recognized Sm site, and does not associate with the canonical Sm proteins. Unlike the other spliceosomal snRNAs which are products of RNA polymerase II, U6 is produced by RNA polymerase III. Since U6 snRNA

has a γ -methyl triphosphate cap structure which is not hypermethylated, and is thought to be largely retained in the nucleus (at least in higher eukaryotes where this has been studied; reviewed in Reddy and Busch, 1988), neither of the two primary roles defined for the Sm core complex (signals for hypermethylation and nuclear import) is generally considered to be required for the biogenesis of the U6 snRNP.

We present here an analysis of the molecular interactions of the yeast Lsm proteins. We show that their genes are required for normal growth, and that seven are required for the maintenance of normal levels of U6 snRNA. We provide evidence that these seven proteins form a complex, are associated with U6 snRNA and influence the efficiency of pre-mRNA splicing through effects on various U6 snRNA-containing complexes in the spliceosome assembly–dissociation cycle.

Results

Two-hybrid interactions between Lsm proteins

A two-hybrid direct mating approach was used to investigate all potential pairwise interactions between the Lsm proteins 1–8. Smx1p was not analysed, as the primary goal was to investigate a potential role in pre-mRNA splicing, and the evidence available suggested no involvement of Smx1p with the splicing machinery (see above). Several canonical Sm protein-encoding ORFs and fragments of Lsm4p were included as controls. Bait and prey fusions in haploid yeasts were combined by mating in all pairwise combinations, with resistance to defined levels of 3-aminotriazole (3-AT) used as a guide to the strength of any interaction seen. As variance between the expression levels and stability of different bait and prey proteins may affect these measurements, they are only a rough guide to the strength of interactions between different pairs of proteins. However, the number of diploid strains capable of surviving on medium containing up to 50 mM 3-AT does suggest a number of strong, and specific two-hybrid interactions (Figure 1). Other than an interaction between the Lsm4 and Smb fusions (only in the absence of 3-AT), the Lsm proteins did not interact with the canonical Sm proteins, whereas the Smb and Smd3 fusions interacted as expected. Instead, the Lsm proteins appeared to form many interactions with each other. Also, no significantly strong homotypic interactions were seen. Thus, those diploids which were able to survive on 3-AT are believed to represent specific associations between the two-hybrid fusion proteins and not aspecific interactions between any two proteins containing the Sm fold. It should be noted, however, that some of the observed interactions may be indirect, as discussed below.

Lsm4p is complexed with each of the other seven Lsm proteins

To facilitate more direct analyses of Lsm protein interactions, functional haemagglutinin (HA)-tagged versions of the proteins were constructed (Materials and methods; Table II). When anti-Lsm4p antibodies (Cooper *et al.*, 1995) were used to immunoprecipitate Lsm4p from extracts containing HA-tagged proteins, each of the other seven HA-tagged Lsm proteins was co-immunoprecipitated, whereas a control HA-tagged protein (Gal4-AD)

Table II. *Saccharomyces cerevisiae* strains

Strain	Genotype	Source
BMA38 ^a	<i>MATa/α trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100</i>	B.Dujon
BMA64 ^a	<i>MATa/α trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100</i>	F.Lacroute
JDY6 ^a	<i>MATa/α trp1Δ99, his3Δ200, ura3Δ99, leu2Δ1, ade2-101, cir^o</i>	J.D.Brown
BMA38α	<i>MATα trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100</i>	this work
BMA64α	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100</i>	this work
AEMY19 ^b	<i>MATα trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100 lsm6Δ::HIS3</i>	this work
AEMY22 ^b	<i>MATα trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100 lsm7Δ::HIS3</i>	this work
AEMY24 ^b	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100 lsm1Δ::TRP1</i>	this work
AEMY28 ^c	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100 HA:LSM1</i>	this work
AEMY29 ^c	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100 lsm5Δ::TRP1 [pACT11st-LSM5]</i>	this work
	<i>lsm3Δ::TRP1 [pBM125-HA:LSM3]</i>	
AEMY31 ^c	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100</i>	this work
AEMY32	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100 [pBM125-HA:LSM3]</i>	this work
AEMY33 ^c	<i>MATα trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100 lsm2Δ::HIS3 [pBM125-LSM2-HA]</i>	this work
AEMY34 ^c	<i>AEMY19 [YCpIF16-LSM6]</i>	this work
AEMY35 ^c	<i>AEMY22 [YCpIF16-LSM7-(intronless)]</i>	this work
AEMY46	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100 lsm8Δ::TRP1 [pBM125-HA:LSM8]</i>	this work
AEMY47	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100 lsm5Δ::TRP1 [pBM125-HA:LSM5]</i>	this work
LMA4-2A ^c	<i>MATα trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100 lsm8Δ::TRP1 [pACT11st-LSM8]</i>	M.Fromont-Racine
MCY4	<i>MATa ade1-101, his3-Δ1, trp1-289, ura3-52, LEU2-GAL1-LSM4</i>	M.Cooper
CG1945	<i>MATa gal4-542, gal80-538, ade2-101, his3Δ200, trp1-901, leu2-3,-112, ura3-52, lys2-801 URA3::</i> <i>(GAL4 17mers)₃-CYC1-lacZlys2::GAL1UAS-HIS3 cyh^r2</i>	Clontech
L40	<i>MATa ade2, gal4-542, gal80-538, trp1-901, leu2-3,-112, his3Δ200, lys2-801am, URA3:: (lexAop)₈-LacZ,</i> <i>lys2:: (lexAop)₄-HIS3</i>	Hollenberg <i>et al.</i> (1995)
Y187	<i>MATα gal4-542, gal80-538, ade2-101, his3Δ200, trp1-901, leu2-3,-112, ura3-52, URA3::GAL1-lacZ</i>	Clontech

^aDiploid strains are isogenic for the auxotrophic markers described.

^bTemperature-sensitive strains.

^cHA epitope tag strains used for immunoprecipitations.

The *lsm* deletion strains carrying the *HIS3* or *TRP1* markers are derived from BMA38 and BMA64, respectively.

Fig. 1. Two-hybrid direct matings of Lsm proteins. Haploid strains expressing the bait or prey fusions indicated were mated on YPDA medium and the growth of diploid cells was assayed by replica-plating to selective medium containing the concentrations indicated of 3-AT. The figure shows the highest 3-AT concentration at which each diploid displayed growth after 3 days at 30°C. Lsm4Δp, amino acids 1–92 of Lsm4p; Lsm4-A, amino acids 1–74 of Lsm4p; Lsm4-B, amino acids 46–86 of Lsm4p.

was not (Figure 2). Thus, each Lsm protein can associate in a complex, or complexes, with the Lsm4 protein. The exact composition of such a complex(es) cannot be determined in this experiment, nor whether all the Lsm proteins can associate simultaneously with Lsm4p.

Characterization of the *LSM* genes

Of the genes encoding Lsm proteins in *S.cerevisiae*, only *LSM4* (*USS1*) has been characterized extensively (Cooper *et al.*, 1995). In order to characterize the other *LSM* ORFs, targeted gene deletions were carried out (Figure 3A; Materials and methods). Three of the genes, *LSM1*, *LSM6* and *LSM7*, were found to be dispensable for cell viability; however, the deletions caused slow growth at 23 and 30°C, and failure to grow at 37°C (Figure 3B and data not shown). The remaining four genes, *LSM2*, *LSM3*, *LSM5* and *LSM8*, were essential for cell viability and, for these genes, a *GAL1*-regulated copy of the ORF (*P_{GAL1}-LSMX*) permitted growth with galactose but not with glucose as sole carbon source (Figure 3B and data not shown). Thus, for each *LSM* gene, a strain displaying a conditional growth phenotype was constructed, either temperature-sensitive (referred to below simply as *lsm* strains; see Table II) or carbon source-regulated. Representative growth curves are shown in Figure 3B.

Analysis of pre-mRNA splicing in the conditional strains

In order to determine whether the *in vivo* depletion of the Lsm proteins causes a defect in pre-mRNA splicing, Northern blot analysis was performed. Figure 4 shows

Fig. 2. Co-immunoprecipitation of Lsm proteins with antiserum against Lsm4p. Extracts from strains (AEMY28, AEMY29, AEMY31, AEMY33, AEMY34, AEMY35 and LMA4-2A), each producing a different HA-tagged Lsm protein or BMA64 [pACT11st] (HA-tagged Gal4-AD control), were subjected to immunoprecipitation with Lsm4p antibodies. All incubations and washes contained 150 mM NaCl. The precipitated proteins were fractionated by 15% (w/v) SDS-PAGE and electroblotted. Immunodetection was with anti-HA antibodies and anti-mouse HRP-conjugated secondary antibodies, and was visualized by ECL. The positions of the tagged fusion proteins are marked by asterisks.

that following a shift to the restrictive conditions, cells depleted of Lsm2, Lsm3, Lsm4, Lsm5, Lsm6, Lsm7 or Lsm8 protein accumulated intron-containing pre-U3 RNAs, indicating a splicing defect.

Analysis of small RNAs in the conditional strains

Cooper *et al.* (1995) observed that when Lsm4p was depleted *in vivo* there was a concomitant reduction in the level of U6 snRNA. The effect of a shift to the restrictive conditions on the levels of the spliceosomal snRNAs in the *lsm* conditional strains was therefore examined by Northern analysis (Figure 5A). No difference was observed in the levels of the spliceosomal snRNAs extracted from the *lsm1* strain compared with those from the parental strain when the cells were grown at 37°C. However, for each of the other conditional strains, a decrease in the level of U6 snRNA was observed upon a switch to the restrictive conditions. The level of U6 snRNA was lower in the *lsm6* and *lsm7* strains even at the permissive temperature, so the effect of shifting to the restrictive conditions was less obvious than for the gal-regulated strains. Indeed it may be that the defect causing cessation of growth at 37°C in the *lsm6* strain is not directly related to U6 snRNA function, but rather to some other defect arising from the lack of the Lsm6 protein (see Discussion). Although the kinetics of U6 decline were slower for the P_{GALI} -regulated strains, these had the lowest level of U6 after 24 h at the restrictive conditions, resembling the effect of depleting Lsm4p (Cooper *et al.*, 1995). The slow decline of U6 snRNA with these strains may be due to the time needed to titrate out the Lsm protein after repression of transcription, compared with the more rapid effect of heat treatment on the temperature-sensitive strains. From these data, it was concluded that in addition to Lsm4p, functional Lsm2, Lsm3, Lsm5, Lsm6, Lsm7 and Lsm8 proteins are needed for the normal accumulation of the U6 snRNA *in vivo*. A slight decrease in the level of U5_L snRNA was seen in P_{GALI} -LSM2 and P_{GALI} -LSM3,

Fig. 3. Deletion of genes encoding Lsm proteins. (A) Scheme for the production of conditional strains. (B) Examples of growth analyses of the conditional strains. Cells were grown to mid-logarithmic phase, harvested, washed and resuspended in the appropriate pre-warmed medium. OD₆₀₀ was followed with dilutions made where necessary to maintain the cells in logarithmic growth. The growth temperatures or carbon sources (in the case of P_{GALI} -LSM3 and its control grown at 30°C) are indicated to the right of each line. *x*-axis, time in hours after shift to the experimental conditions; *y*-axis, OD₆₀₀ values. These analyses were also performed for the other conditional strains and produced similar results (data not shown).

while the level of U5_S increased in P_{GALI} -LSM2 and P_{GALI} -LSM8 under restrictive conditions. The significance of this remains unclear.

To investigate the specificity of these effects on spliceosomal RNAs, the levels of several other small RNA species were examined: P RNA (the RNA component of RNase P), MRP RNA (the RNA component of RNase MRP) and 5S rRNA. The levels of most of these RNA species were less obviously affected in the conditional strains; however, there was a slight and reproducible reduction in the level of pre-P RNA in P_{GALI} -LSM2, P_{GALI} -LSM5 and P_{GALI} -LSM8 cells grown in glucose (Figure 5B). In addition, the level of pre-5S RNA declined in all except *lsm1* cells under restrictive conditions (data not shown), and mature 5S rRNA was slightly reduced in P_{GALI} -LSM2, P_{GALI} -LSM5 and P_{GALI} -LSM8 cells. Abnormalities in pre-tRNAs were also observed in several *lsm* strains (our unpublished results). Like U6 snRNA, the 5S rRNA, tRNAs and P RNA are products of RNA polymerase III. Although we cannot exclude the possibility that the effects on these RNAs may be due to indirect effects of Lsm protein depletion (e.g. as a result of

disruption of splicing), considering the proposal of Pannone *et al.* (1998) that La protein and Lsm8p collaborate in the stabilization of U6 and other polymerase III-transcribed small RNAs, these results suggest that all except Lsm1p may contribute to this activity.

Effect of U6 snRNA overproduction in conditional strains

Given the reduction in the U6 snRNA levels seen for most of the *lsm* conditional strains, the effect of overproducing U6 snRNA was investigated. U6 overproduction partially suppressed the growth defect of *lsm6* and *lsm7* cells at 37°C, the (previously) restrictive temperature (Figure 6A), and of *P_{GALI}-LSM2*, *P_{GALI}-LSM3* and *P_{GALI}-LSM4* cells on glucose medium (Figure 6B).

The best complementation was for the *P_{GALI}-LSM5* and *P_{GALI}-LSM8* cells, for which overproduction of U6 permitted full growth on glucose medium (Figure 6B). In contrast, overproducing U6 snRNA (confirmed by Northern blot analysis) had no discernible effect on the

growth of the *lsm1* strain at 30 or 37°C. (Figure 6A). Presumably this reflects the fact that the level of U6 snRNA is normal in *lsm1* cells at the restrictive temperature (Figure 5).

Thus, the overproduction of U6 snRNA can, to some extent, compensate for the loss of the Lsm2, Lsm3, Lsm4, Lsm5, Lsm6, Lsm7 and Lsm8 proteins, suggesting that at least part of the physiological function of these proteins is directly related to the stable accumulation of U6 snRNA.

Co-immunoprecipitation of snRNAs with tagged Lsm proteins

Northern analysis of spliceosomal snRNAs co-immunoprecipitating with each HA-tagged Lsm protein showed that Lsm2, Lsm5, Lsm6, Lsm7 and Lsm8 proteins all co-precipitated U6 snRNA and U4 snRNA (Figure 7A), as was previously reported for Lsm3p (Sraphin, 1995) and Lsm4p (Cooper *et al.*, 1995). Some U5 snRNA also co-precipitated with Lsm2p, and Lsm5p, presumably due to precipitation of the U4/U6-U5 tri-snRNP (and supported by co-precipitation of several of the Lsm proteins by antibodies against Prp8p which is present in tri-snRNPs; V.Vidal and J.D.Beggs, unpublished results). No immunoprecipitation of U1 or U2 snRNA was detected for any of the Lsm proteins. When the membranes in Figure 7A were reprobbed for P RNA, MRP RNA and 5S rRNA, none of these RNA species was detected in any of the immunoprecipitates (data not shown), suggesting that these proteins do not possess a general RNA-binding capacity.

With Lsm1p, no snRNAs were immunoprecipitated under these conditions (150 mM NaCl; Figure 7A) despite efficient precipitation of the protein (data not shown). A very low level of U6 snRNA could be co-immunoprecipitated with Lsm1p at 50 mM but not at higher salt concentrations (data not shown).

To determine whether the U6 snRNA that was associated with each of the Lsm proteins was in the free or the U4

Fig. 4. Effect of *in vivo* inactivation or depletion of Lsm proteins on splicing. (A) RNA was extracted from the gal-regulated strains and from the wild-type parent grown continuously in galactose or shifted to glucose medium for 12 h. RNA was separated in a 6% (w/v) denaturing polyacrylamide gel, electroblotted and hybridized with a radiolabelled oligonucleotide complementary to the U3 snoRNAs. (B) RNA was extracted and analysed as above for *lsm* or wild-type cells grown at 30°C. *represents a stable breakdown product of pre-U3 RNA (Hughes and Ares, 1991)

Fig. 5. Effect of *in vivo* inactivation or depletion of Lsm proteins on the levels of small RNAs. (A) The conditional strains and wild-type parents were grown under the restrictive conditions for the times indicated. RNAs were analysed as described for Figure 4 and probed with radiolabelled oligonucleotides complementary to the spliceosomal snRNAs. (B) The same membrane was stripped and reprobbed for P (or pre-P) RNA, MRP RNA and 5S rRNA.

Fig. 6. Effect of overproduction of U6 snRNA on growth of conditional strains. **(A)** Ten-fold serial dilutions of each conditional strain transformed with either vector (YEpl24) or U6 encoded on a high copy number plasmid (pYX117; Hu *et al.*, 1994) were spotted on selective medium and incubated at 30 or 37°C for 3 days. **(B)** Ten-fold serial dilutions of gal-regulated strains transformed with either vector (YEpl3 or YEpl24) or U6 encoded on a high copy number plasmid (pYX172 or pYX117; Hu *et al.*, 1994) were spotted on selective galactose or glucose medium and incubated at 30°C for 3 days.

base-paired form, the co-precipitated RNA was analysed under non-denaturing conditions that preserve the U4/U6 base pairing (Brow and Guthrie, 1988). The HA-tagged Lsm2, Lsm3, Lsm5, Lsm6, Lsm7 and Lsm8 proteins co-precipitated both free U6 and base-paired U4/U6 (Figure 7B and data not shown). In conclusion, all except Lsm1p associate stably with U6 snRNA, remaining associated with this RNA as it interacts with U4 snRNA, and at least transiently as the U4/U6·U5 tri-snRNPs form.

Analyses of snRNP complexes

The state of the U4 and U6 snRNAs was examined in extracts from *lsm6* and *lsm7* cells (deletion strains which contain no Lsm6p or Lsm7p, respectively). Compared with extracts from control strains, the *lsm* extracts contained greatly reduced levels of free U6 snRNP even at 30°C, and accumulated free U4 snRNP (Figure 8A). The majority of the residual U6 was complexed with U4. As *lsm7* extract contains a high level of U4/U6 duplex, it was of interest to investigate the U4/U6·U5 tri-snRNP content of this extract. Prp8p antibodies were used to immunoprecipitate tri-snRNPs (Cooper *et al.*, 1995), and the levels of the U4, U5 and U6 snRNAs that co-

precipitated were compared between *lsm7* and wild-type extracts (Figure 8B). The levels of U5 were similar (Prp8p is a U5 snRNP protein), whereas the levels of U4 and U6 in the *lsm7* precipitate were reproducibly only one-third of wild-type levels. Therefore, although the two extracts contained similar levels of U4/U6 duplexed RNAs, in the absence of Lsm7p less of this was in the form of tri-snRNPs.

Discussion

Searches of the complete *S.cerevisiae* genome database identified 16 putative ORFs encoding proteins with Sm motifs (Fromont-Racine *et al.*, 1997). Three of the encoded proteins had already been identified and characterized as homologues of canonical Sm proteins: Smd1p (Rymond, 1993; Rymond *et al.*, 1993), Smd3p (Roy *et al.*, 1995) and Smep (Bordonné and Tarassov, 1996), while four others had been shown to be associated with U1 snRNA but not characterized further (Neubauer *et al.*, 1997; Gottschalk *et al.*, 1998). Based on the two-hybrid interactions and co-immunoprecipitation experiments presented here, eight Sm-like or Lsm proteins appear capable of associating with one another to form a novel complex or complexes. Seven of these proteins (Lsm2, Lsm3, Lsm4, Lsm5, Lsm6, Lsm7 and Lsm8) associate with U6 snRNA and are required for maintenance of normal U6 snRNA levels and for pre-mRNA splicing.

The two-hybrid data for the Lsm proteins suggest greater promiscuity in their mutual protein interactions than for the canonical Sm proteins, which show preferences for particular Sm protein pairings (Fury *et al.*, 1997; Camasses *et al.*, 1998). However, no strong homotypic interactions were seen and, with the exception of Smbp and Lsm4p, several canonical Sm proteins that were tested did not interact with the Lsm proteins. The failure of Lsm4-Ap and Lsm4-Bp to interact indicates the requirement for both Sm motifs. An analogous situation has been reported for the canonical Sm proteins, where deletion of either of the Sm motifs leads to loss of their interaction (Hermann *et al.*, 1995; Camasses *et al.*, 1998). The two-hybrid approach may also detect indirect interactions, mediated by a bridging protein or RNA. The apparent promiscuity of the Lsm proteins may reflect indirect interactions in some cases. In exhaustive two-hybrid screens of a genomic library using many bait proteins including dozens of splicing factors, Lsm proteins were found as prey almost exclusively by Lsm proteins as baits (A.E.Mayes, M.Fromont-Racine, J.D.Beggs and P.Legrain, unpublished results). Thus the Lsm proteins have strong, highly specific interactions with each other.

It is not known whether either the two-hybrid interactions or the co-immunoprecipitation of pairs of Lsm proteins is facilitated by U6 snRNA. However, it is interesting to note the strong co-immunoprecipitation of HA-Lsm1p with Lsm4p (in 150 mM NaCl), given that U6 snRNA did not co-precipitate with Lsm1p in the same conditions. Thus Lsm1p and Lsm4p may interact in the absence of U6 snRNA.

From the analysis of snRNA levels in the conditional strains, it is evident that the level (and presumably the stability, although this has not been demonstrated formally) of U6 snRNA is dependent on the presence of seven

Fig. 7. Co-immunoprecipitation of snRNAs with Lsm proteins. (A) Extracts from strains each producing an HA-tagged Lsm protein or HA-tagged control protein (all as in Figure 2) were incubated with anti-HA antibodies. Total lanes contain RNA from one-fifth the amount of extract used in the immunoprecipitation reactions. RNAs were analysed as described in Figure 4. Note that HA-tagged Lsm2p is precipitated relatively inefficiently by anti-HA antibodies (data not shown), presumably due to poor availability of the epitope tag. (B) Non-denaturing analysis of U6 snRNA co-immunoprecipitated with Lsm proteins. The RNAs were extracted from the precipitates and analysed by non-denaturing electrophoresis, Northern blotting and probing for U6 snRNA. Total as in (A).

Fig. 8. Analysis of U4/U6 and U4/U6-U5 complexes in *lsm* deletion strains. (A) Non-denaturing analysis of U4 and U6 snRNAs in the *lsm6* and *lsm7* deletion strains. Extracts are from the *lsm6* (AEMY19) and *lsm7* (AEMY22) deletion strains, and *LSM6* (AEMY34) and *LSM7* (AEMY35) controls grown at 30°C. Total RNA from each extract was analysed as in Figure 7B. As a control, a sample of RNA was denatured by boiling prior to loading on the gel. (B) Analysis of tri-snRNPs in the *lsm7* extract. Extracts from the *lsm7* deletion and *LSM7* control strains were incubated with anti-Prp8p antibodies and the immunoprecipitated RNAs were extracted and analysed by denaturing gel electrophoresis as for Figure 4, probing for U4, U5 and U6 snRNAs. The levels of these RNAs were compared by phosphorimaging.

functional Lsm proteins. This is similar to the effect on the U1, U2, U4 and U5 snRNAs of depleting Sm proteins (e.g. Rymond, 1993).

Pannone *et al.* (1998) reported similar effects of La protein in the biogenesis of the U6 snRNP, suggesting that the La protein (which associates with nascent U6 and other RNA polymerase III transcripts) acts as a chaperone, stabilizing the U6 snRNA structure in a conformation suitable for the formation of the U6 snRNP. These authors described a synthetic lethal interaction between the genes

encoding the La and Lsm8 proteins, and proposed that Lsm8p may be the first U6-specific protein that binds to U6 snRNA, although no evidence was presented for a direct interaction between these two proteins. Since the La protein is non-essential, it may be the Lsm complex which is really the chaperone. The La protein may function by 'handing off' (as proposed by Herschlag, 1995) the U6 snRNA from the transcription machinery to the Lsm proteins, which in turn facilitate di- and tri-snRNP formation. Subtle but reproducible effects of Lsm protein depletion on levels of pre-P, pre-5S and pre-tRNAs also support the proposal of Pannone *et al.* (1998). Although none of these other RNAs was observed to co-precipitate with any of the HA-tagged Lsm proteins in extracts, Lsm proteins appear to bind to some of these RNAs *in vitro* (L. Verdone and J.D. Beggs, unpublished results). The specificity of RNA binding by these proteins is currently being investigated more fully. U6 snRNA is different from the other RNAs affected in that this mature species requires continued association with the Lsm proteins for its stable accumulation as an RNP. The mature forms of the other polymerase III products associate with different sets of proteins.

Analysis of the U4 and U6 snRNAs under non-denaturing conditions showed that in the absence of either Lsm6p or Lsm7p, free U6 snRNPs were severely depleted, and free U4 particles accumulated, presumably as a consequence of the reduced level of U6 in these cells. An effect on U4/U6-U5 tri-snRNP formation was also observed; with extract lacking Lsm7p (or Lsm4p; Cooper *et al.*, 1995), the amount of tri-snRNPs co-immunoprecipitating with Prp8p was reduced, although the level of U4/U6 duplex was normal. Thus Lsm4p and Lsm7p, at least, may play a role in tri-snRNP formation and/or stabilization. The co-precipitation of U5 snRNA with some of the HA-tagged Lsm proteins (Figure 7)

Fig. 9. Sequence alignment of Sm-like proteins. (A) Putative homologues of Lsm2p were identified by BLAST searches (<http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast?Jform=0>) and aligned using the PILEUP program in GCG10. Identities and similarities are highlighted using BOXSHADE 3.21 (http://www.isrec.isb-sib.ch/software/BOX_form.html). The positions of the Sm motifs 1 and 2 are indicated. White on black represents amino acid identity in at least five of the nine sequences; black on grey represents conservation of the nature of the amino acid at that site. Accepted groupings were M=I=V=L, K=R=H, F=Y=W, S=T, E=D, A=G, Q=N. The accession numbers of the proteins are as follows: *Homo sapiens*, AA315292; *Mus musculus*, U85207; *Branchiostoma floridae*, Z83273; *Drosophila melanogaster*, AA821196; *Brugia malayi*, AA228204; *Caenorhabditis elegans*, Z81118; *Arabidopsis thaliana*, AC005278; *Saccharomyces cerevisiae*, P38203; *Schizosaccharomyces pombe*, AL034491. The sequences shown for the *C. elegans* and *A. thaliana* proteins differ from those in the database due to the predicted use of other splice sites. The conceptual translations used here produce proteins of the size expected for Lsm2p homologues, whilst those in the database give larger products with a disrupted Sm architecture. (B) Sm-like proteins from ancient organisms. Polypeptide sequences containing the Sm motifs were identified from BLAST searches of protein databases. Sequences were aligned with identities and similarities highlighted as in (A). MTH649, *Methanobacterium thermoautotrophicum* protein (accession No. AE000845); MTH1440, *M. thermoautotrophicum* protein (accession No. AE000905); AF362, *Archaeoglobus fulgidus* protein (accession No. AE001079); AF875, *A. fulgidus* protein (accession No. AE001044); *P. minut.*, *Pedinomonas minutissima* protein (accession No. U58510).

and with untagged Lsm4p indicates at least a transient association of Lsm proteins with tri-snRNPs.

The yeast Prp24 protein has been identified as a factor involved in the formation and disassembly of U4/U6 duplexes (Ghetti *et al.*, 1995; Jandrositz and Guthrie, 1995). Interestingly, *prp24* mutants resemble *lsm* mutants in having reduced levels of U6 snRNA and, in extracts containing destabilized U4/U6, Prp24p co-precipitated small amounts of free U4 as well as free U6 and base-paired U4/U6 (Ragunathan and Guthrie, 1998a). Prp24p has been demonstrated to promote the annealing of U4 and U6 snRNPs in yeast extracts; however, the annealing activity of Prp24p with deproteinized U4 and U6 snRNAs was markedly lower, prompting the suggestion that other proteins contribute to the rate of annealing (Ragunathan and Guthrie, 1998a). Thus, Lsm proteins may cooperate with Prp24p in promoting the association of the U4 and U6 snRNAs. In support of such a model, genetic interactions between *LSM4* and *PRP24* have been observed (A.E.Mayes, M.Cooper and J.D.Beggs, unpublished results), and Prp24p has been identified in exhaustive two-hybrid screens with several Lsm proteins (A.E.Mayes, M.Fromont-Racine, J.D.Beggs and P.Legrain, unpublished results).

RNA-protein interactions are expected to respond to changing circumstances faster than RNA-RNA interactions (Herschlag, 1995); thus the Lsm proteins may have a chaperone-like function to facilitate U4/U6 dimer formation and possibly also in tri-snRNP assembly, minimizing the energy required to drive conformational rearrangements. This would be additional to the roles of Prp24p as a 'matchmaker' (Herschlag, 1995) in the association of U4 with U6, and of the RNA-unwinding protein Brr2p in their dissociation (Ragunathan and Guthrie, 1998b).

In summary, we propose that the Lsm proteins form a complex on free U6 RNA that is either newly synthesized (and may be La-associated; Pannone *et al.*, 1998) or released from dissociating spliceosomes. This complex protects the RNA against degradation and may facilitate subsequent conformational rearrangements of the RNA and/or of other proteins involved in the association of U6 with U4 and then with U5 snRNPs as these particles form tri-snRNPs. The weak co-precipitation of U5 snRNA with the Lsm proteins suggests that either the epitopes become masked in the tri-snRNPs or the Lsm proteins dissociate from the tri-snRNPs soon after their formation and/or the destabilization of the U4/U6 interaction in tri-snRNPs.

Following completion of the splicing reaction, the U6 snRNA must reassociate with the U4 snRNA, and in cells depleted of Lsm proteins this U6 snRNA is degraded, resulting in the accumulation of free U4 snRNPs.

An interesting question that remains is the subcellular localization of the Lsm proteins and of U6 snRNA. Most current evidence suggests a nuclear localization for U6 snRNA; however, U6 snRNA free from U4 snRNA has been reported to be present and matured in the cytosol of mouse fibroblasts prior to nuclear import and association with U4 snRNA (Fury and Zieve, 1996). Obviously, an additional role for the Lsm protein complex may be to act as (part of) a nuclear localization signal analogous to the canonical Sm proteins, facilitating the nuclear import of U6 snRNA that might be present transiently in the cytosol, for example immediately after mitosis.

The stoichiometry of the proteins present in the canonical Sm complex has been studied (Raker *et al.*, 1996; Plessel *et al.*, 1997), and a structural model has been proposed in which a single copy of each protein is present in a seven-membered complex (Kambach *et al.*, 1999). Data presented here suggest that seven interacting Lsm proteins associate with U6 snRNA and, although it is not demonstrated that all interact simultaneously in the same U6 complex, a seven-membered complex is attractive by analogy with the canonical Sm complex.

In striking contrast to the others, depletion of Lsm1 had no effect on the level of U6 (or other polymerase III transcripts) or on the efficiency of pre-mRNA splicing, and HA-tagged Lsm1p did not associate stably with U6 snRNA. Therefore, although Lsm1 can interact with the other Lsm proteins, it is not a component of the U6-associated complex. It is therefore conceivable that there is more than one form of Lsm complex with alternative protein compositions, which might have distinct functions and/or substrate specificities. This is also suggested by the fact that, in most cases, the growth defect caused by the depletion of the Lsm proteins is not fully suppressed by the overproduction of U6 snRNA (Figure 6). This would be expected if Lsm proteins have another essential function that is independent of U6. Clearly, the proposed role of Lsm1p in decapping mRNA (Boeck *et al.*, 1998) suggests a possible function for another Lsm complex. Indeed, data from two-hybrid screens indicate that several Lsm proteins interact with factors involved in mRNA turnover (A.E. Mayes, M. Fromont-Racine, J.D. Beggs and P. Legrain, unpublished results), and there is direct evidence that multiple Lsm proteins influence mRNA decapping (R. Parker, personal communication).

Database searches reveal that some Lsm proteins appear to have been conserved through evolution. A sequence alignment (Figure 9A) of Lsm2p structural homologues from a number of higher organisms shows that the sequence identities extend beyond the Sm motifs. Lsm2p has 63% identity (75% similarity) to the human sequence, whilst the other *S. cerevisiae* Lsm proteins all have <32% identity (41% similarity). Also, the identification of an Sm-like protein and a U6-like RNA in the miniaturized genome of *Pedinomonas minutissima* suggests that these are ancient macromolecules (Gilson and McFadden, 1996). The presence of Sm motif sequences in the genomes of *Methanobacterium thermoautotrophicum* and *Archaeo-*

globus fulgidus (Klenk *et al.*, 1997; Smith *et al.*, 1997) reinforces this theory (see Figure 9B). Since neither of these archaeobacteria contains recognizable splicing machinery, these Sm-like proteins may affect processes more fundamental than RNA splicing. Eukaryotes may have enlarged the Lsm protein family, and recruited Lsm proteins to function in pre-mRNA splicing in addition to their roles in other cellular processes.

In this work, we have investigated the association of the Lsm proteins with U6 snRNA and their consequent role in pre-mRNA splicing. However, there are strong indications that these proteins may have more general functions. Hopefully, the functional characterization of a number of interacting factors that were identified in a systematic programme of exhaustive two-hybrid screens with Lsm proteins as baits will give new clues as to the roles, locations and other associations of these novel proteins.

Materials and methods

Yeast manipulations

The genotypes of *S. cerevisiae* strains used in this work are listed in Table II. Yeast cells were propagated and sporulated as described by Cooper *et al.* (1995). Yeast transformations were performed as in Geitz *et al.* (1992).

Two-hybrid direct matings

The two-hybrid bait vectors were pAS2ΔΔ (Gal4 DNA-binding domain; Fromont-Racine *et al.*, 1997) and pBTM116 (LexA DNA-binding domain; Vojtek *et al.*, 1993). Two-hybrid prey were constructed using pACT11st (Fromont-Racine *et al.*, 1997), except those for Smb, Smd1 and Smd3, which were as reported by Fromont-Racine *et al.* (1997). Yeast strains CG1945 (for Gal4 bait fusions), L40 (for LexA bait fusions) or Y187 (for prey fusions) were grown on selective media. Bait and prey strains were mated by replica-plating onto rich medium, and diploids were grown on medium selecting for both the bait and prey plasmids, then tested on histidine-free medium for a successful two-hybrid interaction. The stringency of the interaction was tested by growth of the diploids on selective medium containing up to 50 mM 3-AT.

Gene deletions and complementations

All primers for PCR were based on the coding sequences in the *Saccharomyces* Genome Database (<http://genome-www.stanford.edu/Saccharomyces/>). Gene deletions were made by replacing the entire coding sequence with either a *HIS3* or *TRP1* cassette (Baudin *et al.*, 1993). Each deletion was made in two genetic backgrounds, BMA38 or BMA64 (Table II) and JDY6 (data not shown); the same results were obtained in each strain. The diploids were sporulated at 23°C and viable progeny were scored for the appropriate auxotrophic marker, and for growth at 23, 30 and 37°C. For essential genes, the diploid strain was transformed with a *P_{GALI}*-regulated version of the gene, sporulated and haploid progeny were tested for complementation of the deletion. *LSM5* and *LSM8* extensively overlap uncharacterized ORFs on the other strand (previously unannotated and YJR023c, respectively) which were also disrupted by the knockout deletions. However, prey fusion constructs encoding Lsm5p or Lsm8p, but not the putative products of the other strand, complement the growth defect caused by the deletions, thus confirming that *LSM5* and *LSM8* are essential.

HA-tagging of the Lsm proteins

The Lsm proteins were tagged with a single HA epitope (nine amino acids) at their N-termini, with the exception of Lsm2p which was C-terminally tagged. The HA-tagged *LSM2*, *LSM3*, *LSM5* and *LSM8* genes were cloned in pBM125. *LSM6* and *LSM7* were cloned (with the intron of *LSM7* removed) in YCpIF16 to generate *P_{GALI}*-regulated, HA-tagged versions. The two-hybrid prey fusions of *LSM5* and *LSM8* have an HA tag and were used for the immunoprecipitation studies. The chromosomal *LSM1* gene was tagged by replacing *lsm1Δ::TRP1* in AEMY24, with HA-tagged *LSM1* sequence. The functional ability of each tagged protein was confirmed by rescue of the growth defect of the gene-deleted strains.

RNA extraction and analysis

RNA was extracted from yeast cells by the method of Schmitt *et al.* (1990). Denaturing Northern analysis and probes for detecting the small RNAs were as described by Cooper *et al.* (1995). Oligonucleotides were provided by D.Tollervey for the detection of: P RNA, ATTTCTG-ATAACAACGGTCGG; MRP RNA, AATAGAGGTACCAGGTCAA-GAAGC; 5S rRNA, CTACTCGGTACAGCTC; and U3 (2'-O-methyl-RNA), UUAUGGGACUUGUU. Non-denaturing conditions for the extraction of complexed U4/U6 snRNAs from yeast cell extracts were as in Brow and Guthrie (1988).

Immunoprecipitations

Yeast cell extracts were prepared as described by Lin *et al.* (1985). Immunoprecipitation of RNAs was as described by Cooper *et al.* (1995). Co-precipitated proteins were analysed by SDS-PAGE and Western analysis, and visualized by enhanced chemiluminescence (ECL, Amersham). Antibodies for immunodetection and immunoprecipitation were: anti-HA antibodies (Boehringer Mannheim), rabbit polyclonal anti-Lsm4p (anti-Uss1p) or anti-Prp8p antibodies (Cooper *et al.*, 1995), and anti-mouse horseradish peroxidase (HRP)-conjugated second antiserum.

Acknowledgements

We are very grateful to M.Fromont-Racine for the *LSM8* deletion strain, to M.Fromont-Racine and J.-C.Rain for providing the *LSM8* two-hybrid constructs prior to publication, to M.Cooper for *lsm4Δ*, *lsm4-A* and *lsm4-B* constructs, and to Jeremy Brown, Roy Parker and David Tollervey for helpful comments on this manuscript. A.E.M. was the recipient of a Wellcome Trust Prize Studentship. J.D.B. holds a Royal Society Cephalosporin Fund Senior Research Fellowship. This work was partly funded by Wellcome Trust Grant 044374 and EU Biotech Grant 95007 (TAPIR Network).

References

- Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. and Cullin, C. (1993) A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **21**, 3329–3330.
- 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.
- Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Bernard, H. and Jacob, M. (1982) U2 snRNA shares a structural domain with U1, U4 and U5 snRNAs. *EMBO J.*, **1**, 1259–1265.
- Brow, D.A. and Guthrie, C. (1988) Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature*, **334**, 213–218.
- 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.
- Cooper, M., Parkes, V., 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 ribonucleoprotein particles. *EMBO J.*, **14**, 2066–2075.
- Fischer, U. and Lührmann, R. (1990) An essential signalling role for the m³G cap in the transport of U1 snRNP to the nucleus. *Science*, **249**, 786–790.
- Fromont-Racine, M., Rain, J.-C. and Legrain, P. (1997) Towards a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nature Genet.*, **16**, 277–282.
- Fury, M.G. and Zieve, G.W. (1996) U6 snRNA maturation and stability. *Exp. Cell Res.*, **228**, 160–163.
- Fury, M.G., Zang, W., Christodouloupoloulos, I. and Zieve, G.W. (1997) Multiple protein:protein interactions between the snRNP core proteins. *Exp. Cell Res.*, **237**, 63–69.
- Geitz, D., St Jean, A., Woods, R.A. and Schiestl, R.H. (1992) Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res.*, **20**, 1425.
- Ghetti, A., Company, M. and Abelson, J. (1995) Specificity of Prp24 binding to RNA: a role for Prp24 in the dynamic interaction of U4 and U6 snRNAs. *RNA*, **1**, 132–145.
- Gilson, P.R. and McFadden, G.I. (1996) The miniaturized nuclear genome of a eukaryotic endosymbiont contains genes that overlap, genes that are co-transcribed, and the smallest known spliceosomal introns. *Proc. Natl Acad. Sci. USA*, **93**, 7737–7742.
- Gottschalk, A. *et al.* (1998) A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. *RNA*, **4**, 374–393.
- Hamm, J., Darzynkiewicz, E., Tahara, S. and Mattaj, I.W. (1990) The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. *Cell*, **62**, 569–577.
- Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahms, H. and Lührmann, R. (1995) snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein–protein interactions. *EMBO J.*, **14**, 2076–2088.
- Herschlag, D. (1995) RNA chaperones and the RNA folding problem. *J. Biol. Chem.*, **270**, 20871–20874.
- Hollenberg, S.M., Sternglanz, R., Cheng, P.F. and Weintraub, H. (1995) Identification of a new family of tissue-specific basic helix–loop–helix proteins with a two-hybrid system. *Mol. Cell. Biol.*, **15**, 3813–3822.
- Hu, J., Xu, Y., Schappert, K., Harrington, T., Wang, A., Braga, R., Mogridge, J. and Friesen, J.D. (1994) Mutational analysis of the PRP4 protein of *Saccharomyces cerevisiae* suggests domain structure and snRNP interactions. *Nucleic Acids Res.*, **22**, 1724–1734.
- Hughes, J. and Ares, M. (1991) Depletion of U3 small nucleolar RNA inhibits cleavage in the 5' external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. *EMBO J.*, **10**, 4231–4239.
- Jandrositz, A. and Guthrie, C. (1995) Evidence for a Prp24 binding site in U6 snRNA and in a putative intermediate in the annealing of U6 and U4 snRNAs. *EMBO J.*, **14**, 820–832.
- 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.
- Klenk, H.P. *et al.* (1997) The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature*, **390**, 364–370.
- Krämer, A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.*, **65**, 367–409.
- Lerner, M.R. and Steitz, J.A. (1979) Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc. Natl Acad. Sci. USA*, **76**, 5495–5499.
- Lin, R.-J., Newman, A.J., Cheng, S.-C. and Abelson, J. (1985) Yeast mRNA splicing *in vitro*. *J. Biol. Chem.*, **260**, 14780–14792.
- Mattaj, I.W. (1986) Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding. *Cell*, **46**, 905–911.
- Mattaj, I.W. and DeRobertis, E.M. (1985) Nuclear segregation of U2 snRNA requires binding of specific snRNP proteins. *Cell*, **40**, 111–118.
- Moore, M.J., Query, C.C. and Sharp, P.A. (1993) Splicing of precursors to messenger RNAs 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., 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.
- 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.
- Parkes, V. and Johnston, L.H. (1992) *SPO12* and *SIT4* suppress mutations in *DBF2*, which encodes a cell cycle protein kinase that is periodically expressed. *Nucleic Acids Res.*, **20**, 5617–5623.
- Plessel, G., Lührmann, R. and Kastner, B. (1997) Electron microscopy of assembly intermediates of the snRNP core: morphological similarities between the RNA-free (E.F.G) protein heteromer and the intact snRNP core. *J. Mol. Biol.*, **265**, 87–94.
- Ragunathan, P.L. and Guthrie, C. (1998a) A spliceosomal recycling factor that reanneals U4 and U6 small nuclear ribonucleoprotein particles. *Science*, **279**, 857–860.
- Ragunathan, P.L. and Guthrie, C. (1998b) RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and DEIH-box splicing factor Brt2. *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 a snRNP subcore particle *in vitro*. *EMBO J.*, **15**, 2256–2269.
- Reddy, R. and Busch, H. (1988) Small RNAs: RNA sequences, structure, and modifications. In Birnstiel, M.L. (ed.), *Structure and Function of*

- Major and Minor Small Nuclear Ribonucleoprotein Particles.* Springer-Verlag, Berlin, pp. 1–37.
- Roy,J., Zheng,B., Rymond,B.C. and Woolford,J.L. (1995) Structurally related but functionally distinct yeast Sm D core snRNP 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., Rokeach,L.A. and Hoch,S.O. (1993) Human snRNP polypeptide D1 promotes pre-mRNA splicing in yeast and defines non essential yeast Smd1p sequences. *Nucleic Acids Res.*, **21**, 3501–3505.
- Schmitt,M.E., Brown,T.A. and Trumpower,B.L. (1990) A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **18**, 3091.
- Sraphin,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.
- Smith,D.R. *et al.* (1997) Complete genome sequence of *Methanobacterium thermoautotrophicum*ΔH: functional analysis and comparative genomics. *J. Bacteriol.*, **179**, 7135–7155.
- Staley,J.P. and Guthrie,C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell*, **92**, 315–326.
- van Venrooij,W.J. (1987) Autoantibodies against small nuclear ribonucleoprotein components. *J. Rheumatol.*, **14**, 78–82.
- Vojtek,A.B., Hollenberg,S.M. and Cooper,J.A. (1993) Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell*, **74**, 205–214.
- Zieve,G.W. and Sauterer,R.A. (1990) Cell biology of the snRNP particles. *Crit. Rev. Biochem. Mol. Biol.*, **25**, 1–46.

*Received February 1, 1999; revised June 11, 1999;
accepted June 15, 1999*