

Yeast Pre-mRNA Splicing Requires a Pair of U1 snRNP-Associated Tetratricopeptide Repeat Proteins

MITCH R. McLEAN¹ AND BRIAN C. RYMOND^{1,2*}

*T. H. Morgan School of Biological Sciences¹ and Markey Cancer Center,²
University of Kentucky, Lexington, Kentucky 40506-0225*

Received 19 August 1997/Returned for modification 29 September 1997/Accepted 16 October 1997

The U1 snRNP functions to nucleate spliceosome assembly on newly transcribed pre-mRNA. *Saccharomyces cerevisiae* is unusual among eukaryotes in the greatly extended length of its U1 snRNA and the apparent increased polypeptide complexity of the corresponding U1 snRNP. In this paper, we report the identification of a novel U1 snRNP protein, Prp42p, with unexpected properties. Prp42p was identified by its surprising structural similarity to the essential U1 snRNP protein, Prp39p. Both Prp39p and Prp42p possess multiple copies of a variant tetratricopeptide repeat, an element implicated in a wide range of protein assembly events. Yeast strains depleted of Prp42p by transcriptional repression of a *GAL1::PRP42* fusion gene arrest for splicing prior to pre-mRNA 5' splice site cleavage. Prp42p was not observed in a recent biochemical analysis of purified U1 snRNPs from *S. cerevisiae* (28). Nevertheless, antibodies directed against an epitope-tagged version of Prp42p specifically precipitate U1 snRNA from yeast extracts. Furthermore, Prp42p is required for U1 snRNP biogenesis, because yeast strains depleted of Prp42p formed incomplete U1 snRNPs that failed to produce stable complexes with pre-mRNA *in vitro*. The evidence shows that Prp39p and Prp42p are both required to configure the atypical yeast U1 snRNP into a structure compatible with its evolutionarily conserved role in pre-mRNA splicing.

The basic steps in spliceosome assembly are well conserved between *Saccharomyces cerevisiae* and humans (reviewed in references 17 and 27). An initiating event in intron selection is recognition of the pre-mRNA by the U1 snRNP. Proteins bound to the pre-mRNA and associated with the U1 snRNP (i) stabilize the highly specific 4- to 7-bp interaction between the 5' splice site and the 5' end of the U1 snRNA and (ii) mediate branch point recognition by the U1 snRNP (reviewed in references 12 and 31). The resulting structure, the commitment complex, serves as a substrate for prespliceosome formation through the ATP-dependent addition of the U2 snRNP. The subsequent addition of the U4-U6.U5 snRNP-tri-snRNP complex and an unknown number of non-snRNP proteins completes spliceosome assembly and promotes intron removal.

Yeast uses an atypical U1 snRNP in commitment complex formation. The yeast U1 snRNA is 3.5 times larger than its metazoan equivalent (19, 40), and associated with it are at least six proteins without obvious counterparts in the mammalian U1 snRNP (28). The relevance of this increased yeast U1 snRNP complexity to splice site identification and spliceosome assembly is poorly understood. Early experiments revealed that much of the yeast-specific U1 snRNA could be deleted without significant detriment to cell viability (24, 41). The remaining U1 snRNA appears to fold into a structure similar to that found in the mammalian snRNP (18). Thus, the bulk of the yeast-specific U1 snRNA, while possibly contributing to the efficiency of the splicing event, does not provide an essential function in yeast.

Genetic and biochemical studies have confirmed the presence of the ubiquitous U1-specific proteins in the yeast U1 snRNP. The genes for yeast strains U1-70k (Snp1p [44]) and U1-C (47) were identified from genomic DNA sequence data

based on the phylogenetic conservation of the encoded proteins. Each proved essential for U1 snRNP function. A synthetic lethal screen for mutations that exacerbate a mutant U1 snRNA phenotype led to the identification of the yeast U1-A counterpart (Mud1p [25]). Curiously, Mud1p is not required for pre-mRNA splicing or cellular viability.

In addition to the conserved U1 snRNP protein genes, two genes have been found that encode proteins defined to date only in yeast. *PRP39* was discovered in a screen for temperature-sensitive mutants defective in pre-mRNA splicing (26). *In vitro* studies revealed that Prp39p is a U1 snRNP protein and is required to assemble a productive U1 snRNP-pre-mRNA complex. Prp40p was initially identified as a suppressor of a cold-sensitive C-to-U nucleotide change at position 4 of the U1 snRNA (16). Consistent with a primary role in intron removal, Prp40p is required for pre-mRNA splicing *in vivo* and *in vitro*. A recent investigation of commitment complex assembly revealed that Prp40p interacts with an evolutionarily conserved protein, BBP, which binds to the pre-mRNA branch point sequence (2, 4, 5). Yeast BBP also binds the yeast U2AF counterpart, Mud2p (1, 2). A Prp40p-BBP-Mud2p association may underpin the U1 snRNP-mediated branch point recognition that occurs in commitment complex formation (see references 2, 5, and 27 and references within). A recent yeast two-hybrid study (8) suggests that Prp39p may also interact with Mud2p. If so, Prp39p would provide an additional tether between the U1 snRNP and the branch point region of the intron. Thus, in contrast to the dispensable nature of the excess yeast U1 snRNA sequence, at least two of the yeast-specific proteins are critical for U1 snRNP function.

A number of proteins with probable common ancestry function in pre-mRNA splicing. Examples of such proteins include the Sm core snRNP proteins (14, 35), the snRNP-specific mammalian proteins U1A (42) and U2B'' (13), the serine- and arginine-rich (SR) splicing factors (9), and the family of DEAD/H-box proteins (see references 17 and 27). Here we report the identification of a novel yeast protein, Prp42p, with

* Corresponding author. Mailing address: T. H. Morgan School of Biological Sciences, University of Kentucky, Lexington, KY 40506-0225. Phone (606) 257-5530. Fax: (606) 257-1717. E-mail: rymond@pop.uky.edu.

extensive similarity to the U1 snRNP protein, Prp39p. Embedded within the Prp39p and Prp42p proteins are multiple copies of a 34-amino-acid repeat (tetratricopeptide repeat [TPR]) similar to the TPR elements found in the crooked neck (crn) protein of *Drosophila melanogaster* (52). Prp39p and Prp42p are not functionally redundant, because each is needed for cellular pre-mRNA splicing. The data show that like Prp39p, Prp42p is required to assemble a stable U1 snRNP capable of productive interaction with cellular pre-mRNA.

MATERIALS AND METHODS

Identification of Prp42p. A BLASTP comparison (3) of Prp39p amino acids 1 to 629 with the *S. cerevisiae* nonredundant protein database was performed at site www-genome.stanford.edu/ with the Blossum 62 matrix. The best match for a novel protein, P[N] = $2.5e^{-10}$, was obtained with the 544-amino-acid hypothetical protein YDR235W. Pairwise sequence comparisons of Prp39p and Prp42p were performed with the Gap and BestFit programs (University of Wisconsin GCG package). TPR structures were initially selected on the basis of the presence of five or more conserved tryptophan and hydrophobic residues at positions 4, 7, 8, 11, 21, 24, and 27. The GCG Helical Wheel program and the Protein Sequence Analysis System (45, 50) were then used to screen for amphipathic helical sequences within the putative TPRs. Consensus sequences were derived from the Prp39p and Prp42p TPRs with the GCG Pretty program.

Cloning and characterization of PRP42. The *PRP42* open reading frame (ORF) was amplified from yeast genomic DNA by PCR (upstream primer, 5' AGA GGA TCC ATG GAT AAA TAT ACT GCT TTG ATT CAC 3'; downstream primer, 5' CCA GGA TCC AAT AAA TGA CAA TGC CTT TTG GCT AAG G 3'). The primer-encoded *Bam*HI sites (underlined) were used to fuse the *PRP42* ORF to the *GAL1* promoter of plasmid pBM150 (15) to create *GAL1::PRP42*. *GAL1::PRP42HA* was created in a similar manner, except that the downstream PCR primer contained codons for the nine-amino-acid hemagglutinin (HA) epitope (5' TTT GGA TCC CTA AGC GTA GTC TGG AAC GTC GTA TGG GTA AGG TTC TTC AGT AAA CAT TTC CTC 3'; *Bam*HI site, underlined; HA codons, boldface).

The *PRP42* ORF described above was blunt end ligated into the *Sma*I site of pTZ19U (United States Biologicals). An internal deletion of approximately half of the *PRP42* coding sequence was then created by excision of an internal 0.81-kbp *Bst*BI fragment. This deletion derivative was then modified for genetic selection by blunt end insertion of a 2.2-kbp DNA fragment containing the *LEU2* gene of yeast. The resultant *prp42::LEU2* gene was released from the vector DNA by digestion with *Bam*HI, followed by treatment with mung bean nuclease. Gel-purified *prp42::LEU2* DNA was then used to replace one of the two wild-type copies of *PRP42* in the diploid yeast strain MGD407 (33). The site of insertion was confirmed by PCR analysis of the genomic DNA. Standard techniques (37) were used to induce sporulation and dissect the meiotic progeny of this strain. To create the *GAL1::PRP42* and *GAL1::PRP42HA* strains, the respective plasmids were transformed into the MGD407 heterozygous disruptant prior to sporulation. The asci were dissected on yeast extract-peptone medium containing 2% galactose to permit expression of the plasmid-borne fusion genes. The haploid offspring were then phenotypically scored on selective medium to screen for the presence of the *prp42::LEU2* gene and for galactose-dependent cell growth.

Yeast exact preparation and immune precipitations. Yeast extracts were prepared by grinding liquid nitrogen-frozen cell pellets with a mortar and pestle and processing the lysate as previously described (49). For immune precipitation, 1 μ l of the antibody was bound to 10 μ l of protein G agarose as previously described (26). The antibodies used were the HA-specific HA.11 antibody (Babco) and a nonspecific monoclonal antibody (MAb63) (gift of M. Mendenhall). For typical reactions, yeast extract containing approximately 250 μ g of protein was mixed with 10 μ l of antibody-bound beads in a total volume of 100 μ l of HNT buffer (20 mM HEPES [pH 7.9], 100 mM NaCl, 0.05% Triton X-100) at room temperature for 30 min. The NaCl concentration of the HNT buffer was adjusted between 50 and 300 mM as needed to assay the tightness of the Prp42HAp-U1 snRNP association. The unbound extract was separated from the antibody-bound material by centrifugation in a Marathon 16 KM microcentrifuge (Fisher Scientific) at 3,500 rpm for 1 min. The beads were washed six times with 300 μ l of HNT. To release the antibody-associated snRNAs, 100 μ l of PK buffer (100 mM Tris-HCl [pH 7.5], 12.5 mM EDTA, 150 mM NaCl, 1% sodium dodecyl sulfate, 400 μ g of proteinase K per ml) was added, and the mixture was incubated at 37°C for 10 min. After phenol extraction, the samples were concentrated by ethanol precipitation and assayed by Northern blotting with the previously described snRNA probes (6). Unbound and total snRNA preparations were recovered and assayed in parallel by analogous procedures.

Splicing analysis, spliceosome assembly, and snRNP complexes. The analysis of in vivo pre-mRNA splicing was performed as previously described (6). RNA was extracted from yeast cultures grown on galactose and for various times after the shift to a glucose-based medium. Approximately 25 μ g of total RNA was fractionated on a 1% agarose-formaldehyde gel. A transfer on Nytran-plus

membrane was then hybridized with probes specific for the *RP51A* and *ADE3* genes (6).

In vitro splicing reactions were assembled on *RP51A* transcripts prepared by in vitro transcription of the pSPPrp51A construct (29) with SP6 RNA polymerase. The splicing reaction was fractionated on a 5% polyacrylamide-7 M urea denaturing gel to assay splicing and on a 3% polyacrylamide-0.5% agarose gel to assay the assembly of splicing complexes as previously described (26). To visualize the commitment complex bands, the U2 snRNA was cleaved with the extract's endogenous RNase H activity (34). This was accomplished by a 10-min preincubation with 0.1 μ g of the anti-U2 oligonucleotide 5' CAG ATA CTA CAC TTG 3' prior to pre-mRNA addition. Yeast snRNP complexes were resolved on 3 or 4% polyacrylamide-0.5% agarose gels as previously described (26), except that the heparin (Sigma) concentration was increased to 1 mg/ml. The samples were incubated on ice for 10 min prior to electrophoresis. Where antibodies were used for gel shift analysis, 1 μ l of the HA.11 antibody or MAb63 was preincubated with the yeast extract at room temperature for 10 min prior to the addition of R buffer (50 mM HEPES [pH 7.5], 2.0 mM magnesium acetate, 20 mM EDTA, 1 mg of heparin per ml).

Glycerol gradient fractionation. Prior to fractionation of the yeast snRNPs, 1 volume of yeast extract was mixed with 2 volumes of buffer A (50 mM Tris-HCl [pH 7.4], 25 mM NaCl, 5 mM MgCl₂). The samples were applied to the top of a 10 to 30% linear glycerol gradient made in buffer A and centrifuged for 14 h at 37,000 rpm in a Beckman SW41 rotor. The positions of the various snRNP complexes were established by recovery of RNA from sequential 500- μ l samples from the top of the gradient. Each gradient fraction was phenol extracted, and the RNA was precipitated with ethanol. The recovered RNAs and control samples were fractionated on a 5% polyacrylamide-7 M urea denaturing gel and assayed by Northern blotting for snRNAs as described previously (6). The relative intensities of the snRNA bands were determined with an LKB model 2222-010 Ultrascan XL laser densitometer.

RESULTS

Identification of a novel protein, Prp42, with sequence similarity to Prp39p. A BLASTP search (3) of the *S. cerevisiae* nonredundant protein database revealed a previously uncharacterized protein of 544 amino acids, YDR235w (henceforth called Prp42p), that shares 50% sequence similarity with Prp39p (Fig. 1). The protein sequence alignment shows that the regions of Prp39p-Prp42p similarity are scattered throughout the protein coding sequences, with somewhat lower levels of sequence identity at the extreme amino and carboxyl termini. Many of the Prp39p-Prp42p sequence identities are clustered in contexts reminiscent of the TPRs of the *Drosophila* crn protein (52) (Fig. 2). A global BLAST search of the public protein databases highlighted this fact, because both Prp39p and Prp42p produced matches to the *Drosophila* crn protein, and, in each case, the similarities were restricted to the TPR elements (unpublished observations). Yeast repeats 42.1 to 42.4 and 39.1 to 39.5 show the greatest similarities to the *Drosophila* crn TPRs; each aligned with one or more of the crn repeats 3, 4, 7, 8, 10, 11, 12, 15, and 16 (see reference 52 for sequence definitions of the numbered repeats). The Prp39p-Prp42p TPR elements match best when arranged in register (e.g., 39.1 with 42.1 and 39.2 with 42.2). This pattern is most obvious for the first three TPR elements because the alignment based on maximum similarity places the TPRs in phase with one another. Elsewhere the Prp39p and Prp42p proteins appear to have evolved to such a degree that the TPR elements no longer align.

In addition to the proposed TPRs, a possible nuclear localization sequence, KKKLKK, is present at amino acids 231 to 235 of the Prp42p protein. This element may not be conserved within Prp39p, because none of the clustered basic residues of Prp39p match closely the canonical simian virus 40 T antigen or nucleoporin bipartite nuclear localization sequence motifs. No other strong matches to known functional sequence motifs were observed in Prp42p.

PRP42 encodes a protein required for pre-mRNA splicing in vivo. The structural similarity between Prp42p and the U1 snRNP protein Prp39p suggested that Prp42p might also function in pre-mRNA splicing. Since *PRP39* is essential in yeast

```

Prp39p 1 MPDETNPFTIEDIEPRFDALRGLDTQFLQDNTALVQA..YRGLDWSDISL 48
Prp42p 1 .....MDKYTALIHDFNSTLTLNVRSY 23
Prp39p 49 TQMVDVIEQTVVKYGNPNDSIKLALETILWQILR.....KYPLLFGP 90
Prp42p 24 PKSLAYWEKLLNVIKASAPICKSTEPQLLKLIRCTYSMLNEFPYLYN 73
Prp39p 91 WKRFATIEYQLFGLKKSITAVLATSVK.WFPTSLLEWCDYLNVLVCVNNPNE 139
Prp42p 74 YIDFALLEYKLGNVSMCHKIFQRGLQAFNQRSLLLWTSYLK.FCNNVISH 122
Prp39p 140 TDFIRNRFIAKDLIGKQLSHPFWDKFIIEFVGG...KNHWNVQRIYE 185
Prp42p 123 QKQLFKKYETAEEYVGLHFFSGEFWDLYLEQISSRCTSSKKYWNVLRK.. 170
Prp39p 186 YIIIEVPLHQYARFFTSYKFLNEKNLKTTRNIDIVLRKKTQTTVNEIWFQE 235
Prp42p 171 .ILEIPLHSFSKPYALW.....LQRIDDIM....DLKQLSOLT 203
Prp39p 236 SKIKQPFNGLQVLLNDDLENWSRYLKFVTDPSKSLDK...EFVMSVDFR 281
Prp42p 204 SKD....ELKKLRIDINYSGRKGPYLQDAKKLLKIKTKEMYVMVQYQV 248
Prp39p 282 CLIPFHYENTWMMYIKWLLKKNISDEVV..VDIYQKANTFLEDFKRTL 329
Prp42p 249 LEIYSIFESKIYINYYSPETLVSSDIEIETWIKYLDYTIITLQDLSLTHLN 298
Prp39p 330 YD.....FLRFLRKYRSNNTLFNINIFNETVSRYLKIWPNIDIL 367
Prp42p 299 FQALLPLAHYDLVWIKYSKWLINSKNDLLG.....AKNVL 334
Prp39p 368 LMTEYLCMLKRHSFK..NSLDQSPKEILEKQTSFTKILETSITNYINQI 415
Prp42p 335 LMGLKPSLKKTEIIRLLYSVICKLNEVLLRNLLKKT.ESSYSNDVENV 383
Prp39p 416 DAKV....HLQTLINDKNLSIVVVELIKTTLWLVLKNMQRKRYFNLYQK 460
Prp42p 384 DFEIFWDYLFQKFTQCNLSYSSRYSDSQSNGLLNKELFDKVKRLSCKEK 433
Prp39p 461 ....NILKNSVFFWLTYYKFEKSNVNFTRLNKF..IRELGVIEYLPTTV 504
Prp42p 434 KSGQEILNLNLVQF.....YSKDTVEFVEKNIFQKIIEFGWGYLQNGM 477
Prp39p 505 MNDILTQYKFTYLLTHSNIVTYESSIIDSNTFDPILYPELKMNSPKYDVP 554
Prp42p 478 .....FWNCYCRILYFDTS..... 491
Prp39p 555 NTTANVDWHKKTTEWKBAGHIGITERRPQISNSIIECNSGTLIQKPISLPN 604
Prp42p 492 .....RSYLDKRQYIVRKIWPQIDKKFPAQSVLPSLTF..FCBSY 528
Prp39p 605 FRNLEKINQVKINDLYTEBFLKEGK 629
Prp42p 529 FPE....EMDTLEEMFTEEP..... 544

```

FIG. 1. Comparison of the Prp39p and Prp42p protein sequences. An amino acid alignment between Prp39p and ORF YDR235w (i.e., Prp42p) was performed with the BestFit program (GCG, Inc.) to show the locations of sequence identities (vertical lines) and conservative amino acid substitutions (colons, strong similarity; single dots, moderate similarity). The positions of the putative TPRs are overlined (Prp39p) and underlined (Prp42p).

(26), it seemed unlikely that *PRP42* acted simply as an alternate source of the Prp39p activity. Consistent with an independent cellular function, a genomic disruption of the *PRP42* ORF by the yeast *LEU2* gene proved lethal (see Materials and Methods). The *prp42::LEU2* mutation was complemented by a plasmid-borne *PRP42* allele in which the natural *PRP42* promoter was replaced by the nutritionally regulated *GAL1* promoter. The *GAL1::PRP42* colonies were not quite as large as those from wild-type yeast when assayed on galactose medium. While the reason for this discrepancy is not known, the viability of the *GAL1::PRP42* yeast clearly required expression of the fusion gene. Transfer of this strain to glucose-based medium repressed *GAL1::PRP42* transcription and blocked colony formation (Fig. 3A). These experiments established that *PRP42* is an essential gene and provided a valuable genetic tool with which to experimentally manipulate the intracellular levels of Prp42p.

The intron-containing *ACT1*, *RP51A*, *CYH2*, and *SNR17* genes were assayed to determine the impact of *GAL1::PRP42*

```

1 34
39.1 GFWKRFATIEYQLFGLKKSITAVLATSVKWFPTSL 39.6
39.2 EFLWCDYLNVLVCVNNPNETDFERNFEEAKDHTG 42.1
39.3 HFPWDFLEBFEGGKKNHNVQRIYEVYDVPVHQ 42.2
39.4 HNNWSEYLFKFTDPSKSLKEKEMVSVLDRCTPC 42.3
39.5 ENTWMMYIKWLLKKNISDEVVVDIYQKANTFPL 42.4
39.6 WPFWLTYYKFEKSNVNFTRLNKFTRELGVEIYLP 42.5
42.1 ENYIEVPLHQYARFFTSYKFLNEKNLKTTRNIDIV 42.1
42.2 HLLWTSYLKFCNNVISHQKQKPKKYETAEEYVGL 42.2
42.3 GEFWDLYLEQISSRCTSSKKYWNVLRKLEIPH 42.3
42.4 HETWIKYLDYTIITLQDLSLTHLN 42.4
42.5 DLVWIKYKYSKWLINSKNDLLGAKNVLMLGLKPSL 42.5
Consensus LEFWIKYLFKFE--L-NSSK-HKN*-RKILEFLLL
: : : | | | : | | : : : : : : : : :
crn VKLWIKYAREEELLKEIDRAREIYERALEFLPRD
TPR W LG Y A F A P

```

FIG. 2. Alignment of 11 proposed TPRs within Prp39p and Prp42p. The repeats are numbered consecutively from the amino termini of the Prp39p (39.1 to 39.6) and Prp42p (42.1 to 42.5) proteins presented in Fig. 1. Columns with greater than 40% (light shading) or 70% (dark shading) amino acid sequence similarity among the Prp39p-Prp42p repeats are boxed. The amino acid residues were grouped as polar charged (D, E, H, K, and R), polar uncharged (N, Q, S, and T), nonpolar hydrophobic (L, I, V, M, F, Y, and W), small (A and G), or other (C and P). The Prp39p-Prp42p consensus sequence was assembled by BestFit and Pretty analyses (University of Wisconsin, GCG program). In many but not all instances, the calculated consensus residue was the most common amino acid within a boxed group. The asterisk indicates the position of the conserved hydrophobic residue at position 24. The vertical lines indicate sequence identity, and the colons and single dots represent degrees of conservative substitutions. Underlined in the crn sequence (52) are the box A and box B elements. Below the crn sequence is a more generalized repeat present in a diverse set of TPR proteins that consists of seven hydrophobic residues and a proline (20).

transcriptional repression on the efficiency of *in vivo* splicing (Fig. 3B and data not shown). RNA was extracted from wild-type (*PRP42*) and *GAL1::PRP42* cultures grown continuously on galactose medium and from the same cultures after transfer to glucose-based medium. Yeast that expressed the *GAL1::PRP42* gene showed an mRNA/pre-mRNA ratio indistinguishable from that observed with the wild-type allele. By this measure, the *GAL1::PRP42* allele functioned as well as the endogenous chromosomal allele in pre-mRNA splicing. In contrast, splicing arrest resulted when *GAL1::PRP42* transcription was repressed; the mRNA/pre-mRNA ratios of the interrupted genes decreased as a function of time in the glucose-based medium. Elevated pre-mRNA levels became apparent approximately 8 h after the shift of the media, while cell division slowed and then stopped some 10 to 15 h later. No differences in RNA mobilities were observed for transcripts from the intronless *ADE3*, *SNR19*, *SNR20*, and rDNA genes (Fig. 3B and data not shown), indicating that Prp42p is required specifically for the processing of intron bearing pre-mRNA.

Prp42p is a U1 snRNP-associated protein. To assist in the characterization of Prp42p, the nine-codon HA epitope sequence was added to the 3' end of the *GAL1::PRP42* ORF. This derivative, *GAL1::PRP42HA*, retained biological activity, because it complemented the growth and splicing defects of the chromosomal *prp42::LEU2* null allele. Extracts prepared from the *GAL1::PRP42HA* culture and from control cultures were assayed for coprecipitation of spliceosomal snRNAs with the HA-tagged proteins (Fig. 4). U1 snRNA was specifically recovered from the anti-HA (HA.11) immune pellets of the epitope-tagged Prp42HAp and Prp39HAp extracts (Fig. 4, lanes 5 and 7) (26). The coprecipitation of U1 snRNA with Prp42HAp was rather salt sensitive; peak U1 snRNA recovery was achieved below 100 mM NaCl (Fig. 4, lanes 2 to 4). The U1 snRNA recovery was epitope specific, however, because no detectable U1 snRNA coprecipitated at 100 mM NaCl from

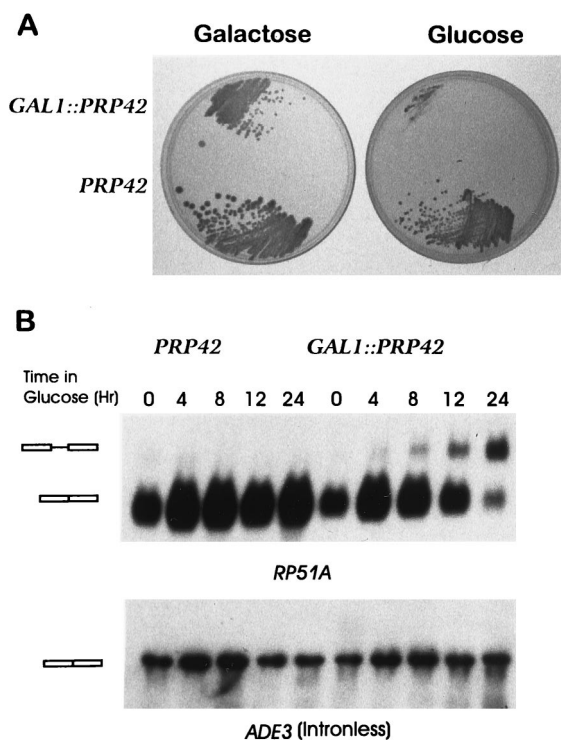


FIG. 3. *PRP42* encodes a protein required for cellular viability and pre-mRNA splicing. (A) To assay for *PRP42*-dependent growth, yeast strains that expressed the endogenous wild-type copy of *PRP42* or the glucose-repressible *GAL1::PRP42* fusion gene were plated on galactose-based and glucose-based rich media. (B) RNA was extracted from the *PRP42* and *GAL1::PRP42* cultures grown continuously on galactose medium (time 0) or after a shift to glucose medium for the indicated number of hours. The Northern transfer in the upper panel was hybridized with the intron-containing *RP51A* gene. The identities of the mRNA and pre-mRNA (rectangles) were confirmed by primer extension analysis (data not shown). The lower panel presents the same filter after hybridization with the intronless *ADE3* gene.

the extract metabolically depleted of Prp42HAp or from an extract that lacked an HA-tagged protein (Fig. 4, lanes 6 and 8). Coprecipitation was also antibody specific, because only background levels of U1 snRNA were recovered from the *GAL1::PRP42HA* extract when the irrelevant MAb63 was substituted for HA.11 (reference 26 and unpublished observations). Because the U1 snRNA levels varied less than twofold in all extracts tested (see below), the failure to recover U1 snRNA with the HA.11 antibody after Prp42HAp depletion was not a consequence of extensive U1 snRNA degradation. Rather, the recovery of U1 snRNA from *GAL1::PRP42HA* extracts correlated directly with the presence or absence of the Prp42HAp protein.

As a more direct means of viewing the Prp42HAp-U1 snRNP association, the *GAL1::PRP42HA* extract was resolved by native gel electrophoresis with or without prior preincubation with the HA.11 antibody (Fig. 5A). In the absence of HA.11, we routinely observed two distinct U1 snRNP bands upon hybridization (Fig. 5A) (26). The upper band was previously deduced to be the mature form of the U1 snRNP (26). In contrast, the lower band appeared to be an incomplete U1 snRNP, since it was formerly shown to migrate very near the position of naked U1 snRNA and it lacked the U1 snRNP-specific protein, Prp39p (26). The incomplete U1 snRNP band was also detected in wild-type yeast extracts, but the abundance of the incomplete form increased in the epitope-tagged

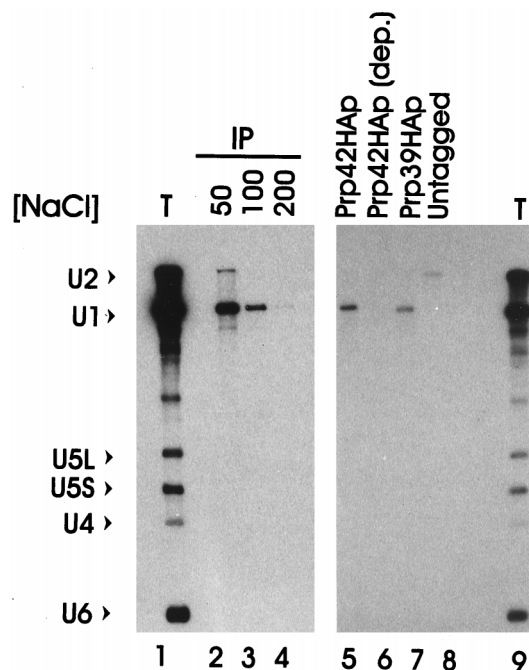


FIG. 4. Prp42HAp specifically associates with the U1 snRNP. A splicing-competent yeast extract prepared from the *GAL1::PRP42HA* yeast strain was immunoprecipitated with the HA-specific antibody HA.11 under conditions of increasing NaCl concentration (lanes 2 to 4). RNA present in the immune pellets (IP) or the unfractionated total extract (T) was analyzed by Northern blotting with probes specific for the spliceosomal snRNAs (U1, U2, U4, U5, and U6). In lanes 5 to 9, the immune precipitation was repeated at 100 mM NaCl with the splicing-competent *GAL1::PRP42HA* extract (lane 5), an extract prepared from *GAL1::PRP42HA* cultures metabolically depleted of Prp42HAp (Prp42HAp dep.), an extract with an HA-tagged *PRP39* allele (Prp39HAp), and an extract without an HA-tagged gene (Untagged).

Prp39HAp and Prp42HAp extracts. This observation suggests that the HA epitope reduces the stability or inhibits the function of the tagged proteins. However, the negative impact of epitope addition is clearly limited, because the Prp39HAp and Prp42HAp extracts both contain fully assembled U1 snRNPs and are splicing competent. When the HA.11 antibody was added to the *GAL1::PRP42HA* extract, 100% of the upper U1 band was shifted (Fig. 5A), consistent with Prp42HAp being present in each of the fully assembled U1 snRNPs. Approximately 50% of the lower U1 band was shifted by the HA.11 antibody. We interpret this to signify heterogeneity within the incompletely assembled U1 particles (i.e., some possessed the Prp42HAp protein and therefore shifted, while others either lacked Prp42HAp or had it sequestered in an antibody-inaccessible location). The U1 snRNP complexes did not change mobility when the HA.11 antibody was added to an untagged extract or when the irrelevant MAb63 was substituted for HA.11 in the *GAL1::PRP42HA* extract (reference 26 and unpublished observations).

In contrast to the U1 hybridization results, none of the U2, U4, U5, or U6 snRNP complexes showed decreased electrophoretic mobility after HA.11 antibody addition. Occasionally, changes in the relative levels of the U4-U6 versus U4-U6.U5 snRNP complexes were noted after antibody addition (Fig. 5A, asterisks, compare U4 and U6 hybridizations). This appeared to be a nonspecific consequence of the antibody preparation and was independent of the presence of an epitope-tagged protein in the extract (26a).

Metabolic depletion of Prp42HAp caused a profound, yet

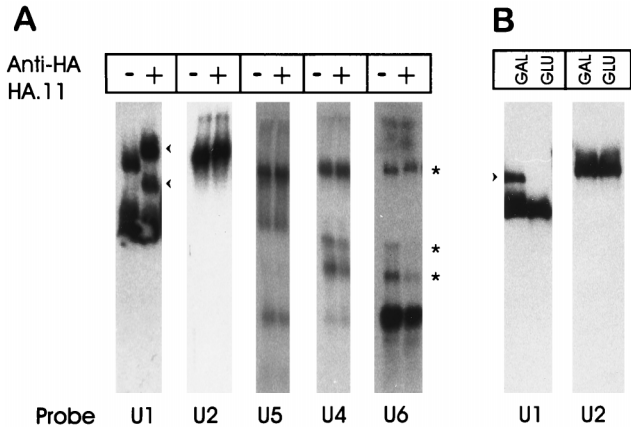


FIG. 5. Prp42p contributes directly to U1 snRNP structure. (A) Splicing-competent *GALI::PRP42HA* extracts were incubated with (+) or without (-) the HA.11 antibody and then resolved on a 3% polyacrylamide-0.5% agarose gel. After electrophoresis, parallel lanes of the Northern transfer were probed for the U1, U2, U4, U5, and U6 snRNA-bearing snRNP complexes. The arrowheads to the right of the U1 snRNA hybridization show the positions of the novel bands present after incubation with the HA.11 antibody. The asterisks to the right of the U6 hybridization identify the location of the U4-U6.U5 tri-snRNP (top band) and two distinct U4-U6-hybridizing bands (based on comigration of the U4, U5, and U6 snRNA signals). (B) *GALI::PRP42HA* extracts prepared before (GAL) and after (GLU) depletion of Prp42HAp were resolved by native gel electrophoresis and probed for the U1 and U2 snRNA-bearing complexes. The arrowhead to the left of the U1 hybridization indicates the position of the fully assembled U1 snRNP.

specific, increase in the electrophoretic mobility of the U1 snRNP and the loss of splicing competence (Fig. 5B and see below). Although incomplete, the Prp42HAp-depleted extract was not otherwise denatured and could be functionally reconstituted by the addition of a micrococcal nuclease-treated (i.e., RNA free) wild-type extract (unpublished observations). No shift in electrophoretic mobility was observed with the U2 snRNP or any other snRNP complex after Prp42HAp depletion (Fig. 5B and data not shown).

Since the magnitudes of electrophoretic shifts after subunit removal were not necessarily predictable, glycerol gradient sedimentation was performed as an independent assay for Prp42HAp-dependent changes in snRNP structure (Fig. 6). A reproducible reduction in the U1 snRNA signal was observed for the fastest sedimenting and presumably most complete U1 snRNP particles after Prp42HAp depletion (compare lanes 13 through 21, Fig. 6A and B). In contrast, the relative amounts and positions of the U4 and U5 snRNAs did not vary significantly between these two parallel gradient preparations. Densitometric scans showed an approximate 70% reduction in the ratio of the U1 to U5L plus U5S snRNA signals in the heaviest gradient fractions after Prp42HAp depletion (averaged ratio of fractions 19 to 21 = 1.1 in Fig. 6A and 0.3 in Fig. 6B). (Note that fraction 23 contains nonspecific aggregates from the tube bottom.) The U1/U5L plus U5S ratio increased roughly two-fold for the bulk of the remaining U1 snRNP fractions (averaged ratio of fractions 13 to 17 = 2.0 in Fig. 6A and 3.9 in Fig. 6B), which is consistent with the recruitment of the Prp42HAp-deficient particles into the lighter U1 snRNP fractions.

The U1 snRNA decreased by ~35% and the U6 snRNA decreased by ~50% after Prp42HAp depletion (compare Fig. 6A and B). SnRNA reductions have been reported after the inactivation or removal of certain other spliceosomal proteins (for examples, see references 6, 32, and 33) and appear to be caused by enhanced snRNA degradation upon snRNP perturbation or splicing arrest. The decrease in U1 snRNA could be accounted for by modestly decreased U1 snRNA stability in the absence of Prp42HAp. The reduction in U6 snRNA is less easily explained but likely occurs as an indirect consequence of the splicing arrest. The free U6 snRNP pool (i.e., the lightest gradient fractions) decreased most dramatically after Prp42HAp depletion, suggesting that the higher-order U4-U6 and U4-U6.U5 snRNP complexes stabilize the U6 snRNA.

Pre-mRNA-U1 snRNP complex formation is impaired by Prp42p depletion. The molecular contacts that secure the 5' splice site-U1 snRNP interaction in the commitment complex are poorly understood. Since the Prp42HAp-depleted U1 snRNP particles were stable, the formation of a pre-mRNA-U1

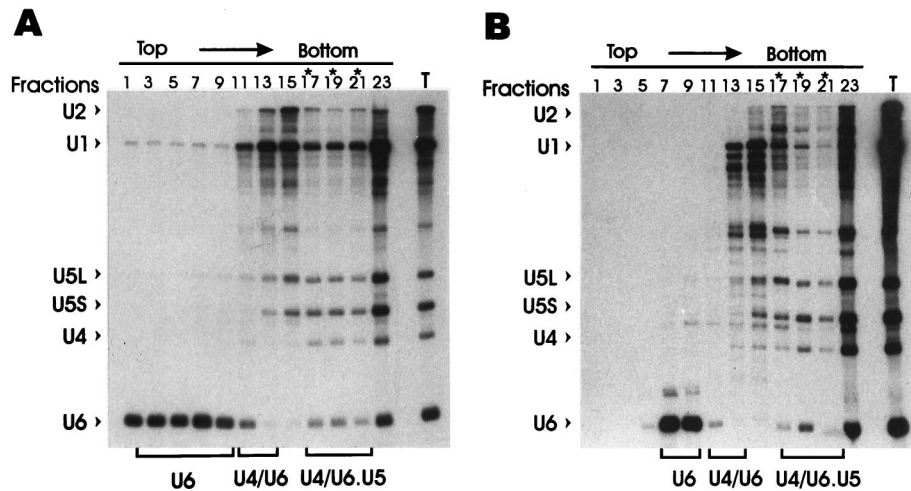


FIG. 6. Sedimentation of the U1 snRNP is altered after Prp42HAp depletion. *GALI::PRP42HA* extracts prepared before (A) and after (B) metabolic depletion of Prp42HAp were fractionated on a 10 to 30% linear glycerol gradient. The snRNAs present in alternate gradient fractions were assayed by Northern blotting with a hybridization cocktail which probed for each of the spliceosomal snRNAs. The U1 snRNP fractions most sensitive to Prp42HAp depletion are indicated by asterisks. The location of the free U6 snRNP, the U4-U6 snRNP, and the U4-U6.U5 tri-snRNP are indicated below the figure. In both panels, sample 23 comes from the bottom of the gradient tube and contains all of the spliceosomal snRNAs, presumably, from native spliceosomes or nonspecific aggregates. T, total, unfractionated RNA.

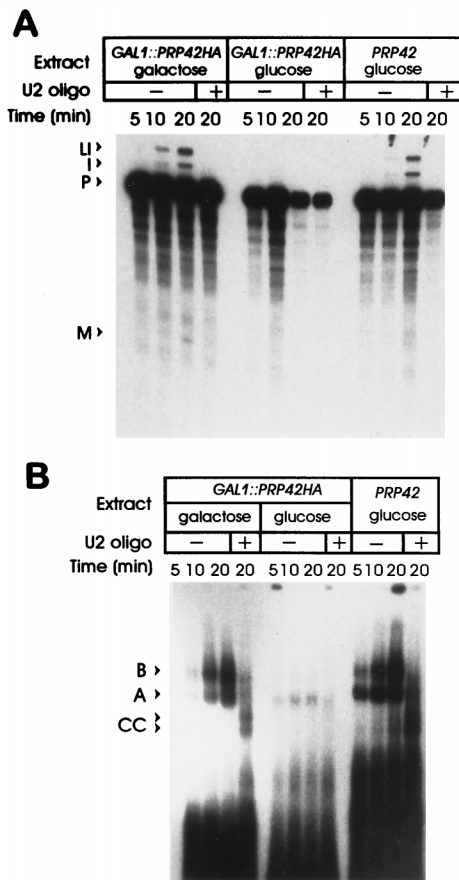


FIG. 7. Spliceosome assembly is blocked by Prp42HAp removal. (A) 32 P-labeled RP51A pre-mRNA was incubated for the indicated times in wild-type extracts (*PRP42* glucose) and in extracts prepared before (*GAL1::PRP42HA* galactose) and after (*GAL1::PRP42HA* glucose) Prp42HAp depletion. The pre-mRNA (P) and spliced RNA products (lariat intermediate [LI], excised intron [I], and mRNA [M]) were then resolved on a 5% polyacrylamide denaturing gel. In some cases (+), an oligonucleotide directed against the 5' end of the U2 snRNA was included to block splicing and arrest spliceosome assembly at the commitment complex stage. (B) A portion of each splicing reaction was resolved in parallel on a 3% polyacrylamide-0.5% agarose gel to monitor spliceosome assembly. The commitment complexes (CC1, upper band; CC2, lower band) and assignments of the prespliceosome (A) and spliceosome (B) forms indicated by the arrowheads were based on our (26, 30) and others' (36) previous characterizations of these complexes. The lane descriptions are the same as those indicated for panel A.

snRNP interaction was credible, despite possible defects in other aspects of U1 snRNP function. In vitro spliceosome assembly and splicing were monitored in the Prp42HAp-complete and Prp42HAp-depleted extracts (Fig. 7). Consistent with the in vivo results, *RP51A* pre-mRNA assembled into prespliceosome and spliceosomal complexes in the Prp42HAp-complete extract (Fig. 7B, *GAL1::PRP42HA* galactose lanes), and the levels of the splicing intermediates and products in the reaction increased with time of incubation (Fig. 7A). When an oligonucleotide directed against the 5' end of the U2 snRNA was preincubated with the Prp42HAp complete extract prior to substrate addition, splicing was blocked and spliceosome assembly was arrested at the commitment complex stage. Roughly equivalent amounts of the 5' splice site-dependent (CC1 [Fig. 7B, upper arrowhead]) and the 5' splice site and branch point-dependent (CC2 [Fig. 7B, lower arrowhead]) commitment complex bands were observed (Fig. 7) (36). Anal-

ogous results were obtained with a wild-type extract prepared from glucose-grown cultures (*PRP42* glucose). In contrast, when pre-mRNA was added to the Prp42HAp-depleted extract (*GAL1::PRP42HA* glucose), a greatly reduced level of spliceosome assembly and no splicing were observed. The low level of splicing complex (mostly prespliceosome) observed was likely due to trace amounts of Prp42HAp remaining in the extract, although inefficient spliceosome assembly in the absence of Prp42HAp cannot be ruled out.

Prominent commitment complex bands were not observed in the Prp42HAp-depleted extracts in either the presence or absence of the U2 oligonucleotide. Since the mobility of the U1 snRNP is changed upon Prp42HAp depletion, it is possible that functional commitment complex 1-commitment complex 2-like complexes formed but did not resolve from the bulk uncomplexed pre-mRNA. This appears unlikely, however, since only a minor amount of spliced pre-mRNA (presumably derived from the residual prespliceosomal complexes) was observed when commitment complex formation was assayed by the substrate challenge method (23) rather than by gel electrophoresis (unpublished observations). These results are consistent with a spliceosome assembly defect prior to stable U1 snRNP addition but do not rule out the possibility that specific, yet weak or transient, U1 snRNP-pre-mRNA interactions occur in the absence of Prp42p.

DISCUSSION

Approximately 10^8 years ago, *S. cerevisiae* is believed to have undergone a whole genome duplication (51). This ancient duplication and subsequent genetic events (e.g., deletions and single-gene amplifications) left the modern yeast with approximately 800 pairs of duplicated genes. In many instances, members of individual gene pairs have evolved distinct yet related biological functions. With this in mind, we used the amino acid sequence of the Prp39p protein to search for and identify a novel component of the yeast U1 snRNP, Prp42p. Prp42p shares 50% amino acid sequence similarity with Prp39p and, like Prp39p, contains multiple copies of the TPR protein recognition domain. The results of this study show that the *PRP39* and *PRP42* genes are not functionally redundant; each protein coded for by these genes is required at an early stage of splicing to assemble a splicing-competent U1 snRNP. The identification of structurally related, snRNP-specific proteins is unprecedented and indicates that the U1 snRNP is more complex than originally envisioned.

Several lines of evidence implicate Prp42p as a legitimate component of the U1 snRNP. First, the structure of Prp42p is quite similar to that of Prp39p, a protein shown genetically and biochemically to be part of the U1 snRNP complex (26, 28). Second, an antibody directed against an epitope-tagged version of Prp42p specifically coprecipitates U1 snRNA and supershifts the U1 snRNP in native polyacrylamide gels. Third, the U1 snRNP in extracts depleted of Prp42HAp shows increased electrophoretic mobility and a decreased sedimentation rate, indicative of a change in U1 snRNP structure or stability. Finally, certain mutations within Prp42p are synthetically lethal with a biologically active U1 snRNA deletion derivative (30a), again consistent with an intimate association between Prp42p and the U1 snRNP.

The absence of Prp42p from a biochemically purified preparation of the yeast U1 snRNP (28) can possibly be resolved by the observation that the Prp42HAp-U1 snRNP interaction is quite salt sensitive. If the native protein behaves like Prp42HAp, then the 300 mM salt washes used during affinity purification would strip the Prp42p from the U1 snRNP. This

apparently weak interaction might reflect a mainly protein-based contact between Prp42p and the U1 snRNP. The snRNP association of yeast U1-C with the U1 snRNP is likely protein mediated and is nearly as salt sensitive as Prp42p (47). In contrast, the U1-snRNA-binding proteins Snp1p and Mud1p remain bound to the snRNP at NaCl concentrations of at least 200 mM (Mud1p [25]) and 500 mM (Snp1p [21]). Whatever the mode of interaction, all available data are consistent with an essential, although possibly salt-sensitive, interaction of Prp42p with the U1 snRNP.

It seems likely that Prp39p and Prp42p are present in the same particle and do not substitute for one another in alternative forms of the U1 snRNP with distinct properties (e.g., distinct substrate preferences). The transcripts of every interrupted gene assayed (i.e., *ACT*, *CYH2*, *RP51A*, and *SNR17*) required both Prp39p and Prp42p to be spliced *in vivo*. In addition, the gel shifts associated with anti-HA antibody addition and with Prp42HAp depletion included all of the fully assembled U1 snRNPs independent of whether Prp39HAp or Prp42HAp was being manipulated. If Prp39HAp and Prp42HAp were present in different U1 snRNP populations, these experiments should have revealed a pre-mRNA or snRNP subset insensitive to the manipulation of one or the other factor. Thus, although we have not directly assayed for Prp39p in Prp42p-bearing complexes, no precedent for functionally distinct U1 snRNP subpopulations exists in yeast, and the available data all support Prp39p-Prp42p colocalization in the U1 snRNP.

A consensus element built from the TPR sequences observed in Prp39p and Prp42p is most closely related to that observed in the *Drosophila* crn protein (Fig. 2). The level of sequence match to the TPR consensus element for the individual Prp39p and Prp42p repeats is low but not unlike that observed for many other TPR elements (10, 11). In all cases, the Prp39p and Prp42p TPR elements are predicted to present amphipathic alpha-helical surfaces characteristic of the TPR. TPR proteins can be grouped into distinct subfamilies based on their distinctive primary structure features (10, 11, 39). For instance, domain A of the *Drosophila* crn repeats generally has charged residues at positions 6 and 9, aromatic residues at positions 7 and 10, and an aspartic acid residue at position 11. Other than a reduced prevalence of the position 11 aspartic acid, each of these amino acids is well conserved within the Prp39p and Prp42p repeats. Other TPR proteins, including the one other known TPR protein of the spliceosome, Prp6p (22), show different distributions of amino acids in these positions (10, 11, 39). The remaining positions from amino acid 1 through the end of domain A of the Prp39p-Prp42p repeats are mostly perfect matches or conservative substitutions for the crn consensus sequence. The domain B structure is less conserved in primary sequence, but individual repeats generally match the crn or an elaborated TPR consensus (52) at multiple positions and show the overall alpha-helical character of the TPR. In 7 of 11 cases, the predicted domain B alpha helix terminates with a characteristic proline residue (10, 11, 39) located with a more relaxed positioning between TPR residues 30 and 34.

A number of polypeptide targets for TPR interaction have been identified. For instance, the TPR region of the yeast Cyc8p protein binds to the homeodomain of the yeast $\alpha 2$ protein (43, 48), while the TPRs of yeast Cdc23p bind to a helix-loop-helix region of Sin1p (38). In addition, a mutation in the TPR region of the Cdc27p protein reduces its ability to associate with the Cdc23p TPR protein, suggesting a possible TPR-TPR interaction (20). While these target polypeptides differ considerably in primary sequence, each presents a helical

surface for interaction with the corresponding TPR helices. The unique sequence characteristics of the Prp39p and Prp42p repeats (as well as those of other TPR proteins) likely reflect distinctions in the complementary surfaces of their interacting ligands. Genetic screens and direct protein assays are currently under way to identify the natural ligands of Prp39p and Prp42p interaction.

Database comparisons revealed multiple TPRs of the Prp39p-Prp42p-crn domain A sort in only one other protein set, the yeast Rna14p, *Drosophila* Su(f), and human CstF77 proteins, in which six to eight repeats were present (this study and reference 50a). These proteins are part of the RNA cleavage stimulation factor required for pre-mRNA 3' end processing (see reference 46 and references within). Thus, with the possible exception of crn, all of the known proteins with this variant TPR are involved in RNA processing. Mutations in the *Drosophila* crn gene show a pleiotropic embryonic lethal phenotype with impaired neurological (52) and muscle (7) development. Based in part on the reduced DNA synthesis observed in mutant embryos, it was suggested that crn may be involved in the cell cycle. We have cloned the likely yeast homolog of crn and are currently investigating its activity in the yeast cell cycle and RNA processing (6a). A role for yeast crn in splicing or polyadenylation would support the view that the crn-like TPR is reserved for the assembly or intracellular trafficking of complexes involved in RNA maturation.

Prp39p and Prp42p are essential in yeast, while mammalian homologs have not been found. Either these proteins provide a truly yeast-specific function or they are present in mammals but lost in the U1 snRNP isolation procedures used to date. In the first case, the Prp39p and Prp42p proteins might substitute, for instance, for one or more of the mammalian SR proteins absent in yeast but required for early spliceosome assembly events in mammals (e.g., ASF/SF2 and SC35 [reviewed in references 9 and 17]). However, given the general high level of conservation observed between the yeast and mammalian basal spliceosomal components (17, 27), perhaps a more likely view is that the mammalian equivalents of Prp39p and Prp42p exist but are tenuously associated with the U1 snRNP. The initial failure to identify the phylogenetically conserved, yet weakly bound SF3a and SF3b proteins with the U2 snRNP offers precedent for this (see references in references 17 and 27). The identification of mammalian Prp39p and Prp42p equivalents would reveal the yeast U1 snRNP as more conserved than is suggested by its exaggerated snRNA length.

ACKNOWLEDGMENTS

We thank Charles Query, Martha Peterson, John Woolford, and our laboratory colleagues for helpful comments on the manuscript. Carol Williams and Kevin O'Hare are thanked for pointing out the TPRs within Su(f). Seyung Chung is gratefully acknowledged for assistance with the Prp39p-Prp42p TPR alignments, and Elizabeth Otte is acknowledged for help with the quantitative analysis of the snRNA levels.

This work was supported by grant GM42476 from the National Institutes of Health to B.C.R.

REFERENCES

1. Abovich, N., X. C. Liao, and M. Rosbash. 1994. The yeast MUD2 protein: an interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. *Genes Dev.* 8:843-854.
2. Abovich, N., and M. Rosbash. 1997. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* 89:403-412.
3. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403-410.
4. Arning, S., P. Gruter, G. Bilbe, and A. Krämer. 1996. Mammalian splicing factor SF1 is encoded by variant cDNAs and binds to RNA. *RNA* 2:794-810.
5. Berglund, J. A., K. Chua, N. Abovich, R. Reed, and M. Rosbash. 1997. The

- specificity factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* **89**:781–787.
6. **Blanton, S., A. Srinivasan, and B. C. Rymond.** 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.
 - 6a. **Chung, S., and B. C. Rymond.** Unpublished observations.
 7. **Drysdale, R., E. Rushton, and M. Bate.** 1993. Genes for embryonic muscle development in *Drosophila melanogaster*. *Roux's Arch. Dev. Biol.* **202**:276–295.
 8. **Fromont-Racine, M., J. C. Rain, and P. Legrain.** 1997. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat. Genet.* **16**:277–282.
 9. **Fu, X.-D.** 1995. The superfamily of arginine/serine-rich splicing factors. *RNA* **1**:663–680.
 10. **Gindhart, J. G., Jr., and L. S. Goldstein.** 1996. Tetratricopeptide repeats are present in the kinesin light chain. *Trends Biochem. Sci.* **21**:52–53.
 11. **Goebel, M., and M. Yanagida.** 1991. The TPR snap helix: a novel protein repeat motif from mitosis to transcription. *Trends Biochem. Sci.* **16**:173–177.
 12. **Guthrie, C.** 1991. Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science* **253**:157–163.
 13. **Habets, W. J., P. T. Sillekens, M. H. Hoet, J. A. Schalken, A. J. M. Roebroek, J. A. M. Leuissen, W. J. M. van de Ven, and W. J. Venrooij.** 1987. Analysis of a cDNA clone expressing a human autoimmune antigen. Full length sequence of the U2 small nuclear RNA-associated B' antigen. *Proc. Natl. Acad. Sci. USA* **84**:2421–2425.
 14. **Hermann, H., P. Fabrizio, V. A. Raker, K. Fouaki, H. Hornig, H. Brahm, and R. Lührmann.** 1995. SnRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J.* **14**:2076–2088.
 15. **Johnston, M., and R. W. Davis.** 1984. Sequences that regulate the divergent *GAL1-GAL10* promoter in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **4**:1440–1448.
 16. **Kao, H. Y., and P. G. Siliciano.** 1996. Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. *Mol. Cell. Biol.* **16**:960–967.
 17. **Krämer, A.** 1996. The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annu. Rev. Biochem.* **65**:367–409.
 18. **Kretzner, L., A. Krol, and M. Rosbash.** 1990. *Saccharomyces cerevisiae* U1 small nuclear RNA secondary structure contains both universal and yeast-specific domains. *Proc. Natl. Acad. Sci. USA* **87**:851–855.
 19. **Kretzner, L., B. C. Rymond, and M. Rosbash.** 1987. *S. cerevisiae* U1 RNA is large and has limited primary sequence homology to metazoan U1 snRNA. *Cell* **50**:593–602.
 20. **Lamb, J. R., S. Tugendreich, and P. Hieter.** 1995. Tetratricopeptide repeat interactions: to TPR or not to TPR? *Trends Biochem. Sci.* **20**:257–259.
 21. **Lauber, J., P. Fabrizio, S. Teigelkamp, W. S. Lane, E. Hartmann, and R. Lührmann.** 1996. The HeLa 200 kDa U5 snRNP-specific protein and its homolog in *Saccharomyces cerevisiae* are members of the DEXH-box protein family of putative RNA helicases. *EMBO J.* **15**:4001–4015.
 22. **Legrain, P., and A. Choulouka.** 1990. The molecular characterization of PRP6 and PRP9 yeast genes reveals a new cysteine/histidine motif common to several splicing factors. *EMBO J.* **9**:2775–2781.
 23. **Legrain, P., B. Seraphin, and M. Rosbash.** 1988. Early commitment of yeast pre-mRNA to the spliceosome pathway. *Mol. Cell. Biol.* **8**:3755–3760.
 24. **Liao, X., L. Kretzner, B. Seraphin, and M. Rosbash.** 1990. Universally conserved and yeast-specific U1 snRNA sequences are important but not essential for U1 snRNP function. *Genes Dev.* **4**:1766–1774.
 25. **Liao, X. C., J. Tang, and M. Rosbash.** 1993. An enhancer screen identifies a gene that encodes the yeast U1 snRNP A protein: implications for snRNP protein function in pre-mRNA splicing. *Genes Dev.* **7**:419–428.
 26. **Lockhart, S. R., and B. C. Rymond.** 1994. Commitment of yeast pre-mRNA to the splicing pathway requires a novel small nuclear ribonucleoprotein polypeptide, Prp39p. *Mol. Cell. Biol.* **14**:3623–3633.
 - 26a. **Lockhart, S. R., and B. C. Rymond.** Unpublished observations.
 27. **Moore, M. J., C. C. Query, and P. A. Sharp.** 1993. Splicing of precursors to mRNA by the spliceosome, p. 303–357. *In* R. F. Gesteland and J. F. Atkins (ed.), *The RNA world*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 28. **Neubauer, G., A. Gottschalk, P. Fabrizio, B. Seraphin, R. Lührmann, and M. Mann.** 1997. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proc. Natl. Acad. Sci. USA* **94**:385–390.
 29. **Pikielny, C. W., and M. Rosbash.** 1986. Specific small nuclear RNAs are associated with yeast spliceosomes. *Cell* **45**:869–877.
 30. **Pikielny, C. W., B. C. Rymond, and M. Rosbash.** 1986. Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes. *Nature* **342**:341–345.
 - 30a. **Rosbash, M.** Personal communication.
 31. **Rosbash, M., and B. Seraphin.** 1991. Who's on first? The U1 snRNP-5' splice site interaction and splicing. *Trends Biochem. Sci.* **16**:187–190.
 32. **Roy, J., B. Zheng, B. C. Rymond, and J. L. Woolford, Jr.** 1995. Structurally related but functionally distinct yeast Sm D core small nuclear ribonucleoprotein particle proteins. *Mol. Cell. Biol.* **15**:445–455.
 33. **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.
 34. **Rymond, B. C., and M. Rosbash.** 1985. Cleavage of 5' splice site and lariat formation are independent of 3' splice site in yeast mRNA splicing. *Nature* **317**:735–737.
 35. **Seraphin, 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.
 36. **Seraphin, B., and M. Rosbash.** 1991. The yeast branchpoint sequence is not required for the formation of a stable U1 snRNA-pre-mRNA complex and is recognized in the absence of U2 snRNA. *EMBO J.* **10**:1209–1216.
 37. **Sherman, F., G. P. Fink, and J. B. Hicks.** 1986. *Laboratory course manual for methods in yeast genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
 38. **Shungin, S., A. Liberzon, H. Bangio, E. Yona, and D. J. Katcoff.** 1996. Association of yeast SIN1 with the tetratricopeptide repeats of CDC23. *Proc. Natl. Acad. Sci. USA* **93**:8274–8277.
 39. **Sikorski, R. S., W. A. Michaud, J. C. Wootton, M. S. Buguski, C. Connelly, and P. Hieter.** 1991. TPR proteins as essential components of the yeast cell cycle. *Cold Spring Harbor Symp. Quant. Biol.* **56**:663–673.
 40. **Siliciano, P. G., M. H. Jones, and C. Guthrie.** 1987. *Saccharomyces cerevisiae* has a U1-like small nuclear RNA with unexpected properties. *Science* **237**:1484–1487.
 41. **Siliciano, P. G., W. J. Kivens, and C. Guthrie.** 1991. More than half of yeast U1 snRNA is dispensable for growth. *Nucleic Acids Res.* **19**:6367–6372.
 42. **Sillekens, P. T., W. Habets, R. P. Beijer, and W. J. van Venrooij.** 1987. cDNA cloning of the human U1 snRNA-associated A protein: extensive homology between U1 and U2 snRNP-specific proteins. *EMBO J.* **6**:3841–3848.
 43. **Smith, R. L., M. J. Redd, and A. D. Johnson.** 1995. The tetratricopeptide repeats of Ssn6 interact with the homeo domain of alpha 2. *Genes Dev.* **9**:2903–2910.
 44. **Smith, V., and B. G. Barrell.** 1991. Cloning of a yeast U1 snRNP 70k protein homologue: functional conservation of an RNA binding domain between humans and yeast. *EMBO J.* **10**:2627–2634.
 45. **Stultz, C. M., J. V. White, and T. F. Smith.** 1993. Structural analysis based on state-space modeling. *Protein Sci.* **2**:305–314.
 46. **Takagaki, Y., and J. L. Manley.** 1994. A polyadenylation factor subunit is the human homologue of the *Drosophila* suppressor of forked protein. *Nature* **372**:471–481.
 47. **Tang, J., N. Abovich, M. L. Fleming, B. Seraphin, and M. Rosbash.** 1997. Identification and characterization of yeast homolog of U1 snRNP-specific protein C. *EMBO J.* **16**:4082–4091.
 48. **Tzamaris, D., and K. Struhl.** 1995. Distinct TPR motifs of Cyc8 are involved in recruiting the Cyc8-Tup1 corepressor complex to differentially regulated promoters. *Genes Dev.* **9**:821–831.
 49. **Umen, J. G., and C. Guthrie.** 1995. A novel role for a U5 snRNP protein in 3' splice site selection. *Genes Dev.* **9**:855–868.
 50. **White, J. V., S. M. Stultz, and T. F. Smith.** 1994. Protein classification by stochastic modeling and optimal filtering of amino acid sequences. *Math. Biosci.* **119**:35–75.
 - 50a. **Williams, C., and K. O'Hare.** Personal communication.
 51. **Wolfe, K. H., and D. C. Shields.** 1997. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**:708–713.
 52. **Zhang, K., D. Smouse, and N. Perrimon.** 1991. The *crooked neck* gene of *Drosophila* contains a motif found in a family of yeast cell cycle genes. *Genes Dev.* **5**:1080–1091.