The EMBO Journal Vol.16 No.13 pp.4082–4091, 1997

Identification and characterization of a yeast homolog of U1 snRNP-specific protein C

Jie Tang, Nadja Abovich, Margaret L.Fleming, Bertrand Séraphin[1] and Michael Rosbash[2]

Howard Hughes Medical Institute and Department of Biology, Brandeis University, Waltham, MA 02254, USA and [1]Gene Expression Program, EMBL, Meyerhofstrasse 1, D-69117 Heidelberg, Germany

[2]Corresponding author
e-mail: rosbash@binah.cc.brandeis.edu

U1C is one of the three human U1 small nuclear ribonucleoprotein (snRNP)-specific proteins and is important for efficient complex formation between U1 snRNP and the pre-mRNA 5′ splice site. We identified a hypothetical open reading frame in *Saccharomyces cerevisiae* as the yeast homolog of the human U1C protein. The gene is essential, and its product, YU1C, is associated with U1 snRNP. YU1C depletion gives rise to normal levels of U1 snRNP and does not have any detectable effect on U1 snRNP assembly. YU1C depletion and YU1C ts mutants affect pre-mRNA splicing *in vivo*, and extracts from these strains form low levels of commitment complexes and spliceosomes *in vitro*. These experiments indicate a role for YU1C in snRNP function. Structure probing with RNases shows that only the U1 snRNA 5′ arm is hypersensitive to RNase I digestion when YU1C is depleted. Similar results were obtained with YU1C ts mutants, indicating that U1C contributes to a proper 5′ arm structure prior to its base pairing interaction with the pre-mRNA 5′ splice site.

Keywords: pre-mRNA/*Saccharomyces cerevisiae*/ 5′ splice site/U1 snRNP-specific protein C

Introduction

RNA splicing is a process that removes intervening sequences (introns) precisely and joins two flanking sequences (exons) together. Some highly structured large introns (group I and group II introns) can undergo self-splicing, i.e. the splicing reaction is catalyzed by the RNA itself (for reviews, see Saldanha *et al.*, 1993; Michel and Ferst, 1995). In the case of nuclear pre-mRNA, splicing takes place in a large ribonucleoprotein complex called the spliceosome. Four small nuclear ribonucleoprotein (snRNP) particles, U1, U2, U5 and U4/U6, assemble into spliceosomes through an ordered, dynamic pathway. U1 snRNP recognizes the pre-mRNA 5′ splice site, U2 snRNP binds to the branch point sequence and U4/U5/U6 tri-snRNP enters the pathway subsequently to form an active spliceosome in which the splicing reactions take place (Rymond and Rosbash, 1988; Guthrie, 1991; Lamond, 1993; Madhani and Guthrie, 1994).

As the major components of the spliceosome, each snRNP contains one or two small nuclear RNAs (snRNAs) and >10 proteins. A group of low molecular weight proteins (Sm proteins) is present in all snRNPs, and there are various numbers of specific proteins in each snRNP (Lührmann *et al.*, 1990; Mattaj, 1993; Séraphin, 1995).

The U1 snRNP in higher eukaryotic cells is made up of one U1 snRNA, eight Sm proteins and three U1-specific proteins. Several regions in U1 snRNA are functionally important and are highly conserved among different species. For instance, the 5′ arm of U1 snRNA base-pairs with the pre-mRNA 5′ splice site; stem–loops A and B are the binding sites of the U1snRNP-specific proteins U1-70K and U1A, respectively; and the Sm site, a single-stranded region near the 3′ end of the molecule, serves as the binding site for Sm proteins (Hamm *et al.*, 1990; Lührmann *et al.*, 1990). The U1-70K and U1A proteins bind to their target sequences via their N-terminal RNA-binding domains (RBDs). The U1-70K protein contains an arginine-rich region at its C-terminus, which probably interacts with SR proteins to facilitate the interaction between U1 snRNP and the pre-mRNA (Wu and Maniatis, 1993; Kohtz *et al.*, 1994). There is an additional RBD at the C-terminus of U1A. This domain is not required for the binding of U1A to the U1 snRNA but may play a role in polyadenylation (Lutz and Alwine, 1994). Unlike U1-70K and U1A, the third U1-specific protein, U1C, does not contain an RBD, nor does it show high affinity binding to U1 snRNA. U1C carries a zinc finger-like structure at its N-terminus, which is required for the association of U1C with U1 snRNP (Nelissen *et al.*, 1991). This association is, at least in part, mediated by protein–protein interactions (Nelissen *et al.*, 1994). There is also evidence that U1C is required for a stable interaction between U1 snRNP and the pre-mRNA 5′ splice site (Heinrichs *et al.*, 1990). Although the exact biochemical role of the U1C protein is unclear, it is the only U1 snRNP protein with a well-described functional role during *in vitro* splicing.

The components of yeast (*Saccharomyces cerevisiae*) U1 snRNP have been partially characterized. Similarly to its higher eukaryotic counterpart, the yeast snRNP contains a single U1 snRNA molecule, a group of U1-specific proteins and a group of common snRNP proteins (Fabrizio *et al.*, 1994). These common Sm proteins resemble their mammalian counterparts quite well (Roy *et al.*, 1995; Séraphin, 1995). Yeast U1 snRNA is 5-fold larger than metazoan U1 snRNA, despite the conservation of most functionally important regions, e.g. the 5′ arm, stem–loop II and Sm site (Kretzner *et al.*, 1987, 1990; Siliciano *et al.*, 1987; Liao *et al.*, 1990). The yeast snRNP also contains more than seven specific proteins, in contrast to three specific mammalian U1 snRNP proteins (Fabrizio *et al.*, 1994). Two of these additional yeast U1 snRNP-specific proteins, PRP39p and PRP40p, have been cloned.

Fig. 1. The alignment of the human, *Xenopus*, *C.elegans*, *S.pombe* and *S.cerevisiae* U1C proteins. This multiple sequence alignment was generated by the Clustal W program (Thompson *et al.*, 1994), and edited by hand. The conserved residues among the five sequences are shaded in gray. Arrows point to the residues involved in the zinc finger-like structure.

Although they are essential for yeast U1 snRNP function, no homologs are known in higher eukaryotic cells (Lockhart and Rymond, 1994; Kao and Siliciano, 1996). Additional yeast U1 snRNP-specific proteins have been identified, but these results are not yet published (J.Tang, unpublished data).

The yeast counterparts of two of the three mammalian U1 snRNP-specific proteins, U1-70K and U1A, have been identified and characterized. They resemble their metazoan counterparts in basic structure and U1 snRNA binding (Smith and Barrell, 1991; Kao and Siliciano, 1992; Liao *et al.*, 1993; Tang and Rosbash, 1996). Yeast U1-70K contains one RBD, and the protein binds to the conserved stem–loop II in U1 snRNA (Smith and Barrell, 1991; Kao and Siliciano, 1992). Yeast U1A binds to a large loop on U1 snRNA stem III through its N-terminal RBD. Its C-terminal RBD also contributes to pre-mRNA splicing, but its precise function and binding site are unknown (Tang and Rosbash, 1996). The yeast counterpart of the third U1 snRNP-specific protein, U1C, has not yet been identified.

Through database searching, we found an open reading frame (ORF) (as determined by the yeast genome sequencing project) that shares significant homology with the human U1C protein. By demonstrating that this protein is essential for splicing and is associated with U1 snRNP, we claim that it is the yeast version of the U1C protein. A combination of *in vivo* and *in vitro* assays indicate that the yeast protein contributes to U1 snRNP function like the contribution of U1C to mammalian U1 snRNP function. Moreover, the U1C-depleted snRNP and U1C ts mutants show a dramatic structural alteration at the 5′ end of snRNA. This is almost certainly related to the failure of the altered snRNPs to undergo a proper base-pairing interaction with the pre-mRNA 5′ splice site.

Results

Identification of a putative yeast U1C protein encoded by an essential gene

Using the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990), we searched sequence databases with the human U1C protein sequence. The closest relative of human U1C in the yeast *S.cerevisiae* is a 231 amino acid (27 kDa) hypothetical protein named L8003.21 that we will refer to as YU1C in this manuscript. The sequence was generated by the yeast genome sequencing project. When YU1C was used as a query for a BLAST search, the human and *Xenopus* U1C proteins gave the highest scores. With a more sensitive search method (see Materials and methods), we also identified the U1C homologs in *Caenorhabditis elegans* and *Schizosaccharomyces pombe*, both of which were determined by systematic genome sequencing projects. A multiple sequence alignment of the human, *Xenopus*, *C.elegans*, *S.pombe* and *S.cerevisiae* proteins showed that the high homology region is limited to the first 38 residues (Figure 1). The human U1C protein has a zinc finger-like motif in this region, which is essential for its binding to U1 snRNP (Nelissen *et al.*, 1991). The newly identified proteins (*C.elegans*, *S.pombe* and *S.cerevisiae*) are ~50% identical to human U1C in this region, with the critical cysteine and histidine residues of the zinc finger-like sequence conserved. The C-terminal parts of these proteins are quite divergent, ranging from being proline rich in human and *Xenopus*, to lysine, serine and arginine rich in yeast.

The gene encoding the YU1C protein (named *YHC1* for yeast homolog of U1C) was cloned by PCR amplification of genomic yeast DNA with a pair of *YHC1*-specific primers. One copy of the endogenous *YHC1* gene subsequently was deleted and replaced by the *LEU2* gene by homologous recombination in a diploid strain. A successful disruption was determined by Southern blotting (data not shown). The heterozygous diploid was sporulated, and tetrads were dissected. All tetrads gave rise to at most two viable spores. All of the viable spores were LEU^-, indicating that the disruption of the *YHC1* gene was lethal (data not shown; see Materials and methods for details). Four spore tetrads were obtained from sporulation of the heterozygous diploids transformed with a *URA3* plasmid carrying either the wild-type *YHC1* or a hemagglutinin (HA)-tagged version under the control of the GAL pro-

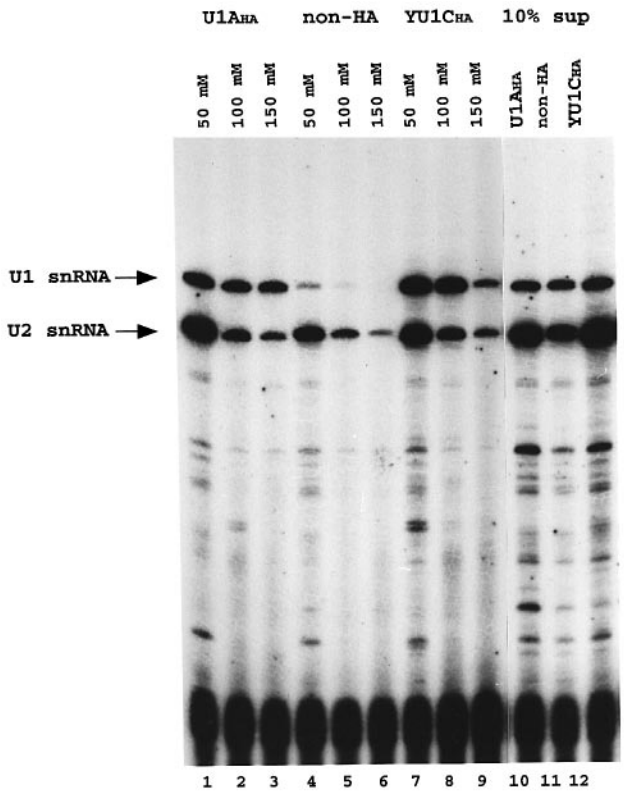

Fig. 2. The HA-tagged YU1C protein co-immunoprecipitates with U1 snRNA. Splicing extracts from the indicated strains were incubated with Gamma-Bind beads pre-coated with 12CA5 antibody. The beads were washed with increasing salt. RNA co-immunoprecipitated on beads was recovered and reverse transcribed with U1 and U2 snRNA-specific primers. U1 and U2 snRNA are indicated by arrows. The KCl concentrations in the washing buffer are indicated above. Lanes 1–3, extract with tagged U1A (U1A$_{HA}$); lanes 4–6, extract without tagged protein (non-HA); lanes 7–9, extract with tagged YU1C (YU1C$_{HA}$); lanes 10–12, 10% of the supernatant from the immunoprecipitation reactions indicated above.

moter. In this case, all the LEU$^+$ colonies were also URA$^+$ and did not grow on 5-fluoro-orotic acid (5-FOA)-containing plates (data not shown).

YU1C is associated with U1 snRNP and required for in vivo splicing

To test whether the YU1C protein is associated with U1 snRNP, we tagged the YU1C protein with the HA epitope and immunoprecipitated yeast splicing extracts with the 12CA5 antibody (a monoclonal antibody recognizing the HA epitope). U1 snRNA is specifically co-immuno precipitated from YU1C$_{HA}$ extract (Figure 2; compare YU1C$_{HA}$ with the non-HA control). YU1C is more salt-sensitive than yeast U1A protein (Figure 2; compare YU1C$_{HA}$ with U1A$_{HA}$), suggesting that YU1C may be a more loosely associated U1 snRNP protein.

In order to conditionally deplete the yeast U1C protein *in vivo*, we put the *YHC1* coding sequences under a GAL-controlled promoter (GAL-YHC1). A strain that carried the GAL-YHC1 plasmid as the only source of YU1C protein grew well in galactose-containing medium and somewhat more slowly in glucose-containing medium. The ΔYHC1/GAL-YHC1 strain grown in glucose was harvested at several time points and total yeast RNA was extracted. Pre-U3 and U3 snRNA were detected by primer extension. The pre-U3 level increased dramatically during growth in glucose (Figure 3A), indicating that *in vivo* splicing was significantly impaired by depletion of the yeast U1C protein.

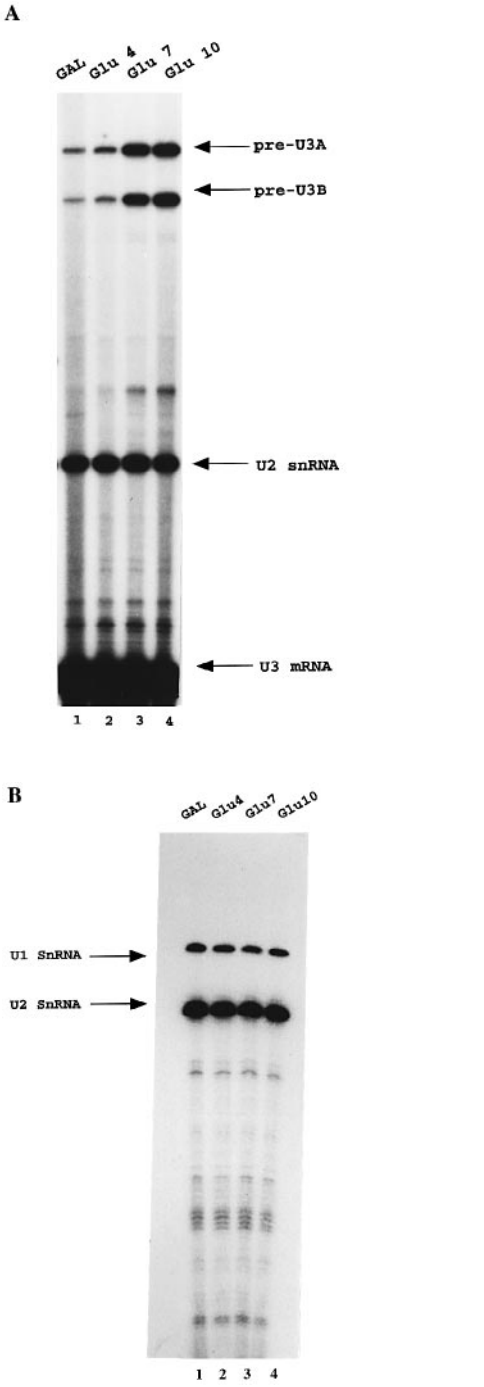

Fig. 3. *In vivo* splicing in the yeast U1C-depleted strain. (**A**) The ΔYHC1/GAL-YHC1 strain was grown in galactose-containing medium (lane 1) for 10 h or in glucose-containing medium for 4 (lane 2), 7 (lane 3) or 10 h (lane 4). Total RNA was extracted from each culture and the U3 snRNA level was assayed by reverse transcription with the U3-specific primer DT1967. Pre-U3A and pre-U3B as well as U3 snRNA are indicated by arrows. The reverse transcription of U2 snRNA served as an internal loading control. (**B**) Same as (A), except the U1 snRNA level was assayed by primer extension with U1 snRNA-specific primer DT586. The U1 and U2 snRNA bands are indicated by arrows.

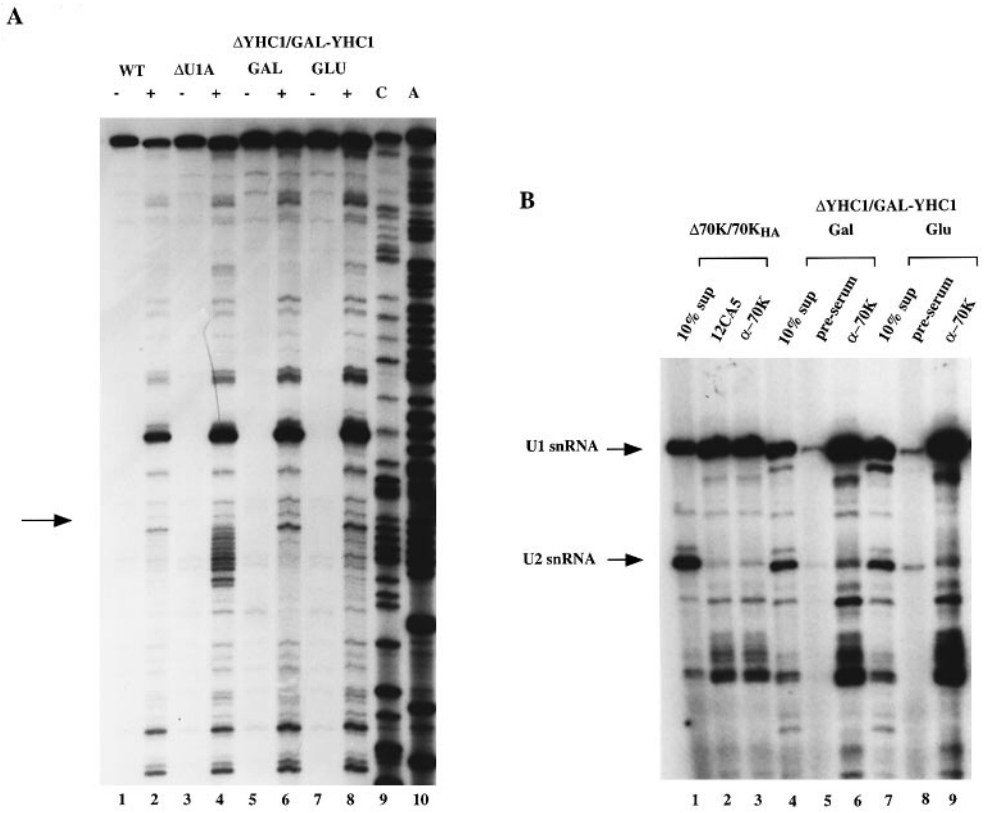

Fig. 4. YU1C-depleted snRNP retains other specific proteins. (**A**) U1A protein protects its binding site in the YU1C-depleted snRNP. Twenty-five ml cultures of the wild-type strain (lanes 1 and 2), the ΔU1A strain (lanes 3 and 4) and the ΔYHC1/GAL-YHC1 strain [grown in galactose- (lanes 5 and 6] or glucose- (lanes 7 and 8) containing medium] were treated with 200 μl of 1:2 diluted DMS. Total yeast RNA was extracted from DMS-treated strains and DMS modification patterns were assayed by reverse transcription with U1-specific primer DT2502. Lanes 1, 3, 5 and 7, negative controls without DMS treatment; lanes 2, 4, 6 and 8, DMS-treated samples. Lanes 9 and 10, U1 snRNA sequence ladders. U1 snRNA loop IIIc, the U1A protein-binding site, is indicated by the arrow. (**B**) U1-70K protein associates with YU1C-depleted U1 snRNP. Splicing extracts from the ΔSNP1/U1–70K$_{HA}$ strain (lanes 1–3) or from the ΔYHC1/GAL-YHC1 strain grown in galactose- (lanes 4–6) or glucose- (lanes 7–9) containing medium are immunoprecipitated by anti-70K antibody. The U1 snRNA co-immunoprecipitated with U1-70K protein was recovered from beads and assayed by reverse transcription with ^{32}P-labeled U1- and U2-specific primers. The 12CA5 antibody against the HA epitope was used as a positive control (lane 2) and the pre-immune serum was used as a negative control (lanes 5 and 8); 10% of supernatant is shown to indicate immunoprecipitation efficiency (lanes 1, 4 and 7). The U1 and U2 snRNA bands are indicated by arrows.

Lack of an effect of YU1C on U1 snRNP assembly

U1 snRNA levels were monitored during glucose depletion (Figure 3B). Remarkably, there was no detectable decrease. Because the cells grow normally during the first few doublings in glucose, the lack of an effect on U1 snRNA levels indicates that U1 snRNP accumulation and stability are largely unaffected by the absence of the YU1C protein.

Several additional approaches were applied to assess the effect of YU1C on snRNP assembly. An *in vivo* dimethylsulfate (DMS) modification assay indicates that the recently identified U1A-binding site is protected in the YU1C-depleted strain (Tang and Rosbash, 1996; Figure 4A, lane 8). Lane 4 shows the pattern in this region when the U1A protein is absent. *In vitro*, an anti-70K antibody co-immunoprecipitates U1 snRNA from a YU1C-depleted extract as efficiently as from a wild-type extract (Figure 4B, compare lanes 6 and 9). Both assays indicate that other U1 snRNP-specific proteins assemble into U1 snRNP in the absence of the YU1C protein, consistent with the lack of a detectable effect of YU1C depletion on U1 snRNP accumulation and stability.

The yeast U1C protein is important for the U1 snRNP–pre-mRNA interaction in vitro

To verify that YU1C depletion affects U1 snRNP function, we made splicing extracts from the ΔYHC1/GAL-YHC1 strain grown in galactose- or glucose-containing medium and compared their ability to form commitment complexes and spliceosomes with pre-mRNA. The extracts from the glucose-grown strain formed much less commitment complex and spliceosomes than the extracts from the galactose-grown strain (Figure 5A), indicating that the YU1C-depleted U1 snRNP is unable to form a stable pre-mRNA complex. Consistent with the commitment complex phenotype, the glucose-grown extract also failed to support substantial levels of *in vitro* splicing (Figure 5B).

In another *in vitro* assay, a biotinylated pre-mRNA substrate was pre-bound to streptavidin beads and then incubated with YU1C-depleted or wild-type extracts under splicing conditions. The U1 snRNA associated with the substrate was assayed by reverse transcription with a U1-specific primer. In the depleted extract, much less U1 snRNA was associated with the pre-mRNA (Figure 6A). To verify that the base-pairing potential of the 5′ arm was affected by YU1C depletion, we substituted the pre-mRNA substrate in the previous experiment with a biotinylated 2′-*O*-methyl RNA oligonucleotide complementary to the U1 snRNA 5′ arm (Figure 6B). A similar result was obtained, i.e., the YU1C-depleted snRNP binds poorly to the oligonucleotide (Figure 6B). Taken together, these *in vitro* results suggest that the YU1C protein potentiates the base-

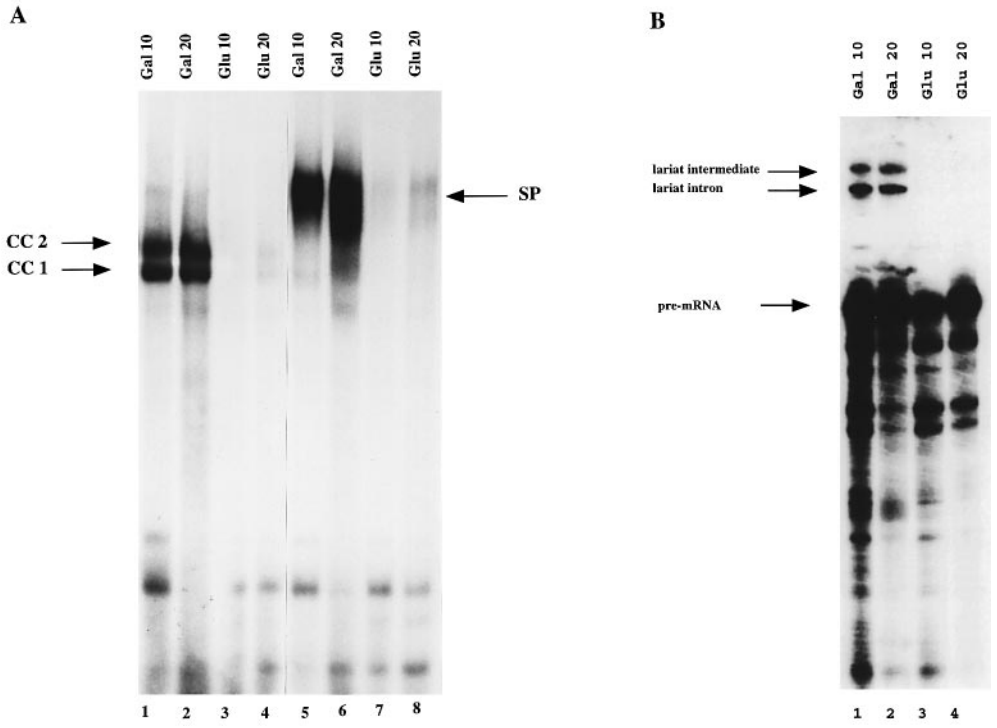

Fig. 5. Yeast YU1C-depleted extracts form less commitment complexes and spliceosomes. (**A**) Splicing extracts were made from the ΔYHC1/GAL-YHC1 strain grown in either galactose-containing medium for 10 (lanes 1 and 5) or 20 h (lanes 2 and 6) or glucose-containing medium for 10 (lanes 3 and 7) or 20 h (lanes 4 and 8). ^{32}P-Labeled pre-mRNA was incubated with different extracts under splicing conditions with (lanes 1–4) or without (lanes 5–8) oligonucleotide-mediated RNase H digestion of U2 snRNA, and the complexes were analyzed by native gel electrophoresis. The positions of commitment complexes (CC) and spliceosome (SP) are indicated. (**B**) The same reactions from lanes 5–8 of (A) were phenol–chloroform extracted and ethanol precipitated. The ^{32}P-labeled pre-mRNA was assayed on a 15% denaturing polyacrylamide gel. The positions of the lariat intermediate, the lariat intron product and the pre-mRNA are indicated by arrows.

pairing interaction between the 5′ arm of U1 snRNA and a 5′ splice site sequence.

Effects of YU1C on U1 snRNP structure

Although glucose depletion of YU1C protein does not grossly affect the accumulation or stability of U1 snRNP (Figure 3B), we searched for a more subtle effect on U1 snRNP structure that might accompany the functional deficit of the YU1C-depleted extracts. U1 snRNP was digested with several RNases, to see if the missing YU1C causes any detectable change in U1 snRNA sensitivity. RNase V1, RNase A and nuclease S1 partial digestions did not detect any difference between the YU1C-depleted and wild-type extracts (data not shown). This is consistent with the presence of the U1A and the U1-70K proteins. However, RNase I probing showed cleavage sites, specific for the YU1C-depleted extract (Figure 7A). These two RNase I-sensitive sites reside in the U1 snRNA 5′ arm (Figure 7B), at the CC sequence that base-pairs with the highly conserved GG at the 5′ splice site.

To verify and extend this observation, we isolated several YU1C temperature-sensitive mutants (see Materials and methods). Three YU1C ts mutants that grow well at 25 but not at 37°C were characterized in some detail. First, the splicing efficiency of the U3 snRNA was assayed in these ts strains. Even at 25°C, these strains accumulated much more pre-U3B RNA than the wild-type control strain (Figure 8 and data not shown), indicating that *in vivo* splicing of these ts strains is already compromised at the permissive temperature. After switching the ts strains to 37°C for 12 h,

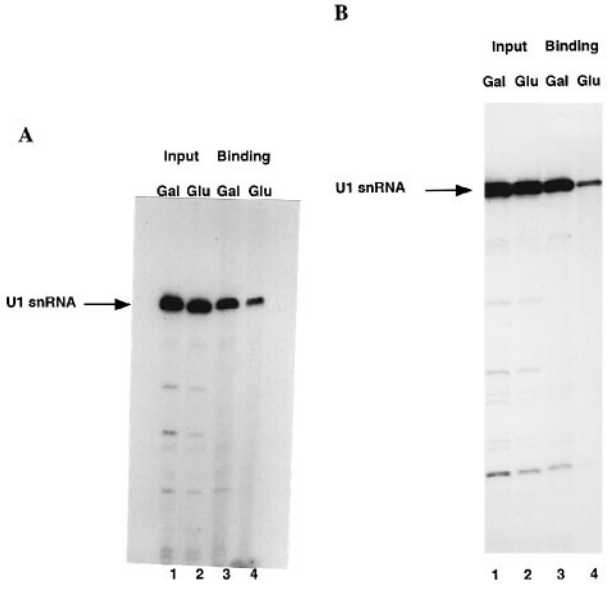

Fig. 6. YU1C-depleted U1 snRNP associates with pre-mRNA less efficiently. (**A**) Extracts from strain ΔYHC1/GAL-YHC1 grown in galactose- (lanes 1 and 3) or glucose- (lanes 2 and 4) containing medium were incubated with biotinylated pre-mRNA (pre-bound to streptavidin beads) under splicing conditions. U1 snRNA associated with pre-mRNA was assayed by reverse transcription with U1-specific primer DT586. Lanes 1 and 2, 10% of input; lanes 3 and 4, U1 snRNA bound to pre-mRNA. The U1 snRNA is indicated by an arrow. (**B**) Same as (A), except a biotinylated 2′-*O*-methyl RNA oligonucleotide complementary to the U1 snRNA 5′ arm region was used instead of pre-mRNA.

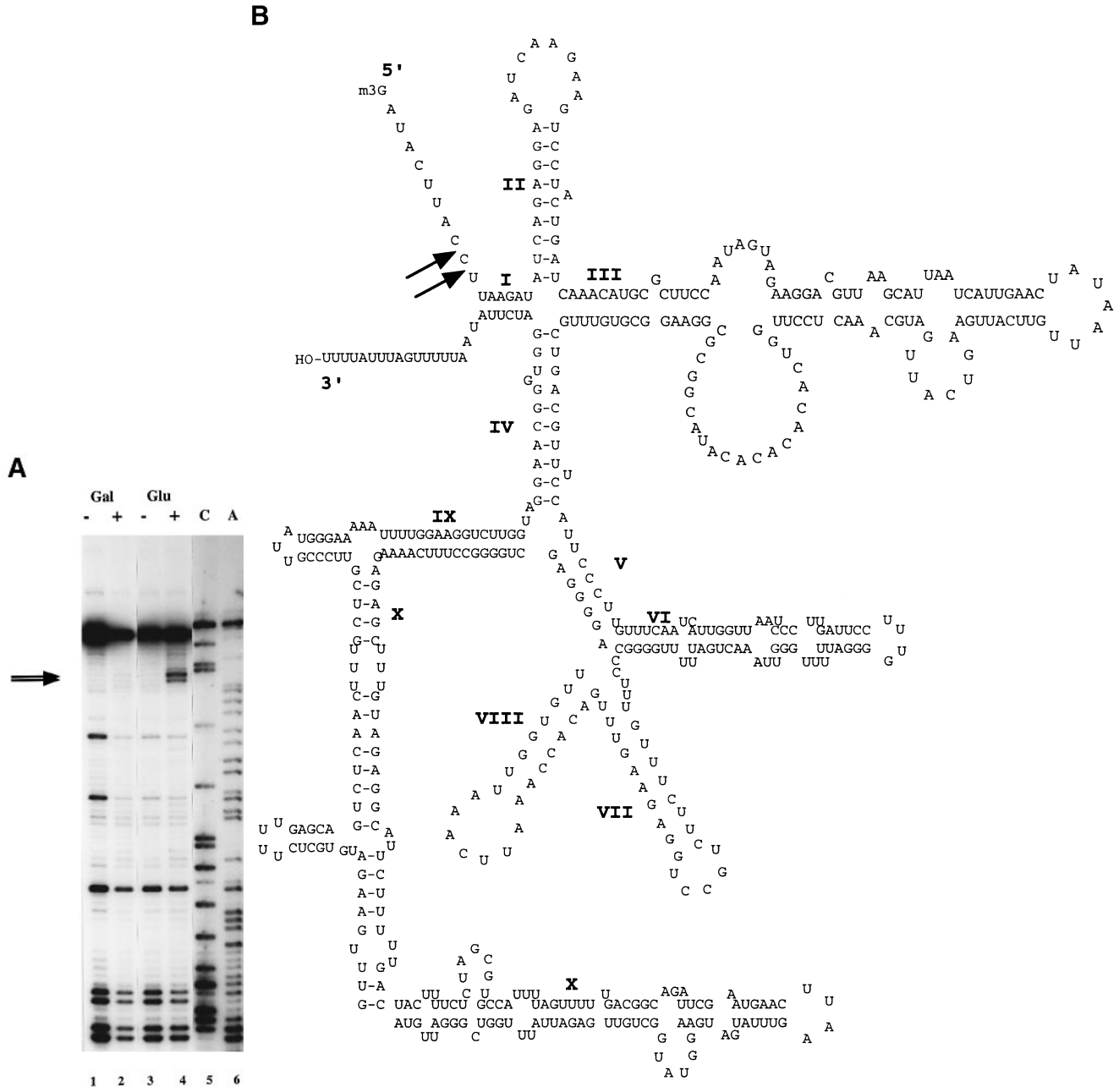

Fig. 7. U1 snRNP structure probing with RNase I. (**A**) Extracts from strain ΔYHC1/GAL-YHC1 grown in galactose- (lanes 1 and 2) or glucose- (lanes 3 and 4) containing medium were partially digested with RNase I. The U1 snRNA digestion pattern was assayed by reverse transcription of total yeast RNA with a U1-specific primer. (–) control without RNase I; (+) RNase I partial digestion. Lanes 5 and 6, U1 snRNA sequence ladders (an *in vitro* transcribed U1 snRNA, which contains one extra nucleotide at its 5′ end, is used as a template). RNase I-hypersensitive sites are indicated by arrows. (**B**) U1 snRNA secondary structure. The RNase I-hypersensitive sites in the YU1C-depleted extract are mapped on the 5′ arm of U1 (indicated by arrows).

the pre-U3B RNA levels were unaffected but U3 snRNA levels were dramatically decreased (Figure 8; to allow a comparison of U3 snRNA levels, the gel was dramatically underexposed compared with Figure 3A). In the wild-type control strain, U3 RNA levels were indistinguishable between 25 and 37°C (Figure 8, compare lanes 7 and 8). The result shows that the three YU1C ts mutants are already defective in splicing at 25°C, which becomes even more prominent after a shift to 37°C. We then made extracts from these three ts strains, incubated either at permissive or non-permissive temperature. In all cases, no complexes or splicing were observed (data not shown). Partial digestion of the 25°C extracts with RNase I indicated the hypersensitive cleavage sites within the U1 snRNA 5′ arm (Figure 9, lane 4; data not shown), consistent with the observation that the ts strains are compromised in splicing even at the permissive temperature. For the ts strains shifted to 37°C, the cleavages are stronger than those at 25°C (compare lane 6 with lane 4; and data not shown). The observations on the ts mutants confirm the glucose depletion data and link the presence or function of the YU1C protein to the structure of the U1 snRNA 5′ arm.

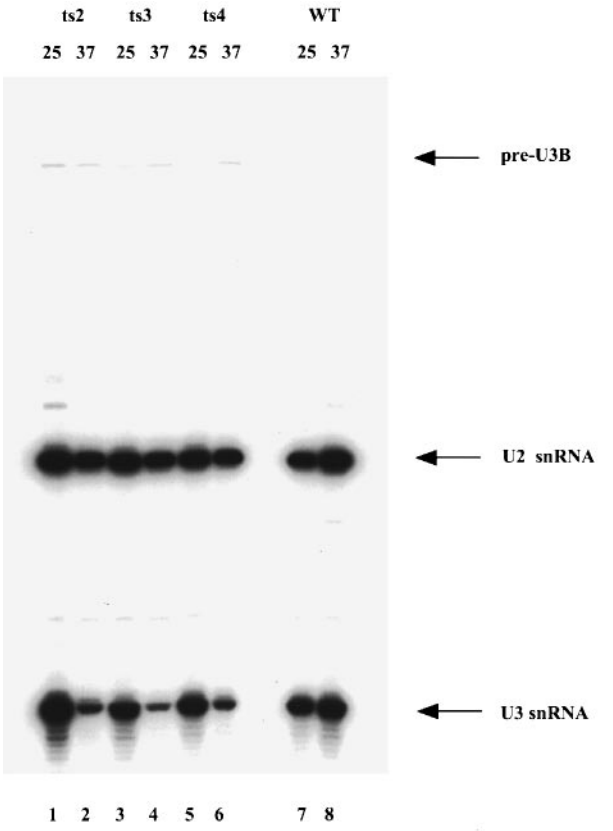

Fig. 8. *In vivo* splicing phenotype of YU1C ts strains. YU1C ts strains and a wild-type control strain were incubated at 25 or 37°C for 12 h. Total yeast RNA was extracted, and U3 snRNA levels were assayed by reverse transcription with a U3-specific primer. A U2 snRNA-specific primer was mixed with the U3 primer as an internal loading control. Lanes 1 and 2, YU1C ts2; lanes 3 and 4, YU1C ts3; lanes 5 and 6, YU1C ts4; lanes 7 and 8, wild-type control. Lanes 1, 3, 5 and 7, strains incubated at 25°C; lanes 2, 4, 6 and 8, strains incubated at 37°C. The pre-U3B, U2 snRNA and U3 snRNA bands are indicated by arrows. The pre-U3A bands are too faint at this level of exposure.

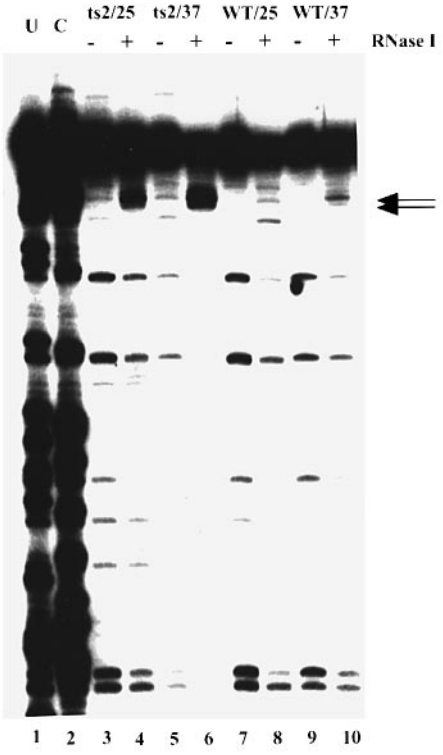

Fig. 9. RNase I structure probing of YU1C ts strains. YU1C ts or wild-type strains were incubated at 25 or 37°C for 12 h. Splicing extracts were made from those strains and subjected to RNase I partial digestion. Digestion patterns were assayed by reverse transcription with a U1 snRNA-specific primer. (–) control without RNase I; (+) RNase I partial digestion. Lanes 1 and 2, RNA sequence ladders; lanes 3 and 4, YU1C ts2 strain at 25°C; lanes 5 and 6, YU1C ts2 strain at 37°C; lanes 7 and 8, wild-type strain at 25°C; lanes 9 and 10, wild-type strain at 37°C. RNase I-hypersensitive sites are indicated by arrows. YU1C ts3 and ts4 strains show similar phenotypes (data not shown).

Discussion

The results of this study indicate that the completely conserved 10 nucleotides at the 5′ end of U1 snRNA have an RNP structure that is important for the base-pairing interaction with a 5′ splice site. There are few other cases in which a missing or mutant snRNP protein has been associated with an alteration in snRNP structure (Tang and Rosbash, 1996). Moreover, there are no other cases of which we are aware in which a change in snRNP structure is paralleled by a change in *in vitro* snRNP function.

Although the binding results (Figures 5 and 6) are consistent with those from previous *in vitro* studies that demonstrated a comparable lack of activity of U1C-depleted mammalian snRNPs (Heinrichs *et al.*, 1990), an effect on snRNP structure was not predictable. For example, the U1 snRNP–pre-mRNA base-pairing might be stabilized by one or more U1 snRNP proteins without any detectable impact of these proteins on snRNP structure prior to the snRNP–substrate interaction. Enhanced nuclease sensitivity implies that access to the 5′ arm is normally blocked by an interaction with one or more snRNP proteins. A further implication is that this interaction positively contributes to the base-pairing between the U1 snRNP 5′ arm and the 5′ splice site. Either YU1C interacts directly with the 5′ arm and subserves this role, or YU1C plays a more indirect role and affects the direct interaction between the 5′ arm and other U1 snRNP proteins. Biochemical approaches will be necessary to identify the proteins that interact directly with the 5′ arm.

The conservation between the yeast and human U1C protein is limited to the common N-terminal zinc finger-like region (C2H2 motif), ~38 amino acids in length. Thus it was not obvious that the identified ORF would encode the yeast U1C equivalent. However, it has been shown that the first 40 residues of human U1C are necessary and sufficient for U1C–U1 snRNP association (Nelissen *et al.*, 1991), and this region is also required for U1C homodimerization (Gunnewiek *et al.*, 1995). The high degree of phylogenetic conservation of the C2H2 motif is consistent with the idea that this motif is very important for U1C function. However, a truncated yeast U1C protein carrying the first 80 residues fails to rescue the lethality of the ΔYHC1 strain (data not shown), indicating that the C2H2 motif may not be sufficient for the function of the protein. The C2H2 motif also exists in other splicing factors such as Prp6p, Prp11p and Prp9p (Legrain and Choulika, 1990). Although the C2H2 motifs are not required for the protein–protein interactions within the Prp9p–Prp11p–Prp21p complex, the deletion of one or two motifs from

Prp9p or Prp11p is dominant lethal, suggesting a functional role for this kind of motif (Legrain and Chapon, 1993; Legrain *et al.*, 1993). Whether this zinc finger-like motif binds to zinc, to RNA (the 5′ arm?) or to other proteins remains unknown.

The C-terminal part of the yeast U1C protein is rich in lysine, arginine and serine. A similar feature can also be found in the C-terminus of the yeast U1-70K protein. The function of this type of sequence motif has not been determined. It may resemble the function of the arginine/serine-rich domain in human U1-70K and other metazoan splicing factors, which is involved in protein–protein interactions (Wu and Maniatis, 1993; Kohtz *et al.*, 1994). It could also facilitate protein–RNA interactions via its positively charged residues. Alternatively, it may only act indirectly, by aiding the maintenance of a proper YU1C protein conformation. The corresponding region of human U1C is proline rich and has not been assigned a functional role.

The co-immunoprecipitation of U1 snRNA with HA-tagged U1C protein ($YU1C_{HA}$) is salt sensitive, i.e. U1 snRNA co-immunoprecipitation decreases substantially when the salt concentration is >150 mM (Figure 2). The HA-tagged yeast U1-70K and U1A proteins are co-immunoprecipitated with U1 snRNA even at 500 mM salt (J.Tang, unpublished data). This difference in the salt sensitivity suggests that the yeast U1C is a more loosely associated U1 snRNP protein. It is possible that the salt sensitivity of $YU1C_{HA}$ is an artifact of the extra tagging sequence. However, the $YU1C_{HA}$ protein is functionally indistinguishable from the wild-type YU1C, e.g. it rescues the lethal phenotype of the ΔYHC1 strain, and the $YU1C_{HA}$-containing extract forms normal commitment complexes and spliceosomes (Figure 5A).

If YU1C is indeed a loosely associated snRNP protein, this would fit well with the abundant evidence that much of U1 snRNP assembly occurs normally without YU1C. There is no U1 snRNA level change in the YU1C-depleted strain (Figure 3B), suggesting that U1 snRNP assembly is normal in the absence of the YU1C protein. Assayed by *in vivo* DMS modification, U1 snRNA loop IIIc (the yeast U1A protein-binding site) is protected in the YU1C-depleted strain (Figure 4A), suggesting that U1A binding is not affected by the lack of YU1C. U1 snRNA can be co-immunoprecipitated from the YU1C-depleted extracts by an anti-70K antibody, suggesting that U1-70K also associates normally with the snRNP in the absence of YU1C (Figure 4B). Taken together, the observations suggest that YU1C is a late addition to a nearly fully formed snRNP; in its absence, there are no profound changes in U1 snRNP quantity, integrity or gross structure *in vivo* or *in vitro*.

Yet the Δ*YHC* strain is not viable, indicating that the yeast U1C protein is essential for growth. Thus, the viability of the ΔYHC/GAL-YHC1 strain in glucose-containing medium is almost certainly due to residual synthesis from the GAL promoter. Previously we have put the *MUD2*, *MSL1* and *SNP1* genes under GAL control, and in all these cases some expression is detectable (Abovich *et al.*, 1994; Tang *et al.*, 1996; J.Tang, unpublished data). Despite the weak growth defect, *in vivo* splicing is affected by depletion of the YU1C protein (Figure 3A). More strikingly, the YU1C-depleted extract forms much less commitment complex (Figure 5A), suggesting that *in vitro* U1 snRNP function is severely compromised by the absence of the YU1C protein. Similar *in vitro* results were obtained with the three ts mutant extracts (data not shown). These YU1C mutations have not been reduced to single amino acid changes, so it is not known what regions of the protein are responsible for the splicing phenotypes. Nor can we exclude that the effects on U1 snRNP function are an indirect consequence of a more primary effect on snRNP assembly. However, the defects in U1 snRNP function are paralleled by a structural changes in the U1 snRNP 5′ arm. As this is the only snRNA region known to interact directly with the pre-mRNA substrate and the only known region affected by YU1C depletion, it may be directly affected by YU1C and directly related to the defective commitment complex phenotype of the depleted and ts snRNPs.

The U1C protein could lead to a stabilization of the U1 snRNA–pre-mRNA complex, by interacting with the U1 snRNA 5′ arm–5′ splice site duplex. Alternatively, U1C could facilitate the rate of snRNP–pre-mRNA interaction, a possibility that is favored by the apparent pre-structuring of the U1 snRNA 5′ arm. This kinetic interpretation is also favored by the results of the 2′-*O*-methyl RNA oligonucleotide binding. This substrate was designed to base-pair with all 10 residues in the U1 snRNA 5′ arm. This interaction should be much stronger or more stable than base-pairing between U1 snRNA and the 5′ splice site. However, the YU1C-depleted snRNP also fails to associate efficiently with this oligonucleotide (Figure 7B). It appears as if the 5′ arm of U1 snRNA is not available for base-pairing in the depleted or mutant snRNP. We speculate that YU1C maintains the U1 snRNA 5′ arm in a proper conformation, poised for a productive base-pairing interaction with a complementary RNA. In this view, YU1C may function as a dedicated local RNA chaperone.

Materials and methods

Strains and plasmids

Plasmid GAL-$YU1C_{HA}$ was constructed by inserting the HA epitope after the last amino acid of the YU1C protein. The tagged coding sequence was inserted in place of the *MUD2* coding sequence in plasmid GAL-MUD2HA (Abovich *et al.*, 1994). Plasmid GAL-YHC1 was modified from GAL-$YU1C_{HA}$ by replacing the 2 μm sequences with the *CEN3* sequence. Plasmids GAL-$YU1C_{HA}$ or GAL-YHC1 were transformed into a diploid strain heterozygous for a *YHC1* deletion. Tetrads were dissected to give rise to ΔYHC1/GAL-$YU1C_{HA}$ or ΔYHC1/GAL-YHC1 strains.

In the ΔSNP1/U1-$70K_{HA}$ strain, the endogenous *SNP1* gene was replaced by the *LEU2* gene and the HA-tagged U1-70K protein was expressed. In the ΔU1A strain, the endogenous *MUD1* gene was replaced by an *ADE2* gene.

Oligonucleotides

Oligonucleotides DT586, DT2502, DT58 and DT1967 are complementary to U1 (nucleotides 113–135 for DT586, nucleotides 266–287 for DT2502), U2 (nucleotides 110–120) and U3 (5′ CCAAGTTGGATTCAGTGGCTC 3′), respectively. The 2′-*O*-methyl RNA oligonucleotide (5′ GCCAGGUAAGUAU 3′) is complementary to U1 snRNA nucleotides 1–10.

Database searching

The full-length human U1C protein sequence was used as a query to search the database with BLAST (Altschul *et al.*, 1990). When the yeast U1C protein was identified, a profile (Gribskov *et al.*, 1987) was built

from the conserved N-terminus of the human, *Xenopus* and yeast U1C protein sequences. This profile was used to search the protein database using the PROFILESEARCH application of the GCG package (Devereux *et al.*, 1984) or nucleic acid databases using the TPROFILESEARCH program, allowing the identification of fragments of the *S.pombe* and *C.elegans* proteins in genomic sequences. Intron–exon structures for the corresponding genes were assigned by using consensus sequences for intron boundaries in the corresponding organism, similarity with the yeast and human U1C proteins and partial cDNA sequences available in the case of the *C.elegans* U1C protein.

YHC1 disruption

A plasmid (pU1CKO) in which the *LEU2* gene is flanked by the 1085 nucleotides immediately upstream of the *YHC1* initiating ATG and 837 nucleotides of downstream sequence, effectively deleting all but 50 nucleotides of coding sequence, was constructed. Unique *Bam*HI and *Sph*I sites were created by PCR to generate a linear fragment for yeast transformation and one-step gene disruption.

A wild-type diploid strain was obtained by crossing MGD353-13D (Séraphin *et al.*, 1988) to MGD 353-46D (α, *trp1-289*, *ura3-52*, *leu2-3*, *112*, *his3-Δ1* cyh r.). This diploid strain was transformed to leucine prototrophy with the *Bam*HI–*Sph*I fragment described above. Successful deletion of the *YHC1* gene was confirmed by Southern blotting.

YU1C temperature-sensitive mutants

A 1.7 kb fragment including 727 bp of the YU1C ORF was amplified by PCR using the primers 5′-CATCAACCGGAATTCGTGTATGTA-GATTTGGAAG-3′ and 5′-TGGTCCGCGGTAGCATAGGGGATGAG-ATCT-3′. Mutagenic conditions were modified from Cadwell and Joyce (1992) with the addition of 5% formamide to the reaction buffer. The PCR product was cloned into centromeric vector pRS414 (Stratagene). The purified plasmids were used to transform the DYHC1/GAL-YHC1 strain, which has an interrupted *YHC1* locus and contains a *URA3* plasmid carrying the wild-type *YHC1* gene. Transformants were grown at 25°C for 72 h and subsequently replica-plated on 5-FOA to select for cells cured of *URA3* plasmid. Cells that grew on 5-FOA plates were then replica-plated at 25 and 37°C. Eight (of 4000) colonies were unable to grow at 37°C. Plasmid DNA was recovered from heat-sensitive yeast cells, and ts phenotypes were verified by retransformation. The mutant YU1C ORF was sequenced. Plasmid ts2, ts3 and ts4 had six, four and seven amino acid changes, respectively.

Immunoprecipitation

Immunoprecipitations of the tagged proteins were performed as described previously (Banroques and Abelson, 1989) except that extracts were incubated with 12CA5 monoclonal antibody pre-bound to Gamma-Bind beads in NET-50 buffer [50 mM Tris–HCl (pH 7.5), 50 mM NaCl, 0.05% NP-40]. After incubation, the beads were washed with NET-50, NET-100 or NET-150 (the numbers indicating the NaCl concentrations in mM).

The rabbit polyclonal antibody against yeast U1-70K protein was generously provided by P.Siliciano's laboratory. Immunoprecipitations with anti-70K antibody were performed as described above, except 25 ml of affinity-purified antibody or pre-immune serum was used. The antibody on beads was incubated with splicing extract in NET-100 buffer and then washed with NET-150.

In vivo DMS modification assay

In vivo DMS modifications were done as described previously (Ares and Igel, 1990) except that 200 µl of DMS (diluted 1:2 with ethanol) was added to 25 ml of yeast cultures (OD_{600} ⊤1) and incubated for 4 min at 30°C. RNA was extracted from DMS-treated cells by hot phenol extraction (Ares and Igel, 1990). Then 10 µg of total yeast RNA from each sample was analyzed by reverse transcription with ^{32}P-labeled U1-specific primer DT2502. The products were separated by electrophoresis on 5% denaturing polyacrylamide gels.

In vivo splicing

The ΔYHC1/GAL-YHC1 strain was grown for 10 h in 25 ml of rich medium containing 3% galactose, 1% sucrose or in rich medium containing 4% glucose for 4, 7 and 10 h. The YU1C ts strains were grown for 12 h in 25 ml of rich medium at 25 or 37°C. Total RNA was extracted from these strains and 10 µg of RNA was used for reverse transcription reactions. ^{32}P-labeled DT1967, DT586 and DT58 primers were used to detect U3, U1 and U2 snRNA, respectively. The reverse transcription products were electrophoresed on a 6% polyacrylamide gel.

In vitro complex formation

Splicing extracts were made from the ΔYHC1/GAL-YHC1 strain grown in either galactose- or glucose-containing medium for 10 or 20 h. The preparation of splicing extracts and the analysis of splicing complexes on native gels were performed as described previously (Abovich *et al.*, 1990). The absolute level of spliceosome and commitment complex formation is variable between different extract preparations, in part due to the cell density and different extract preparation methods.

Biotinylated RNA binding assay

The biotinylated pre-mRNA was transcribed *in vitro* with a 50:1 UTP/Bio-11-UTP ratio to obtain an average of only one biotinylated residue per molecule. Biotinylated 2′-*O*-methyl RNA oligonucleotides were chemically synthesized with the biotin incorporated at their 3′ ends. The biotinylated RNA was pre-incubated with streptavidin magnetic beads (Dynal) in dialysis buffer [20 mM HEPES (pH 7.9), 50 mM KCl, 0.2 mM EDTA, 20% glycerol] for 30 min at 4°C. The beads were washed twice with dialysis buffer before incubating with splicing extract mixed with an equal amount of splicing salts [150 mM potassium phosphate (pH 7.0), 6.25 mM $MgCl_2$, 7.5% PEG] for 30 min at 4°C. The beads were washed three times with NET-100 buffer. The snRNAs associated with the biotinylated RNA were recovered from the beads by incubation in PK buffer [0.1 M Tris–HCl (pH 7.5), 12 mM EDTA, 0.15 M NaCl, 1% SDS] at 65°C for 10 min, followed by phenol–chloroform extraction. U1 snRNA was assayed by reverse transcription with ^{32}P-labeled DT586 primer, and the reaction products were electrophoresed on a 6% polyacrylamide gel.

RNase I protection

RNase I (Promega, 10 U/ml) was diluted 2000- or 4000-fold in dialysis buffer and 5 µl of the dilution was added to 45 µl of splicing extract. The digestion was carried out at room temperature for 10 min and stopped by phenol–chloroform extraction. The RNA was ethanol precipitated and assayed by reverse transcription with U1-specific primer DT586. The primer extension products were electrophoresed on a 7.5% polyacrylamide gel.

Acknowledgements

We thank H.V.Colot for constructive comments on the manuscript, and T.Blumenthal for help with the *C.elegans* sequence. We thank P.Siliciano for providing us with the anti-U1-70K antibodies. This work was supported in part by a grant from the National Institutes of Health to M.R. (GM23549).

References

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

Abovich,N., Liao,X.C. and Rosbash,M. (1994) The yeast MUD2 protein: an interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. *Genes Dev.*, **8**, 843–854.

Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.

Ares,M.J. and Igel,A.H. (1990) Lethal and temperature-sensitive mutations and their suppressors identify an essential structural element in U2 small nuclear RNA. *Genes Dev.*, **4**, 2132–2145.

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

Cadwell,R.C. and Joyce,G.F. (1992) Randomization of genes by PCR mutagenesis. *PCR Methods Appl.*, **2**, 28–33.

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

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

Gribskov,M., McLachlan,A.D. and Eisenberg,D. (1987) Profile analysis: detection of distantly related proteins. *Proc. Natl Acad. Sci. USA*, **84**, 4355–4358.

Gunnewiek,J.M., van Aarssen,Y., Wassenaar,R., Legrain,P., van Venrooij,W.J. and Nelissen,R.L. (1995) Homodimerization of the human U1 snRNP-specific protein C. *Nucleic Acids Res.*, **23**, 4864–4871.

Guthrie,C. (1991) Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science*, **253**, 157–163.

Hamm,J., Dathan,N.A., Scherly,D. and Mattaj,I.W. (1990) Multiple domains of U1 snRNA, including U1 specific protein binding sites, are required for splicing. *EMBO J.*, **9**, 1237–1244.

Heinrichs,V., Bach,M., Winkelmann,G. and Lührmann,R. (1990) U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5′ splice site. *Science*, **247**, 69–72.

Kao,H.-Y. and Siliciano,P.G. (1992) The yeast homolog of the U1 snRNP protein 70K is encoded by the SNP1 gene. *Nucleic Acids Res.*, **20**, 4009–4013.

Kao,H.-Y. and Siliciano,P.G. (1996) Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. *Mol. Cell. Biol.*, **16**, 960–967.

Kohtz,J.D., Jamison,S.F., Will,C.L., Zuo,P., Lührmann,R., Garcia-Blanco,M.A. and Manley,J.L. (1994) Protein–protein interactions and 5′-splice-site recognition in mammalian mRNA precursors. *Nature*, **368**, 119–124.

Kretzner,L., Rymond,B.C. and Rosbash,M. (1987) *S.cerevisiae* U1 RNA is large and has limited primary sequence homology to metazoan U1 snRNA. *Cell*, **50**, 593–602.

Kretzner,L., Krol,A. and Rosbash,M. (1990) *Saccharomyces cerevisiae* U1 small nuclear RNA secondary structure contains both universal and yeast-specific domains. *Proc. Natl Acad. Sci. USA*, **87**, 851–855.

Lamond,A.I. (1993) The spliceosome. *BioEssays*, **15**, 595–603.

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

Legrain,P. and Choulika,A. (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.

Legrain,P., Chapon,C. and Galisson,F. (1993) Interactions between PRP9 and SPP91 splicing factors identify a protein complex required in prespliceosome assembly. *Genes Dev.*, **7**, 1390–1399.

Liao,X., Kretzner,L., Séraphin,B. and Rosbash,M. (1990) Universally conserved and yeast-specific U1 snRNA sequences are important but not essential for U1 snRNP function. *Genes Dev.*, **4**, 1766–1774.

Liao,X.C., Tang,J. and Rosbash,M. (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.

Lockhart,S.R. and Rymond,B.C. (1994) Commitment of yeast pre-mRNA to the splicing pathway requires a novel U1 small nuclear ribonucleoprotein polypeptide, Prp39p. *Mol. Cell. Biol.*, **14**, 3623–3633.

Lutz,C.S. and Alwine,J.C. (1994) Direct interaction of the U1 snRNP-A protein with the upstream efficiency element of the SV40 late polyadenylation signal. *Genes Dev.*, **8**, 576–586.

Lührmann,R., Kastner,B. and Bach,M. (1990) Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochim. Biophys. Acta*, **1087**, 265.

Madhani,H.D. and Guthrie,C. (1994) Dynamic RNA–RNA interactions in the spliceosome. *Annu. Rev. Genet.*, **28**, 1–26.

Mattaj,I.W. (1993) RNA recognition: a family matter? *Cell*, **73**, 837–840.

Michel,F. and Ferst,J.L. (1995) Structure and activities of group II introns. *Annu. Rev. Biochem.*, **64**, 435–461.

Nelissen,R.L.H., Heinrichs,V., Habets,W.J., Simons,F., Lührmann,R. and van Venrooij,W.J. (1991) Zinc finger-like structure in U1-specific protein C is essential for specific binding to U1 snRNP. *Nucleic Acids Res.*, **19**, 449–454.

Roy,J., Zheng,B., Rymond,B.C. and Woolford,J.L.,Jr (1995) Structurally related but functionally distinct yeast Sm D core small nuclear ribonucleoprotein particle proteins. *Mol. Cell. Biol.*, **15**, 445–455.

Rymond,B.C. and Rosbash,M. (1988) A chemical modification/interference study of yeast pre-mRNA spliceosome assembly and splicing. *Genes Dev.*, **2**, 428–439.

Saldanha,R., Mohr,G., Belfort,M. and Lambowitz,A.M. (1993) Group I and group II introns. *FASEB J.*, **7**, 15–24.

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

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

Siliciano,P.G., Jones,M.H. and Guthrie,C. (1987) *Saccharomyces cerevisiae* has a U1-like small nuclear RNA with unexpected properties. *Science*, **237**, 1484–1487.

Smith,V. and Barrell,B.G. (1991) Cloning of a yeast U1 snRNP 70K protein homolog: functional conservation of an RNA-binding domain between humans and yeast. *EMBO J.*, **10**, 2627–2634.

Tang,J. and Rosbash,M. (1996) Characterization of yeast U1 snRNP A protein: identification of the N-terminal RNA binding domain (RBD) binding site and evidence that the C-terminal RBD functions in splicing. *RNA*, **2**, 1058–1070.

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

Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.

Wu,J.Y. and Maniatis,T. (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell*, **75**, 1061–1070.

Received on February 17, 1997; revised on April 3, 1997