J. Mol. Biol. (1989) **208**, 371–379

Schizosaccharomyces pombe U4 Small Nuclear RNA Closely Resembles Vertebrate U4 and is Required for Growth

Thomas Dandekar†, Véronique Ribes† and David Tollervey†

Génie Microbiologique, Institut Pasteur
75724 Paris Cedex 15, France

(Received 23 February 1989)

The single-copy gene *snu4*, which encodes the small nuclear RNA (snRNA) U4, has been cloned and sequenced. *Schizosaccharomyces pombe* U4 is 128 nucleotides in length, similar in size to vertebrate U4 and shows substantial primary and secondary structure homology. The gene lacks sequences closely resembling vertebrate snRNA transcription signals, but has a TATA box at −33 to −30; TATA sequences flanked by several additional conserved nucleotides are found in the same position in the 5′ regions of other snRNA genes from *Schiz. pombe.*

The cloned *snu4* gene was disrupted by transposon mutagenesis and used to replace one chromosomal copy of *snu4* in a diploid strain. On sporulation *snu4*$^-$ haploid strains could not be recovered, demonstrating that U4 is required, at least for spore germination. Haploid *snu4*$^-$ strains are viable if they also carry *snu4*$^+$ on a replicating plasmid but are unable to loose the plasmid under non-selective growth, demonstrating a continuous requirement for U4 for viability.

1. Introduction

The small nuclear RNAs (snRNAs‡) are a group of stable low molecular weight RNA species found in the nuclei of all eukaryotes examined to date, which are believed to play roles in a variety of RNA processing reactions. The most studied species, designated U1, U2, U4, U5 and U6, are all required for pre-messenger RNA (pre-mRNA) splicing *in vitro* (for a review, see Maniatis & Reed, 1987). Splicing of pre-mRNA occurs within a complex structure termed a spliceosome. Major events in spliceosome assembly include the association of U1 and U2 small nuclear ribonucleoproteins (snRNPs) with the 5′ splice junction and site of branch formation, respectively. U4 and U6 snRNA are associated by hydrogen-bonding and are found in a single snRNP particle (Bringmann *et al.*, 1984; Hashimoto & Steitz, 1984) and may associate with the spliceosome as a U4/U5/U6 complex; U4 is released from the spliceosome before intron cleavage (Pikielny *et al.*, 1986; Cheng & Abelson, 1987; Lamond *et al.*, 1988). The site and mechanism of action of U4 snRNP remains obscure but it is essential for splicing activity *in vitro*. We wished therefore to use molecular genetic techniques to study the function of U4 snRNPs *in vivo*.

Genetic analyses in the budding yeast *Saccharomyces cerevisiae* (*S. cerevisiae*) have shown that the major spliceosome-associated snRNAs, including U4, are all essential for viability (Siliciano *et al.*, 1987; for a review, see Guthrie & Patterson, 1988). However, it is clear that, while many features are common between pre-mRNA splicing in *S. cerevisiae* and higher eukaryotes, there are also distinctive differences in the system of *S. cerevisiae*. We therefore considered it to be useful to analyse genetically the role of snRNPs in the fission yeast *Schizosaccharomyces pombe* (*Schiz. pombe*), in which both the RNA and protein components of the snRNPs as well as signals in pre-mRNA introns, are rather closer to those of higher eukaryotes (Tollervey & Mattaj, 1987; Käufer *et al.*, 1985; Brennwald *et al.*, 1988; D. Tollervey, G. Tessars & R. Lührmann, unpublished results).

We have previously reported the identification of snRNAs from *Schiz. pombe* including the analogues of vertebrate U1–U6 and the cloning of the genes encoding several such snRNAs, including the single-copy gene *snu4* (Tollervey & Mattaj, 1987; Tollervey, 1987; Dandekar & Tollervey, 1989).

Schiz. pombe U4 is similar in size to vertebrate U4 and has many features characteristic of snRNAs: (1) it is localized in the nucleus; (2) it carries the distinctive 2,2,7-trimethylguanosine cap structure

† Present address: European Molecular Biology Laboratory, Meyerhofstrasse 1, 6900 Heidelberg, F.R.G.
‡ Abbreviations used: snRNAs, small nuclear RNAs; snRNPs, small nuclear ribonucleoproteins; bp, base-pairs; kb, 10^3 bases or base-pairs.

0022-2836/89/150371-10 $03.00/0

© 1989 Academic Press Limited

unique to snRNAs; (3) it is in a ribonucleoprotein particle, associated with proteins that are immunologically related to the vertebrate "Sm-proteins" common to U1, 2, 4 and 5; and (4) it retains a conserved, functional binding site for such proteins since *Schiz. pombe* U4 is able to associate with *Xenopus* Sm-proteins following microinjection into *Xenopus* oocytes.

Since, both the RNA and protein components of *Schiz. pombe* U4 snRNPs appear to have been highly conserved with respect to those of vertebrates, it is likely to be of value in analysing the function of this snRNP *in vivo*. As a first step in such an analysis, we have determined whether U4 is required for viability.

2. Materials and Methods

(a) *Strains and media*

Growth and handling of *Schiz. pombe* have been described by Gutz *et al.* (1974). The diploid strain used for gene disruption carries h^+, *leu1.32/h*$^+$, *leu1.32; ura4*$^-$*D18/ura4*$^-$*D18; ade 6.704/ade 6.704* (P. Nurse, personal communication).

(b) *Chromosomal DNA*

DNA was prepared by a modification of the technique of Holm *et al.* (1986). Crude nuclei fractionated as described by Dandekar & Tollervey (1989) were used to prepare chromosomal DNA. Nuclei were resuspended in 2·2 ml of 100 mM-NaCl, 20 mM-Tris–HCl (pH 8), 50 mM-EDTA. To this was added 0·24 ml of 20% *N*-lauryl sarcosyl and 10 mg of proteinase K, and the tubes were incubated at 37 °C for 1 h. The lysate was cleared by centrifugation in an Eppendorf centrifuge and the supernatant was loaded onto 7·5 ml of saturated $CsCl_2$ in TE buffer (10 mM-Tris (pH 8), 1 mM-EDTA) and centrifuged in an SW41 rotor. Fractions containing chromosomal DNA, identified by their high viscosity, were pooled and dialysed against TE buffer.

(c) *Sequence analysis*

The sequence analysis was carried out using the assistance of the GCG software package (Devereux *et al.*, 1984). The alignments for the primary structure were obtained using the program "gap" of the above package for the pairwise comparisons to *Schiz. pombe* (gap weight 5·0, gap length weight 0·3) before the sequences were aligned together and the common alignment was further optimized.

(d) *RNA*

RNA was prepared by the hot phenol/guanidinium method (Maniatis *et al.*, 1982) as modified by Tollervey & Mattaj (1987). Immunoprecipitations of RNA were carried out as described by Tollervey & Mattaj (1987). The U4 RNA 3′ sequence was determined enzymically (Donis-Keller *et al.*, 1977) using gel-purified immunoprecipitated RNA. The 5′ RNA sequence was determined by primer extension as described by Geliebter (1987), using an oligonucleotide complementary to nucleotides 56 to 72 of *Schiz. pombe* U4.

(e) *DNA sequencing*

The sequence of the *snu4* gene was determined by dideoxynucleotide chain termination (Sanger *et al.*, 1977). The region shown was sequenced on both strands.

(f) *Transposon mutagenesis and transformation*

The method of transposon mutagenesis was as described by Seifert *et al.* (1986). The transposon inserted, m-Tn*3* (*LEU2*), carries the *S. cerevisiae* selective marker *LEU2*, which complements the *leu1*$^-$ mutation of *Schiz. pombe* (Beach & Nurse, 1981). *Schiz. pombe* was transformed using the protocol described for *S. cerevisiae* by Ito *et al.* (1983).

3. Results

(a) *Localization and sequence of* snu4

The cloning of the gene *snu4* and several other snRNA genes from *Schiz. pombe* has been reported (Dandekar & Tollervey, 1989). Southern hybridization using probe prepared from purified U4 was used to localize the gene to and within the 940 bp genomic *Hin*dIII–*Eco*RI fragment shown in Figure 1. This region was sequenced using dideoxy nucleotides. The 3′ end of the coding region was identified by comparison with the 3′ end RNA sequence obtained from purified U4 snRNA. The 5′ end of the coding region was identified by primer extension with an oligonucleotide complementary to nucleotides 56 to 72 of the RNA. These gave a single 3′ and a single 5′ end-point at the positions indicated. The length of the RNA given by these end-points is very close to our estimates of the size of U4 based on gel mobility. However, the slight possibility that the RNA might be a few nucleotides longer and that the stop in the primer extension reaction might be due to reasons other than the true end of the RNA cannot be completely excluded. The 128-nucleotide RNA sequence and the DNA sequence extending to 208 nucleotides 5′ and 463 nucleotides 3′ of the coding region are shown in Figure 2.

No sequences with high homology to the consensus vertebrate snRNA gene transcription signals were detected. However, a TATA sequence (underlined in Fig. 2) is present at −33 to −30, a position similar to that of TATA sequences of other genes of *Schiz. pombe* transcribed by polymerase II (Russel, 1983). The *snu4* gene carried by YEpsnu4 has only 220 nucleotides of 5′ flanking sequence. Northern hybridization shows that this is sufficient for full expression of the *snu4* gene (data not shown). Comparison of the 3′ flanking region of *snu4* with those of *snu2* (T. Dandekar & D. Tollervey, unpublished results; J. A. Wise, personal communication) and *snu3A* (Porter *et al.*, 1988) reveals a region of moderate sequence conservation (underlined in Fig. 2) at a position similar to that of the transcription termination signals of vertebrate U snRNA genes (for a review, see Dahlberg & Lund, 1988).

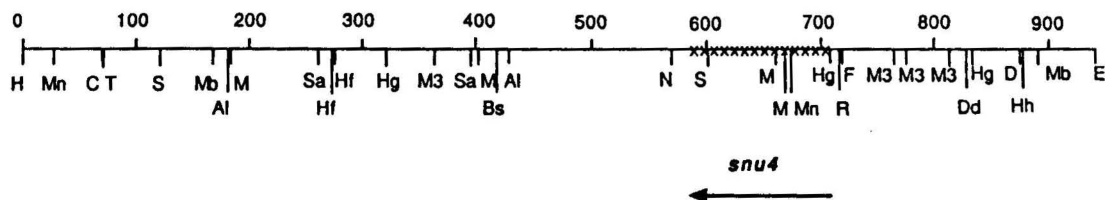

Figure 1. Restriction map of *snu4*. The restriction map of the region surrounding the genomic *snu4* locus is shown. C, *Cla*I; D, *Dra*I; Dd, *Dde*I; Al, *Alu*I; Bs, *Bsm*I; E, *Eco*RI; F, *Fok*I; H, *Hin*dIII; Hf, *Hin*fI; Hg, *Hgi*aI; Hh, *Hha*I; M, *Mae*I; M3, *Mae*III; Mb, *Mbo*II; Mn, *Mnl*I; S, *Ssp*I; Sa, *Sau*IIIa; T, *Taq*I; N, *Nde*I; R, *Rsa*I. The *snu4* coding region is hatched and an arrow indicates the direction of transcription.

(b) *U4 structure*

Comparison of the primary structure of U4 from *Schiz. pombe* and a number of other eukaryotes is shown in Figure 3. These RNAs are clearly homologous. The number of matched nucleotides is 77 for human, 72 for rat, 81 for chicken, 79 for *Drosophila*, 73 for bean and 82 for *S. cerevisiae*.

Figure 4 shows a comparison of a proposed secondary structure for human U4 (Rinke *et al.*, 1985) and a possible secondary structure for *Schiz. pombe* U4. Blocks of identical nucleotides are enclosed in boxes, and a potential Sm-binding site (AGUUUUGG) is drawn in bold type. Another sequence with homology to a consensus Sm-binding site (AAUAUUUGG) is present at positions 112 to 120 (underlined in Fig. 4), and in fact this sequence is aligned against the Sm-binding sites of other organisms in Figure 3. This sequence, however, has an A residue within the poly(U) tract, in contrast to

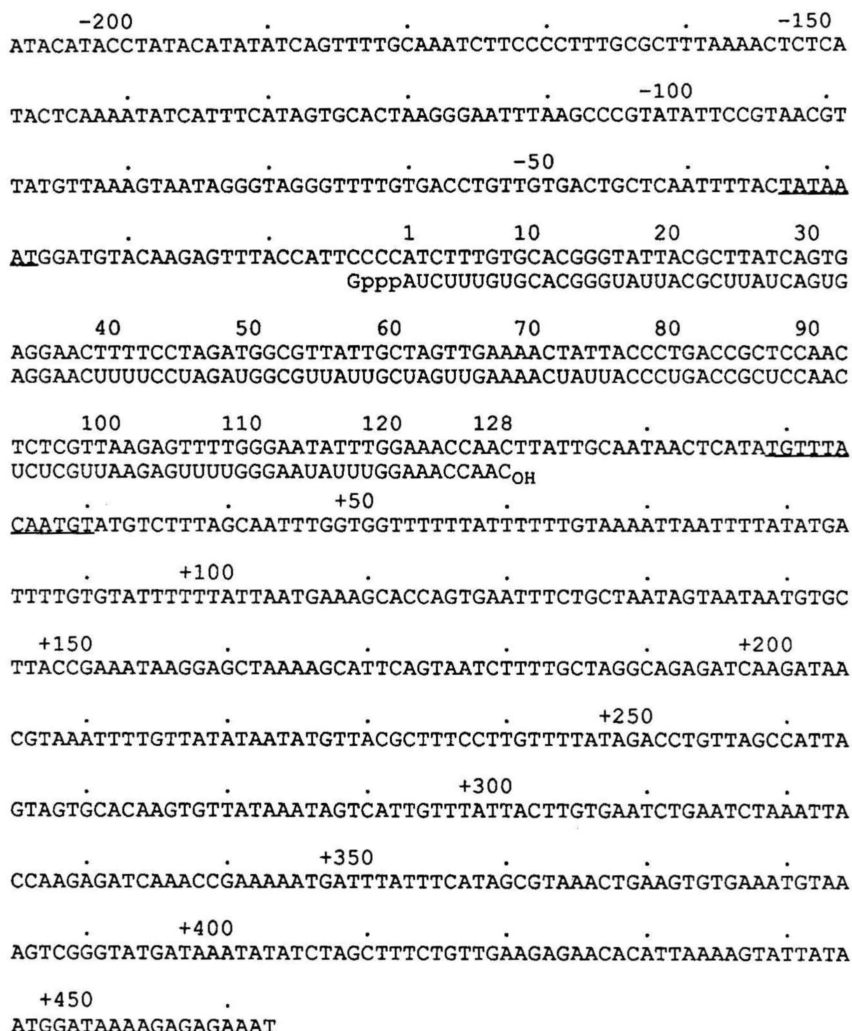

Figure 2. Sequence of *snu4*. The sequence of the *snu4* coding region is shown together with the flanking regions, extending to 208 nucleotides 5′ and 463 nucleotides 3′ of the coding region. The inferred RNA sequence is also shown together with its 5′ cap structure. A TATA-like sequence and a potential termination signal are underlined.

```
Sp.:.AUCUUUGUGCACGGGUAUUACGC.UUAUCAGUGAGGAACUUUUCCUAGAUGGCGUU
Hs.:.AGCUUUGCGCAGUGGCAGUAUCG.UAGCCAAUGAGG..UUUAUCCGAGGCGCGAUU
Rn.:GAACUUU..GCAGUGGCAGUAUCG.UAGCCAAUGAGG..UUUAUCCGAGGCGCGAUU
Gg.:.AGCUUUGCGCAGUGGCAGUAUCG.UAGCCAAUGAGG..UUUAUCCGAGGCGCGAUU
Dm.:.AGCUUAGCGCAGUGGCAAUACCG.UAACCAAUGAAG..CCUCCCUGAGGUGCGGUU
Vf.:.AUCUUUGCGCUUGGGGCAAUGACGUAGCUAGUGAGGU.UCUAACCGAGGCGCGUCU
Sc.:.AUCCUUAUGCACGGGAAAUACGCA.UAUCAGUGAGGA.UUCGUCCGAGAUUGUGUU

Sp.: AUUGCUAGUUGAAAACUAUUACCCUGACCGCUCCAA........CUCUCGUUAAGA
Hs.: AUUGCUAAUUGAAAACUUUUCCCAAUACCCCGCCGU........GACGACUUGCAA
Rn.: AUUGCUAAUUGAAAACUUUUCCCAAUACCCCGCCGU........GACGACUUGAAA
Gg.: AUUGCUAGUUGAAAACUUUUCCCAAUACCCCGCCAU........GACGACUUGAAA
Dm.: AUUGCUAGUUGAAAACU.UUAACCAACCCACGCCAU........GGGACGUGAAAU
Vf.: AUUGCUGGUUGAAAACUAUUUGCCAAACCCCUCUUAGGCUUGGGCUUGGGUUCAAG
Sc.: UUUGCUGGUUGAAAUUUAAUUAUAAACCAGACCGUC........UCCUCAUGGUCA

Sp.: ........GUUUUGGGAAUAUUUGGAAACCAAC_OH (128bp)
Hs.: UAUAGUCGGCAUUGGCAAUUUUUGACAGUCUCUACGGAGACUGG_OH
                                                      (145bp)
Rn.: UAUAGUCGGCAUUGGCAAUUUUUG......UCUACGGAGACUG_OH
                                                      (139bp)
Gg.: UAUAGUCGGCAUUGGCAAUUUUUGAGAGCCUCCCCGGAG_OH
                                                      (141bp)
Dm.: ACCGUCC.ACUACGGCAAUUUUUGGAAGCCCUUACGAGGGCUAA_OH
                                                      (143bp)
Vf.: ........CCUUUGAGAAUUUCUGGAAGGACUCCCUUUGGGGUAAAG_OH
                                                      (154bp)
Sc.: AUUCGUG*AUGUAGGGAAUUUUUGGAAUACCUUU_OH

          (* = UUCGCUUUGAAUACUUCAAGACCU)              (154bp)
```

Figure 3. Alignment of nucleotide sequences of U4 RNAs. The primary sequence of *Schiz. pombe* U4 was aligned with U4 snRNAs from a range of other eukaryotes using the computer program of Devereux *et al.* (1984). Organisms are: Sp, *Schiz. pombe*; Hs., *Homo sapiens*; Rn., *Rattus norvegicus*; Gg., *Gallus gallus* (chicken); Dm., *Drosophila melanogaster*; Vf., *Vicia faba* (broad bean); Sc., *S. cerevisiae*. A region of 24 nucleotides (shown in parentheses) was omitted from the *S. cerevisiae* sequence at the site marked by an asterisk in order to allow the Sm-binding site of *S. cerevisiae* to be aligned with other Sm-binding sites.

Sm-binding sites from other snRNAs (Reddy, 1988) including *Schiz. pombe* U2 (Brennwald *et al.*, 1988). If the sequence at nucleotides 112 to 120 is indeed the Sm-binding site, then the predicted secondary structure would more closely resemble that of *S. cerevisiae*, with stem/loop III extended and stem/loop IV absent. An alternative secondary structure for U4, proposed by Myslinski *et al.* (1984) on the basis of comparison with the *Drosophila* U4 sequence, is modified from that shown in Figure 4 by an extension of stem 1 and reduction of loop II. The structure derived is not compatible with the sequence of *Schiz. pombe* U4.

The loop II region is highly conserved from human to *Schiz. pombe* U4. To determine whether this region is also conserved in U4 snRNAs from other fungi an oligonucleotide complementary to nucleotides 56-72 of *Schiz. pombe* U4 was synthesized. This oligonucleotide hybridizes strongly to U4-sized RNAs from *Aspergillus nidulans*, *Neurospora crassa*, *Candida albicans*, *Saccharomyces diastaticus* and *S. cerevisiae* (data not shown).

Phylogenetic comparisons have led to the proposition of the "Y"-model for the structure of the U4/U6 interaction (Brow & Guthrie, 1988). The sequence of *Schiz. pombe* U6 has been reported

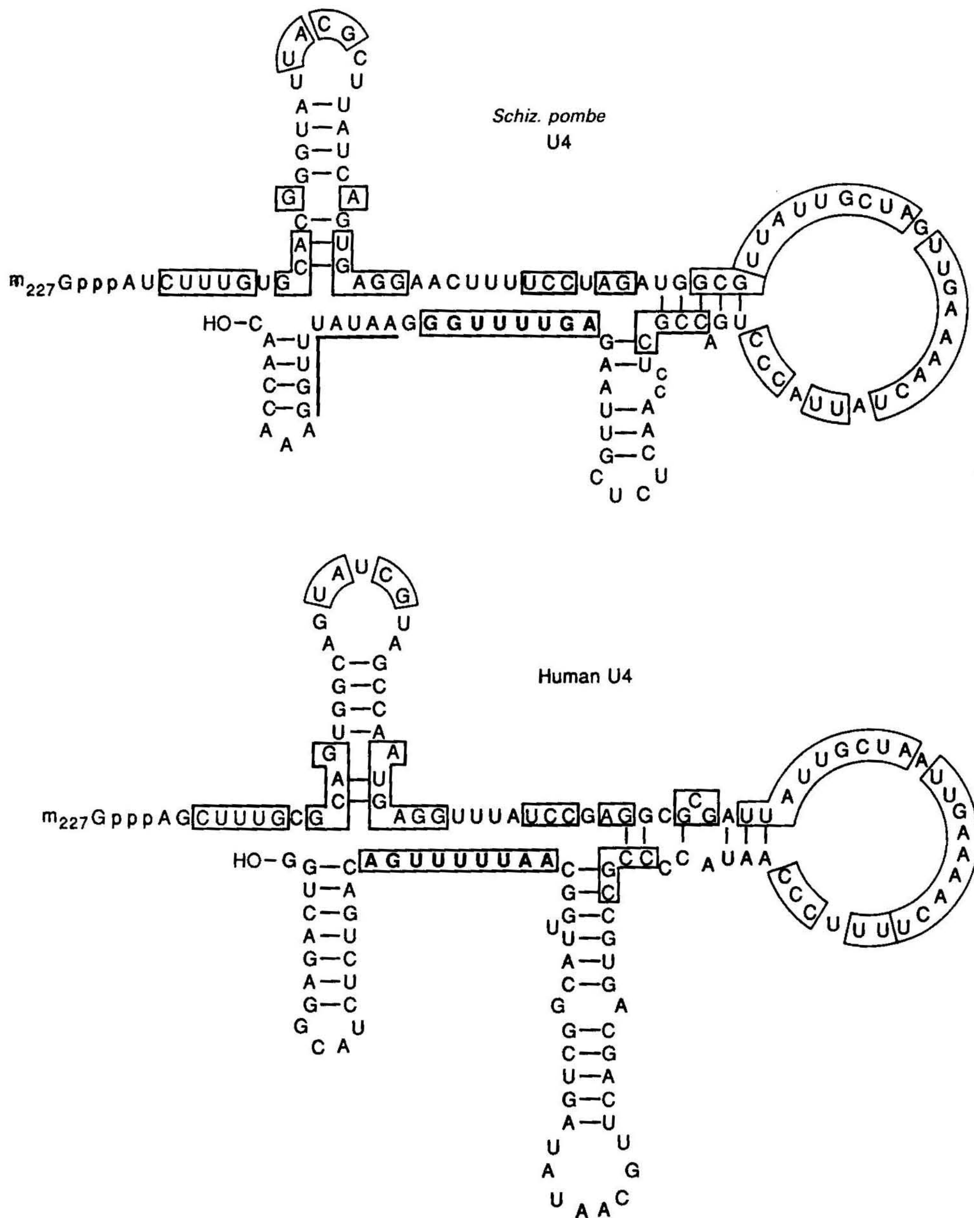

Figure 4. Comparison of predicted secondary structures of *Schiz. pombe* and human U4. Human U4 is drawn according to the secondary structure of Rinke *et al.* (1985). *Schiz. pombe* U4 is drawn to resemble this structure. Boxed regions contain identical sequences. The putative Sm-binding site is shown in bold type and a second sequence with homology to the Sm-binding site is underlined (see the text).

(Tani & Ohshima, 1989). U4 and U6 from *Schiz. pombe* can be drawn in a Y-structure closely resembling that of human U4/U6 (Fig. 5) with a stable stem formed between nucleotides 19 to 28 and 47 to 56 of *Schiz. pombe* U4 giving a loop exactly the size of that in human U4. The proposed *Schiz. pombe* Y structure also gives stable stems between nucleotides 1 and 16 of U4 and 53 and 68 of U6 and nucleotides 58 and 65 of U4 and nucleotides 43 and 50 of U6.

(c) *Disruption of* snu4

The gene *snu4* has been shown to be present in a single copy in haploid cells (Dandekar & Tollervey, 1989). To determine whether this gene is required for viability, a gene replacement experiment was performed. Cloned *snu4* was disrupted by insertion of a Tn*3* transposon carrying AMP^R and the *LEU2* gene from *S. cerevisiae*. Plasmids in which the transposon is inserted into the coding region were identified by Southern hybridization with cDNA probe prepared from purified U4. The disrupted copy of *snu4* was used to replace one chromosomal copy of *snu4* in a diploid strain homozygous for *leu1.32*, which is complemented by the *LEU2* gene carried by the transposon (Beach & Nurse, 1981; Seifert *et al.*, 1986; Ribes *et al.*, 1988). Of four transformed diploids analysed by genomic Southern blots, three gave the pattern of bands predicted for

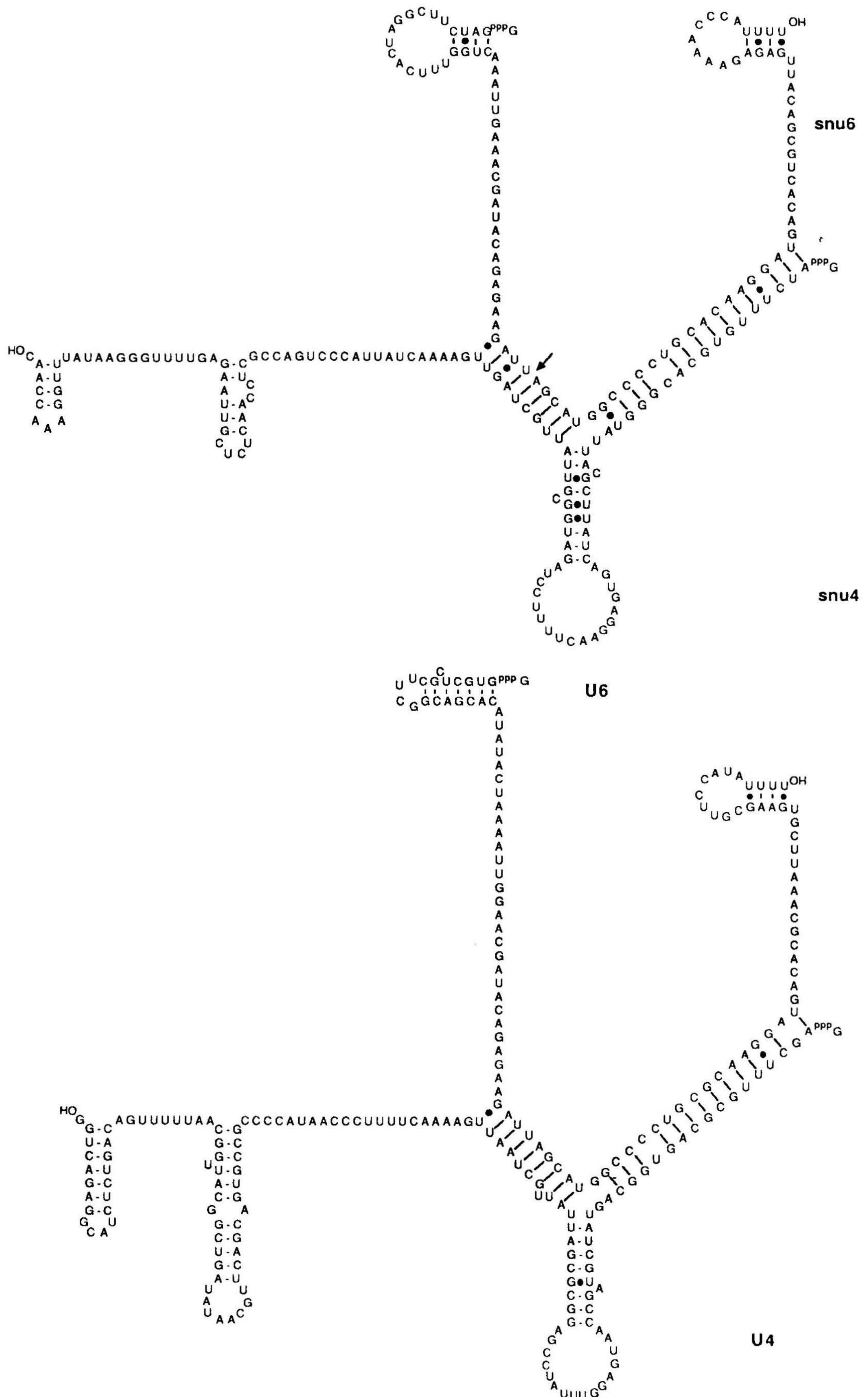

Figure 5. Proposed interaction of U4 and U6. *Schiz. pombe* U4 and U6 (upper panel) and human U4 and U6 (lower panel) are drawn in the Y-structure proposed by Brow & Guthrie (1988). The site of the intron in *Schiz. pombe* U6 (Tani & Ohshima, 1989) is indicated by an arrow. Two potential stem/loop structures, conserved in position between both U6 sequences are drawn at the 5′ and 3′ ends of human and *Schiz. pombe* U6. The positions of Watson–Crick base-pairs (–) and G · U base-pairs (●) are indicated.

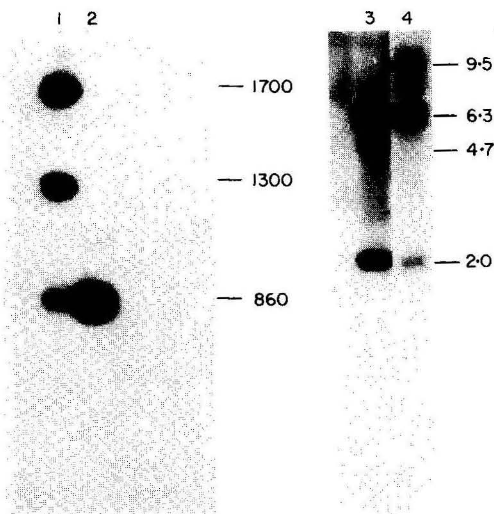

Figure 6. Southern analysis of strains carrying the *snu4* gene disruption and a complementing plasmid. Lane 1, diploid heterozygous for *snu4*:Tn*3(LEU2)/snu4*$^+$; lane 2, wild-type diploid; lane 3, diploid heterozygous for *snu4*::Tn*3(LEU2)/snu4*$^+$; lane 4, haploid strain carrying *snu4*::Tn*3(LEU2)* and the complementing plasmid YEpsnu4. Chromosomal DNA from each strain was digested with *Eco*RI and *Cla*I (lanes 1 and 2) or *Hin*dIII (lanes 3 and 4). Estimated sizes of the bands are indicated in nucleotides on the left and kb on the right. The hybridization probe was prepared from a fragment of DNA extending from the 5′ end of the *snu4* gene to the *Ssp*I site within the coding region (Fig. 1).

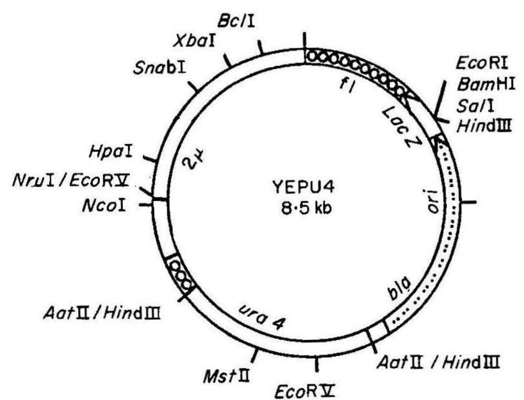

Figure 7. Plasmid YEpU4. The plasmid was constructed by deletion of the *Nru*I–*Eco*RV fragment carrying the *URA3* gene from the plasmid pEMBLYe23 (Cesareni & Murray, 1987) and insertion of a 1·8 kb *Hin*dIII fragment carrying the *ura4* gene of *Schiz. pombe* as a blunt-ended fragment at the single *Aat*II site. Unique restriction sites predicted from the sequences are shown. Restriction digests confirm that all sites shown are present on the plasmid. In *Escherichia coli* the plasmid confers ampicillin resistance, contains the *lacZ* complementing fragment, which permits blue/white colour testing for insertions and carries the F1 replication origin, enabling the single-stranded form of the plasmid to be isolated (orientation as for transcription of *lacZ*) by co-infection and packaging by phage F1. The 1·8 kb *ura4* fragment inserted in YEpU4 corresponds to the region precisely deleted in *ura4-D18* strains (Grimm *et al.*, 1988) and the plasmid therefore has no regions of homology to genomic DNA in such strains.

correct gene replacement at the *snu4* locus (Fig. 6). Digestion with *Eco*RI and *Cla*I gives a 860 nucleotide genomic DNA fragment (Fig. 1), which is visible in Figure 6, lanes 1 and 2. In the transformed diploid (Fig. 6, lane 1) additional bands of 1·3 and 1·7 kb are expected, due to cleavage at the *Eco*RI sites within the transposon (Seifert *et al.*, 1986). The hybridization probe is to a region entirely within the coding sequence of *snu4*, demonstrating that this region has indeed been interrupted by the insertion. The growth of the diploid carrying one disrupted copy of *snu4* is indistinguishable from that of the wild-type parent. Densitometer tracing of Northern hybridizations demonstrate that the ratio of U2/U4 in these strains is identical, indicating that the regulation of *snu4* is such as to provide complete dosage compensation.

The transformed diploid is homozygous h^+/h^+ at the mating-type locus. To obtain haploid spores, rare h^{90}/h^+ revertants were identified (Gutz *et al.*, 1974), which switch mating type to h^-/h^+ and sporulate. Amongst 2700 haploid spores, none was found to carry the inserted *LEU*$^+$ marker, demonstrating that *snu4* is essential, at least for spore germination.

(d) *Genetic complementation with* snu4

To determine whether *snu4* is also required for vegetative growth, and to confirm fully that the lack of U4 is responsible for the lack of growth of the *LEU*$^+$ haploids, a complementation experiment was performed.

A *Schiz. pombe* replicating plasmid, YEpU4, has been constructed (Fig. 7) that carries the *S. cerevisiae* 2 μ origin of replication and the *ura4* gene from *Schiz. pombe*, together with the phage F1 replication origin, which permits the isolation of single-stranded DNA molecules suitable for DNA sequencing or *in vitro* mutagenesis. The 940 bp *Eco*RI–*Hin*dIII region containing *snu4* (Fig. 1) was subcloned into YEpU4 to generate YEpsnu4. The sequence of the *snu4* fragment is not compatible with it containing an intact protein gene, and this region does not hybridize to other small RNA species. The *snu4*::Tn*3(LEU2)/snu4*$^+$ diploids also carry *ura4-D18/ura4-D18* (Grimm *et al.*, 1988), and YEpsnu4 was transformed into such a strain. Spores from the diploid carrying YEpsnu4 were plated on medium lacking uracil to select for those that had retained the plasmid during meiosis, and replicated onto medium with and without leucine. Of 90 spores tested, 53% carried the inserted LEU2 marker. Southern hybridization of two separate *LEU2*$^+$ strains confirmed that they lack an intact genomic copy of *snu4* (Fig. 6). The transformed diploid (Fig. 6, lane 3) and the *snu4*$^-$ haploid (Fig. 6, lane 4) give rise to *Hin*dIII fragments of 2·0 and 6·3 kb, due to the disruption of the 4·7 kb

genomic *Hin*dIII fragment carrying *snu4* by the transposon (3·6 kb) which contains a single *Hin*dIII site (Seifert *et al.*, 1986). In the diploid strain (lane 3), but not in the haploid strain (lane 4), there is an intact chromosomal copy of the *snu4* gene, which gives rise to the 4·7 kb genomic *Hin*dIII fragment. In the haploid strain (lane 4) the 9·5 kb band is the complementing YEpsnu4 plasmid, which carries the intact *snu4* allele, linearized by cleavage at the unique *Hin*dIII site (Fig. 7). The $snu4^-$ strain shown in Figure 6 (lane 4) and a $snu4^+$ sister strain were grown without selection for the YEpsnu4 plasmid for approximately 12 generations. Cells from each culture were plated on non-selective medium and replicated onto media with and without uracil, to test for the presence of the plasmid. From 900 $snu4^+$ colonies, 150 (17%) were found to retain the plasmid, a value in good agreement with our previous estimates of the stability of this plasmid. From 900 $snu4^-$ colonies, 100% retained the YEpsnu4 plasmid. The same result was obtained in three independent experiments.

4. Discussion

U4 snRNA from *Schiz. pombe* shows high primary sequence homology to U4 snRNAs from vertebrates, insects, plants and the budding yeast *S. cerevisiae*. Perhaps surprisingly, overall homology of *Schiz. pombe* U4 compared with any of the other species is similar, with generally around 60 to 63% matched nucleotides. U4 from *Schiz. pombe* can be drawn in secondary structure very similar to the structure of "free" U4 (i.e. not associated with U6) proposed by Rinke *et al.* (1985). In addition, *Schiz. pombe* U4 and U6 can be drawn in a Y structure as proposed for the U4/U6 interaction (Brow & Guthrie, 1988), which also very closely resembles that of human U4/U6.

In current models for spliceosome function, the active enzymic components are RNA molecules, the snRNAs and the pre-mRNA. The high evolutionary conservation of U6 and the release (or less-tight association) of U4 from the spliceosome prior to intron cleavage (Pikielny *et al.*, 1986; Cheng & Abelson, 1987; Lamond *et al.*, 1988) has led to the further suggestion that U6 is the enzymically active component, and that the role of U4 is to act as a carrier for U6 and to block its activity until the appropriate moment. However, it is notable that only a part of the region of highest primary sequence conservation between U4 from *Schiz. pombe* and vertebrates corresponds to the U4/U6 interaction domain. The very high evolutionary conservation of the whole of the loop II region of U4 strongly suggests that U4 plays some more active role in spliceosome function.

Following microinjection into *Xenopus* oocytes, *Schiz. pombe* U4 becomes associated with *Xenopus* snRNP proteins that are stockpiled in the cytoplasm (Zeller *et al.*, 1983; Tollervey & Mattaj, 1987). This demonstrates that *Schiz. pombe* U4 retains a functional "Sm-binding site": a short nucleotide sequence, which is apparently sufficient to permit the association of RNA species with the "Sm-proteins" that are common between U1, 2, 4 and 5 (Branlant *et al.*, 1982; Mattaj & De Robertis, 1985). A sequence (AGUUUUGG) similar to a consensus Sm-binding site is found in a single-stranded region of *Schiz. pombe* U4.

Neither the *Xenopus* U2 nor *S. cerevisiae* U5 genes are detectably expressed in *Schiz. pombe* (D. Tollervey, unpublished results). Consistent with this, we detect no sequences with good homology to the vertebrate snRNA gene consensus distal sequence element (DSE), which acts as an enhancer, or to the proximal sequence element (PSE), which fulfils the role of a TATA box. Genes encoding snRNAs from *S. cerevisiae* also lack sequences homologous to PSE and DSE elements, but have well-conserved TATA sequences located between −80 and −100 (for a compilation of such sequences, see Bally *et al.*, 1988), a position similar to that of TATA boxes for several other genes transcribed by RNA polymerase II in *S. cerevisiae*. The *Schiz. pombe snu4* gene does not have a TATA box in this region but does have a TATA sequence at −33 to −30, a position similar to that of TATA sequences of other polymerase II genes from *Schiz. pombe* (Russel, 1983). It is striking that a TATA sequence is found in identical positions in the *Schiz. pombe snu2* gene (D. Tollervey & T. Dandekar, unpublished results; J. A. Wise, personal communication) and, at a similar position in the *snu3A* gene (Porter *et al.*, 1988); several additional nucleotides around the TATA box are also conserved between these genes, lending increased significance to the conservation of this region. Comparison of six genes encoding U2 from a plant, *Arabidopsis*, reveals that these too have a conserved TATA sequence, surrounded by a number of other completely conserved nucleotides, situated at a distance from the initiation site identical with that of the *Schiz. pombe* TATA sequences (Vankan & Filipowicz, 1988).

As the first step in our genetic analysis of the function of U4 *in vivo*, we have demonstrated that it is essential for viability. The failure to recover $snu4^-$ haploids from a diploid carrying only one intact copy of the *snu4* gene demonstrates that the *snu4* gene is at least required for spore viability or germination. Haploid $snu4^-$ strains can be recovered if they also carry an intact copy of *snu4* on a replicating plasmid, demonstrating that the failure of $snu4^-$ strains to grow is due to a lack of U4 snRNA. These strains are unable to lose the complementing plasmid over many generations of unselected growth, demonstrating that U4 is continuously required for vegetative growth. These experiments also provide the genetic proof that *snu4* is the only gene encoding U4 in *Schiz. pombe* and that the gene-disruption experiments have indeed resulted in replacement at the *snu4* locus.

The techniques used here should be of general use in analysing the structure and function of genes in

Schiz. pombe. The transposon mutagenesis system, developed for *S. cerevisiae* (Seifert *et al.*, 1986) enables essentially random insertions to be made in any gene, even if the sequence is not known. This is particularly useful for making insertions into snRNA genes, which, owing to their small size, frequently lack convenient restriction sites. The plasmid we have constructed permits the complementation of the disrupted gene in haploid cells. Since it carries the phage F1 replication origin, this plasmid can be used directly to construct and sequence *in vitro* mutations for complementation testing.

We thank Jörg Kohli and Kinsey Maundrell for providing the cloned *ura4* gene and sequence, and Paul Russel and Paul Nurse for providing the strain used for the gene disruption experiment. T.D. was supported by the Deutsche Forschungsgemeinschaft.

References

Bally, M., Hughes, J. & Cesareni, G. (1988). *Nucl. Acids Res.* **16**, 5291–5303.

Beach, D. H. & Nurse, P. M. (1981). *Nature (London)*, **290**, 140–142.

Branlant, C., Krol, A., Ebel, J. P., Lazar, E., Haendler, B. & Jacob, M. (1982). *EMBO J.* **1**, 1259–1262.

Brennwald, P., Porter, G. & Wise, J.-A. (1988). *Mol. Cell. Biol.* **8**, 5575–5580.

Bringmann, P., Appel, B., Rinke, J., Reuter, R., Theissen, H. & Lührmann, R. (1984). *EMBO J.* **3**, 1357–1363.

Brow, D. & Guthrie, C. (1988). *Nature (London)*, **334**, 213–218.

Cesareni, G. & Murray, J. A. H. (1987). In *Genetic Engineering* (Setlow, J. K., ed.), vol. 9, pp. 135–154, Plenum Publ. Corp., New York.

Cheng, S.-C. & Abelson, J. (1987). *Genes Develop.* **1**, 1014–1027.

Dahlberg, J. E. & Lund, E. (1988). In *Small Nuclear Ribonucleoprotein Particles* (Birnstiel, M. L., ed.), Springer-Verlag, Berlin and Heidelberg.

Dandekar, T. & Tollervey, D. (1989). *Gene*, in the press.

Devereux, J., Haeberli, P. & Smithies, O. (1984). *Nucl. Acids Res.* **12**, 387–395.

Donis-Keller, H., Maxam, A. M. & Gilbert, W. (1977). *Nucl. Acids Res.* **4**, 2527–2538.

Geliebter, J. (1987). *Focus*, **9**, 5–8.

Grimm, C., Kohli, J., Murray, J. & Maundrell, K. (1988). *Mol. Gen. Genet.* **215**, 81–86.

Guthrie, C. & Patterson, B. (1988). *Annu. Rev. Genet.*, **22**, 387–419.

Gutz, H., Heslot, H., Leupold, U. & Loprieno, N. (1974). In *Handbook of Genetics* (King, R. C., ed.), vol. 1, pp. 395–446, Plenum Press, New York.

Hashimoto, C. & Steitz, J. A. (1984). *Nucl. Acids Res.* **12**, 3283–3293.

Holm, C., Meeks-Wagner, D. W., Fangman, W. L. & Botstein, D. (1986). *Gene*, **42**, 169–173.

Ito, H., Fukuda, Y., Murata, K. & Kimura, A. (1983). *J. Bacteriol.* **153**, 163–168.

Käufer, N. F., Simanis, V. & Nurse, P. (1985). *Nature (London)*, **318**, 78–80.

Lamond, A. I., Konarska, M. M., Grabowski, P. J. & Sharp, P. A. (1988). *Proc. Nat. Acad. Sci., U.S.A.* **85**, 411–415.

Maniatis, T. & Reed, R. (1987). *Nature (London)*, **325**, 673–678.

Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). Editors of *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Mattaj, J. W. & De Robertis, E. M. (1985). *Cell*, **40**, 111–118.

Myslinski, E., Branlant, C., Wieben, E. D. & Pederson, T. (1984). *J. Mol. Biol.* **180**, 927–945.

Pikielny, C. W., Rymond, B. C. & Rosbash, M. (1986). *Nature (London)*, **324**, 341–345.

Porter, G. L., Brennwald, P. J., Holm, K. A. & Wise, J. A. (1988). *Nucl. Acids Res.* **16**, 10131–10152.

Reddy, R. (1988). *Nucl. Acids Res.* 16 *(Suppl.)*, r71–r85.

Ribes, V., Dehoux, P. & Tollervey, D. (1988). *EMBO J.* **7**, 231–237.

Rinke, J., Appel, B., Digweed, M. & Lührmann, R. (1985). *J. Mol. Biol.* **185**, 721–731.

Russel, P. (1983). *Nature (London)*, **301**, 167–169.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 5463–5467.

Seifert, H. S., Chen, E. Y., So, M. & Heffron, F. (1986). *Proc. Nat. Acad. Sci., U.S.A.* **83**, 735–739.

Siliciano, P. G., Brow, D. A., Roiha, H. & Guthrie, C. (1987). *Cell*, **50**, 585–592.

Tani, T. & Ohshima, Y. (1989). *Nature (London)*, **337**, 87–90.

Tollervey, D. (1987). *J. Mol. Biol.* **196**, 355–361.

Tollervey, D. & Mattaj, I. W. (1987). *EMBO J.* **6**, 469–476.

Vankan, P. & Filipowicz, W. (1988). *EMBO J.* **7**, 791–799.

Zeller, R., Nyffenegger, T. & De Robertis, E. M. (1983). *Cell*, **32**, 425–434.

Edited by M. Yanagida