MOLECULAR AND CELLULAR BIOLOGY, Oct. 1992, p. 4456–4463
0270-7306/92/104456-08$02.00/0

Vol. 12, No. 10

# Accumulation of U14 Small Nuclear RNA in *Saccharomyces cerevisiae* Requires Box C, Box D, and a 5′, 3′ Terminal Stem

GUYANG M. HUANG, ARTUR JARMOLOWSKI,† JOACHIM C. R. STRUCK,‡ AND MAURILLE J. FOURNIER*

*Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, Massachusetts 01003*

Received 27 January 1992/Returned for modification 18 March 1992/Accepted 15 July 1992

**U14 is one of several nucleolar small nuclear RNAs required for normal processing of rRNA. Functional mapping of U14 from *Saccharomyces cerevisiae* has yielded a number of mutants defective in U14 accumulation or function. In this study, we have further defined three structural elements required for U14 accumulation. The essential elements include the U14-conserved box C and box D sequences and a 5′, 3′ terminal stem. The box elements are coconserved among several nucleolar small nuclear RNAs and have been implicated in binding of the protein fibrillarin. New mutational results show that the first GA bases of the box C sequence UGAUGA are essential, and two vital bases in box D have also been identified. An intragenic suppressor of a lethal box C mutant has been isolated and shown to contain a new box C-like PyGAUG sequence two bases upstream of normal box C. The importance of the terminal stem was confirmed from new compensatory base changes and the finding that accumulation defects in the box elements can be complemented by extending the terminal stem. The results suggest that the observed defects in accumulation reflect U14 instability and that protein binding to one or more of these elements is required for metabolic stability.**

U14 is one of several nucleolar small nuclear RNAs (snRNAs) linked to processing of rRNA (reviewed in references 37 and 40). Loss of U14 (or U3) in *Saccharomyces cerevisiae* results in impaired processing of 35S pre-rRNA and severe underaccumulation of mature 18S RNA (10, 19; depletion of snR10 also disrupts rRNA processing [36]). Mutational analysis of yeast U14 has thus far identified three sequence elements required for U14 function and two others required for accumulation (12, 18). Accumulation appears to also require a terminal stem domain, a situation indicated by expression analyses with hybrid yeast-mouse U14 RNAs (18). Taken together, the results suggest that metabolic instability is the likely basis for U14 loss, but interference with expression is possible. This study was undertaken to better define the structural elements required for U14 accumulation and to gain insights into the basis of U14 loss.

The two segments essential for accumulation occur near each terminus and include conserved sequence elements known as box C and box D. These elements are two of four sequences noted to be conserved among various U3 RNA species and designated boxes A through D (41). Boxes A and B appear to be unique to U3, whereas boxes C and D occur together in several other nucleolar snRNAs (39; this report). Initial evidence indicating that box C is vital for U14 production came from analysis of a point mutation (A9→G) in the conserved sequence that abolished accumulation (12) (see below).

Box C of human U3 has been linked to binding of fibrillarin, an abundant nucleolar protein commonly associated with nucleolar snRNAs. This role was initially suggested by results showing that fibrillarin antibody precipitated box C- and box D-containing oligonucleotides produced by nuclease treatment of native U3 ribonucleoprotein (RNP) complexes (24). Recent work has shown that complexes of U3 and fibrillarin can be formed in extracts and that mutations in box C interfere with fibrillarin association (3).

Fibrillarin antibodies have been shown to precipitate all of the metazoan and yeast snRNAs linked to rRNA processing as well as several others. The reactive snRNAs include mammalian U3, U8, U13, U14, X, and Y and all 10 nucleolar species found thus far in *S. cerevisiae*, including U14 (37, 39). This property suggests that fibrillarin may be a common component of snRNPs that participate in ribosome biogenesis (20, 39). Consistent with this notion is the finding that repression of fibrillarin synthesis in yeast cells leads to reduced levels of U14 (and snR190) and disruption of rRNA processing (32). Also consistent with this view is our earlier finding that a point mutation in box C leads to loss of U14 (12). Direct binding of fibrillarin to snRNA has not yet been shown in any case.

This study defines features of box C, box D, and the terminal stem that are required for accumulation of yeast U14. The importance of box C was verified by identification of an internal suppressor mutation that corrects a lethal box C defect, apparently by creating a new box C-like sequence at an adjoining site. Finally, it was determined that the essential box C and D functions could be provided by extension of the stem. The results argue that instability is the basis of U14 loss and that one or more of the essential elements is required for binding of protein(s) with stabilizing function.

* Corresponding author.
† Present address: European Molecular Biology Laboratory, 6900 Heidelberg, Germany.
‡ Present address: Henning Berlin GmbH, D-1000 Berlin, Germany.

## MATERIALS AND METHODS

**Strains.** The yeast Gal$^d$ test strain is YS153 (*GAL1*::*SNR128*) containing a galactose-inducible U14 allele de-

scribed previously (12). Nuclear extracts were prepared from protease-deficient strain BJ2407 (43). Cell culturing conditions have been described elsewhere (12).

**Mutagenesis.** Oligonucleotide-directed mutagenesis was carried out with a kit supplied by Bio-Rad (Muta-gene phagemid in vitro mutagenesis kit) according to the instructions provided. The U14 target DNA was in a phagemid vector, pJZ45, containing (i) a 1.3-kb *Cla*I fragment encoding both the *SNR190* and *SNR128* (U14) genes (12), (ii) the yeast *TRP-ARS* elements for replication and selection, and (iii) *CEN3* for maintenance at single copy. Point mutations in boxes C and D were created with eight primer mixes containing equal amounts of correct and incorrect nucleotides (9, 17). The box C multibase substitution mutant was described previously (12), and the same procedure was used to prepare the 3′ insertions in this study, with oligonucleotides of 37 and 40 bases. Mutations were identified by sequencing of *Escherichia coli*-derived plasmids, and functional variants were resequenced after recovery from yeast cells.

**Isolation of suppressor mutants.** A collection of 59 Gal$^i$ isolates was assembled from a culture of ethyl methanesulfonate-treated YS153 cells containing pJZ45 with a *SNR128*/A9→T gene and pLGSD5 with a *GAL10*::*lacZ* allele (8). Cells were spread on plates with glucose and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal), with blue/white screening for glucose repression defects (31); three isolates were blue and excluded from further consideration. Temperature sensitivity was assessed by restreaking at 25, 32, and 37°C on glucose and galactose. Total DNA was prepared and subjected to polymerase chain reaction (PCR) amplification with primers specific for the plasmid-encoded U14/A9→U gene. The DNA was purified and used for double-stranded sequencing to screen for intragenic suppressor mutations.

**PCR and sequencing.** PCR was done with a Perkin-Elmer/Cetus DNA Thermal Cycler. Primers (100 pmol of each) corresponding to bases −99 to −70 and 37 to 66 were added to 1 μg of cell DNA and amplified with *Taq* polymerase (30 cycles of 94°C for 1 min, 55°C for 2 min, and 72°C for 2 min). Product DNA was gel purified and sequenced with a T7 polymerase sequencing kit from Pharmacia.

**RNA analysis.** Procedures for preparing yeast RNA have been described elsewhere (19, 42). U14 levels were estimated by Northern (RNA) hybridization analysis. Approximately 0.4 $A_{260}$ unit of total RNA was fractionated on gels of 10% polyacrylamide–8.3 M urea, transferred to a nylon membrane (Zeta-Probe [Bio-Rad] or GeneScreen [New England Nuclear]), and probed with a $^{32}$P-labeled oligonucleotide complementary to U14 bases 42 to 73 (19). Probes for snR190 and U1 were complementary to bases 158 to 190 and 132 to 160, respectively. The ratio of 18S and 25S rRNAs was determined by densitometric scanning of films of formaldehyde-agarose gels stained with ethidium bromide (19).

**Preparation of nuclear extracts.** Extracts were prepared from a 1-liter YPD culture as follows (2, 22). Cells were grown to 1.0 $A_{600}$, washed with SP buffer (1 M sorbitol, 2% glucose, 0.2% yeast nitrogen base, 0.2% casein amino acids, 100 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES], 50 mM Tris-HCl [pH 8.0]), resuspended in 20 ml of SP buffer containing 2 mM dithiothreitol, and incubated with 10 mg of Lyticase (Sigma) for 1 h at 30°C. The resulting spheroplasts were centrifuged through a cushion of 0.6 M sorbitol–0.6 M sucrose–2% Ficoll–20 mM morpholine ethanesulfonic acid (MES)-Tris (pH 6.5) and resuspended in 10 ml of lysis buffer (10 mM HEPES-HCl [pH 7.0], 10 mM KCl, 1.5 mM $MgCl_2$, 20 μg of heparin [Sigma] per ml, 20 μg of phenylmethylsulfonyl fluoride 20 μg [Sigma] per ml), incubated for 5 min at room temperature, and homogenized by 25 strokes with a 15-ml Dounce homogenizer (pestle B). The homogenate was layered on a prespun (50 min, 27,000 × *g*) gradient consisting of 32.5% Percoll (Sigma) and 67.5% Percoll buffer (1 M sorbitol, 0.5 mM $CaCl_2$ [pH 7.5]) and centrifuged for 40 min at 9,500 × *g*. The nuclei were recovered, washed with Percoll buffer, resuspended in 3 ml of immunoprecipitation buffer (IP buffer; 150 mM NaCl, 50 mM Tris-HCl [pH 7.5], 0.05% Triton X-100), and sonicated three times for 10 s each time at setting 3 on a Branson B15 cell disruptor sonicator with microtip. Glycerol was added to 5%, nuclear debris was removed by centrifugation (10 min, 15,000 × *g*), and 50 to 60 μl of the supernatant (nuclear extract) used for immunoprecipitation in IP buffer.

**Immunoprecipitation analysis.** RNA association with fibrillarin was estimated by immunoprecipitation of nuclear extracts with a mouse monoclonal antibody for fibrillarin (72B9 [30]) or human autoimmune sera determined to be specific for nucleolar antigens (Vitrotec, Carlsbad, Calif.) and reactive with U14 RNP (35). Protein A-Sepharose beads (2.5 mg per sample; Pharmacia) preswollen in IP buffer at room temperature were washed three times (1 ml of IP buffer) and suspended in 500 μl of IP buffer. Ten microliters of antibody was added, and the mix was incubated for 90 min at 4°C with rocking. Beads were washed three times with 1 ml of IP buffer, 20 $A_{260}$ units (about 50 to 60 μl) of nuclear extract was added, and the samples were incubated for 90 min at 4°C with rocking. After five washes with IP buffer, the beads were resuspended in 200 μl of oyster glycogen (100 μg/ml; Sigma), phenol extracted, and precipitated with ethanol. Entire samples were fractionated by electrophoresis on gels of 8% polyacrylamide–urea, transferred to GeneScreen filters (New England Nuclear), and probed for snRNA by Northern hybridization analysis.

## RESULTS

**Conservation of box C.** Variants of the initially defined box C sequence can be found in 10 U3 species from diverse sources and in 4 other nucleolar snRNAs (Table 1). The non-U3 RNAs include mammalian U8, U13, and U14 and yeast U14 and snR190. The number of bases strictly conserved among the box C segments ranges from five to seven for the following comparisons: (i) all U3 species (10 RNAs, five bases [GAUGA]), (ii) all mammalian species (4 RNAs, seven bases [UGAUGAU]), and (iii) yeast species (3 RNAs, five bases [GAUGA]). A consensus sequence, GAUGA, is common to all RNAs, and this sequence can be expanded to PyGAUGA if the two plant U3 sequences are excluded. The conserved nature of the box C-related sequences argues that they may be functionally important and, if so, that the statistically infrequent GAUGA could encode that function. While the RNAs with this sequence are associated with the nucleolus, only a subset of the nucleolar snRNAs contain box C. The minimal box C element occurs in each of the mammalian RNAs precipitated but is present in only three of the precipitable yeast RNAs. Box C occurs in the interior of the U3 sequences, whereas it is close to the 5′ terminus in U14 and snR190 (24, 39, 42). Interestingly, box C and D (consensus PuUCUGA) are coconserved among the snRNAs in which they appear (39).

**Mutagenesis of box C and box D.** Results from earlier functional mapping of yeast U14 identified the box C region as one of several segments required for activity (12). Two

TABLE 1. Box C sequences of the nucleolar snRNAs[a]

| RNA | Organism | Sequence | Reference |
|---|---|---|---|
| U3 | Human | cau UGAUGA ucg | 24 |
| | Rat (U3B) | cau UGAUGA ucg | 28 |
| | *Xenopus laevis* | uau UGAUGA acg | 13 |
| | *X. borealis* | uau UGAUGA acg | 13 |
| | Tomato | agu aGAUGA ucg | 15 |
| | Broad bean | agu aGAUGA ucg | 16 |
| | *Bombyx mori* (U3C) | cuu cGAUGA ucg | 1 |
| | *Dictyostelium discoideum* (D2) | cau UGAUGA ccg | 41 |
| | *Schizosaccharomyces pombe* | aug cGAUGA ucu | 27 |
| | *Saccharomyces cerevisiae* | gcg cGAUGA ucu | 11 |
| U8 | Human | aca UGAUGA uug | 39 |
| | Rat | aaa UGAUGA uug | 29 |
| | Mouse (5.4S RNA) | aaa UGAUGA uug | 29 |
| U13 | Human | gcg UGAUGA uug | 39 |
| U14 | Rat | cug UGAUGA ugu | 38 |
| | Hamster | cug UGAUGA ugu | 38 |
| | Mouse (4.5S hybrid RNA) | cug UGAUGA ugg | 38 |
| | *S. cerevisiae* | cgg UGAUGA aag | 42 |
| snR190 | *S. cerevisiae* | ccc UGAUGA uaa | 42 |
| Consensus | | PyGAUGA | |

[a] The sequences of the box C-related elements in the nucleolar snRNAs are listed along with a suggested consensus sequence, PyGAUGA. Box C is one of four conserved sequence elements (boxes A through D) initially defined among U3 RNAs (41).

mutations in this region led to severe underaccumulation of U14 and a U14-minus lethal phenotype. One was a seven-base substitution that replaced the entire box C sequence. The second was an A→G change in box C (UGAUGA); Fig. 1). The box C and D regions have now been mapped to higher resolution by mutagenesis of the individual bases spanning these sequences and the flanking nucleotides. Mutations were evaluated in a test strain containing a galactose-inducible U14 allele. In this strain, control RNA was derived from a single chromosomal *GAL1::SNR128* gene (U14 was initially designated snR128 [42]), while mutant RNAs were transcribed from normal U14 transcription signals in a single-copy *CEN* vector. U14 depletion was mediated by glucose repression (19). Mutants were screened for (i) growth phenotype at 25, 32, and 37°C on glucose minimal medium, (ii) content of U14 RNA, and (iii) balance of 18S and 25S rRNAs.

The properties of 17 point mutants in the box C segment are summarized in Fig. 2 and 3. The results confirm that box C is essential for U14 accumulation. Two nucleotides appear to be especially important, specifically the first G and A residues (G8 and A9) in the UGAUGA repeat. Although the other sites have not been saturated, the mutations that were characterized had little or no effect on growth or RNA content. Any change at G8 results in a lethal or severely impaired growth phenotype. Mutations at A9 also affect growth substantially, yielding lethal or cold-sensitive pheno-

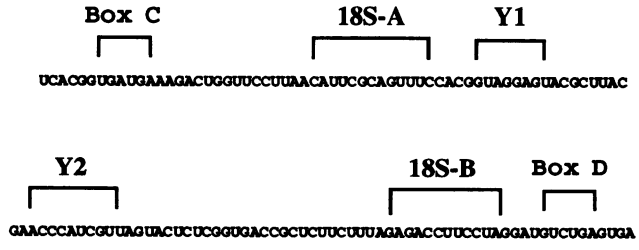



FIG. 1. Sequence of *S. cerevisiae* U14 snRNA. The highlighted segments include box C and box D, which are conserved among several nucleolar snRNAs; elements 18S-A and 18S-B, which are universal among U14 snRNAs; and yeast-specific elements Y1 and Y2. The 18S-A sequence and Y elements are required for yeast U14 function. Mutations in the box C and D regions and terminal stem have been shown to disrupt U14 accumulation (12, 18).

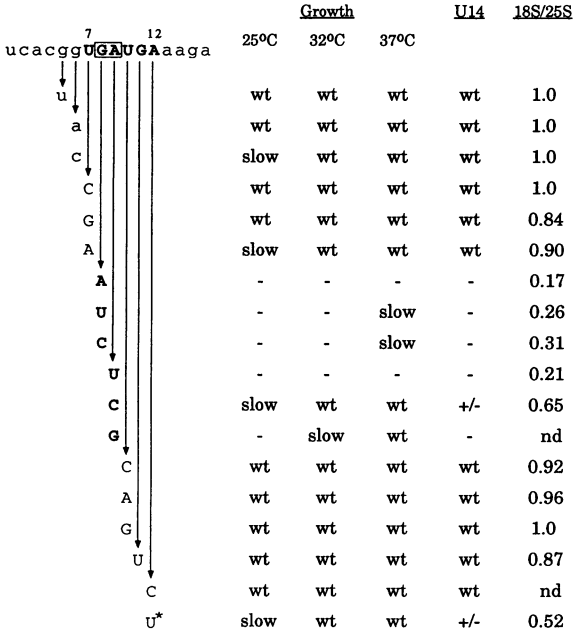



FIG. 2. Effects of box C mutations on growth and RNA accumulation. The sequence of the U14 box C region is shown, with the conserved PyGAUGA sequence in bold letters. Mutations causing U14 loss and impaired growth are highlighted. Slow growth means growth notably slower than wild-type (wt) growth on solid medium. For U14, - indicates trace or undetectable amounts; growth is supported by levels as low as 5% of the wild-type level (12, 42). The A12→U mutant U* also contains A3→C. Northern hybridization data for several mutants are shown in Fig. 3.

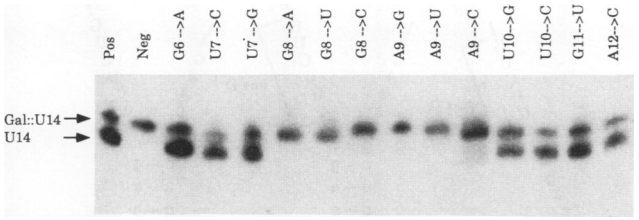
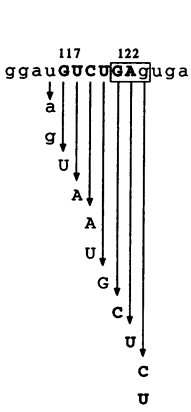



FIG. 3. Effects of box C mutations on U14 content. Results of Northern analysis are shown for several box C mutants listed in Fig. 2 in cells cultured at 32°C. Two U14 bands are evident; these bands correspond to (i) RNA from the chromosomal *GAL1* allele (upper, with five additional 5′ nucleotides) and (ii) mutant RNAs transcribed from wild-type U14 signals in a single-copy *CEN* plasmid (lower). Samples are identified above the lanes. Pos and Neg denote positive and negative control samples from transformants with and without plasmid-encoded wild-type U14 RNA. Slow growth in U14-deficient cells may be due to undetectable U14 RNA (19). Mutations at G8 and A9 block U14 accumulation.

types. Substitutions at either site lead to loss of U14, although apparently to lesser extents for C and G at A9. As anticipated, U14 depletion leads to loss of 18S RNA, reflected in this study as an imbalance in the ratio of 18S to 25S RNAs (12, 19).

Bases in the box D sequence were also found to be important for U14 accumulation (Fig. 4). Analysis of this segment showed that several point mutations in box D and the adjoining 3′ nucleotide led to loss or diminution of U14. U14 loss was most dramatic with changes in the final three bases of box D and the following G, i.e., GUCUGAg. Taken together, the mutagenesis results show that both box C and box D sequences are necessary for U14 accumulation. The results also identify critical bases in each element.

**Accumulation requires terminal base pairing.** The occurrence of a terminal stem in U14 is suggested by both phylogenetic sequence data and results from expression and activity analyses with hybrid yeast-mouse U14 RNAs (18, 21). Although the sequences vary, the U14 RNAs from three mammalian sources and *S. cerevisiae* all have the potential to form four or five base pairs at the ends. Analysis of several yeast-mouse hybrid RNAs in *S. cerevisiae* established that a base pair at wild-type positions A3 and U124 was necessary for accumulation (18).




| | Growth | | | U14 |
|---|---|---|---|---|
| | 25°C | 32°C | 37°C | |
| a | wt | wt | wt | wt |
| g | wt | wt | wt | wt |
| U | wt | wt | wt | wt |
| A | wt | wt | wt | wt |
| A | slow | slow | wt | +/- |
| U | wt | wt | wt | +/- |
| G | - | wt | wt | - |
| C | - | - | slow | - |
| U | - | - | slow | - |
| C | - | - | - | - |
| U | wt | wt | wt | wt |

FIG. 4. Effects of point mutations in the box D region. The sequence of the region is shown, with the conserved box D nucleotides GUCUGA in bold letters. Growth and U14 RNA accumulation results are reported as for Fig. 2. wt, wild type.

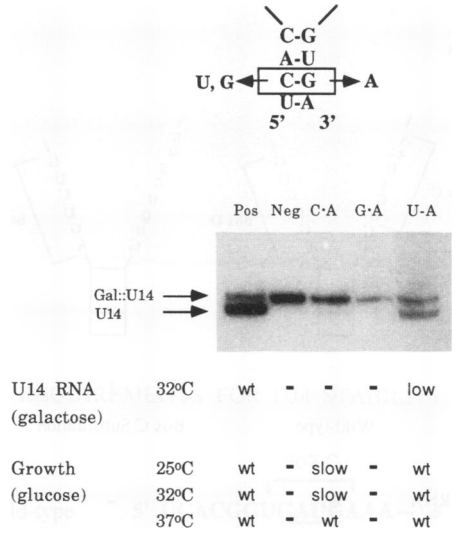



FIG. 5. Effects of terminal base changes on U14 accumulation. Single mutations were made at nucleotide C2 and G125, as indicated. Putative base pairing is disrupted by mutation G125→A (lane C · A) and by double mutations C2→G and G125→A (lane G · A). U14 accumulation and function were recovered by double mutations C2→U and G125→A, which restore base pairing (lane U-A). Growth and U14 assays were as described in Materials and Methods and the legend to Fig. 3. Pos and Neg are as defined in the legend to Fig. 3. wt, wild type.

The existence of terminal base pairing was examined further by mutation of C2 and G125, which define the second of four potential base pairs in the stem. Both positions were mutated to yield two different noncomplementary base pairs and a base pair of different identity (U · A; Fig. 5). RNA failed to accumulate with loss of complementarity for C · A and G · A pairs but did accumulate with the novel A · U base pair (slow growth with mispaired C · A RNA may reflect residual active RNA). These results coupled with the earlier mutational data for the third position constitute strong evidence for pairing at both the second and third positions and, thus, formation of the proposed terminal stem. The results also support the view that metabolic instability is the basis of the accumulation defects observed in the stem region. While a stem of four base pairs is suggested, mutagenic data for G123 in the fourth position indicate that stem function in some cases may be satisfied by only three base pairs. As noted in Fig. 4, a G123→U change is neutral but a G123→C change is lethal, which suggests an important structural context effect.

**Boxes C and D appear to provide stability function.** Loss of U14 with the critical box C and box D mutations could reflect either a block in expression or loss of metabolic stability. In an initial attempt to distinguish between these possibilities, we determined that U14 also failed to accumulate when lethal box C alleles G8→U and G8→C were fused to a plasmid-encoded *GAL1* promoter, whereas a neutral variant (U9→G) accumulated normally (result not shown). If expression from the *GAL1* promoter is resistant to these base changes, this observation argues that U14 loss is due to enhanced turnover. This interpretation seems reasonable in view of the extensive record of successful expressions of heterologous DNA achieved with the *GAL1* promoter (e.g., 23, 33).

**Box C or D function is provided by extension of the terminal**

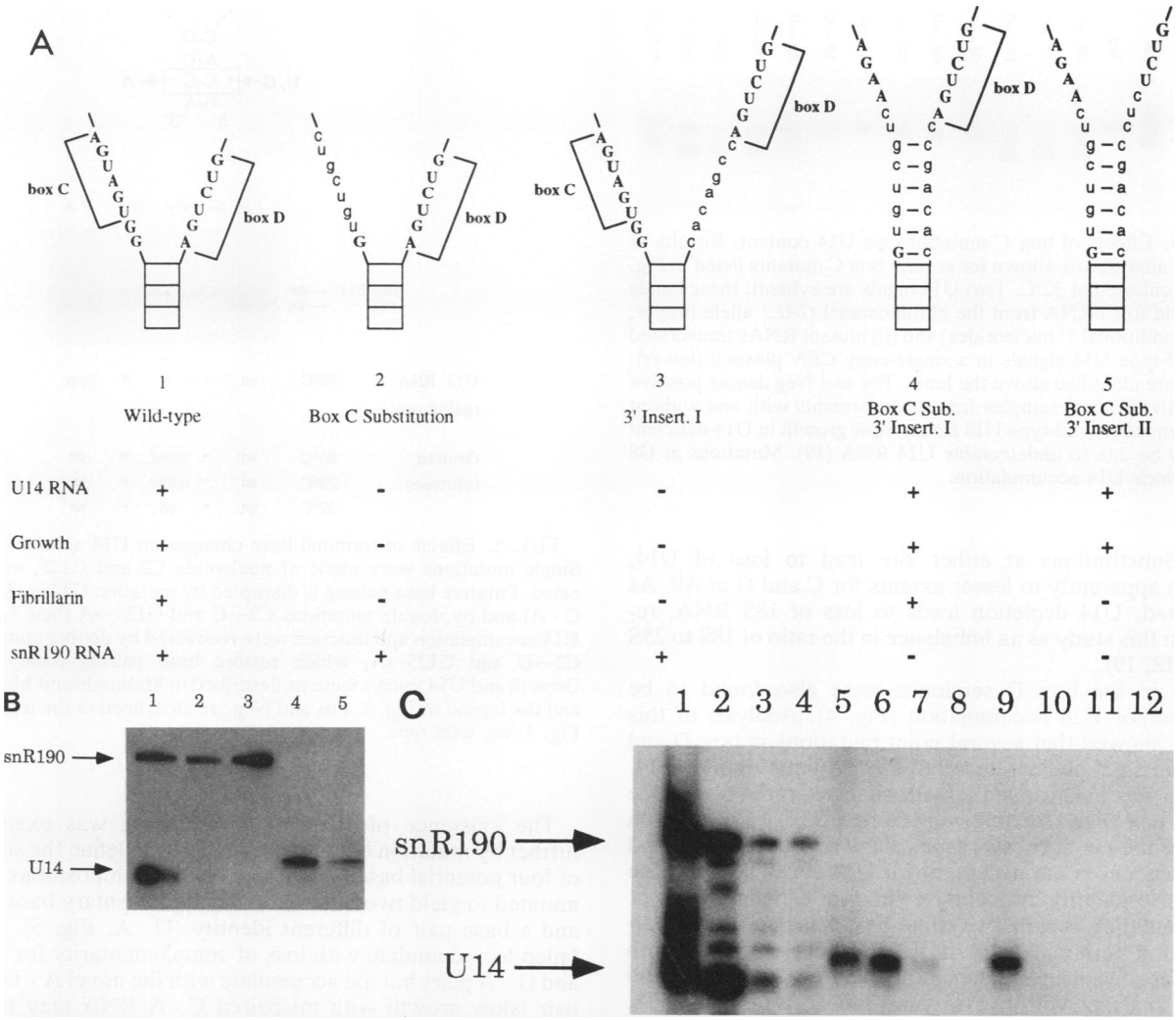



FIG. 6. Structures and properties of extended stem mutants. (A) Hypothetical structures of the stem regions for each RNA variant (1 through 5) and the effects of the various mutations. Properties scored included U14 accumulation, growth, association with fibrillarin, and snR190 level. Fibrillarin association was assayed by immunoprecipitation with a fibrillarin monoclonal antibody. + for the U14 level corresponds to 20 to 30% of the normal levels. ++, +, and − correspond to wild-type-like growth, slow growth, and no growth, respectively. The results indicate that lethal box C and box D mutations are complemented by extension of the 5′, 3′ stem. (B) Accumulation of box C and box D mutants with extended terminal stems. U14 and snR190 levels are shown for YS153 transformants containing plasmids that encode wild-type U14 (lane 1), the lethal box C substitution mutant, i.e., RNA 2 in panel A (lane 2), RNA 3 (lane 3), RNA 4 (lane 4), and RNA 5 (lane 5). The experimental U14 sequences are encoded in pJZ45 DNA, which also encodes snR190. RNA was prepared from cells in YPD medium; samples 2 and 3 (U14-minus) were from cells grown first in SG medium with histidine and uracil. (C) Abilities of the double substitution-insertion mutants to associate with fibrillarin, analyzed with fibrillarin antibodies. Nuclear extracts were treated with antibody, and the immunoprecipitated snRNAs were analyzed by Northern hybridization. Results are presented for three U14 variants probed with three antibody preparations. The extracts are as follows: wild type (lanes 1 to 4), RNA 4 (lanes 5 to 8), and RNA 5 (lanes 9 to 12). The lanes in each sample set correspond to (i) untreated extract, (ii) antibody 72B9, (iii) antibody 7, and (iv) antibody 8; antibodies 7 and 8 were human autoimmune sera specific for nucleolar antigens. Some degradation is apparent in the wild-type snRNA sample.

**stem.** If box C or box D is required for stability, we reasoned that it might be possible to replace these elements with another type of stabilizing domain. We determined that both elements are dispensable if the adjoining terminal stem is extended (Fig. 6). The four constructs used for the analysis were (i) a lethal seven-base substitution mutant described previously (12) in which the final four bases of box C have been replaced, (ii) a seven-base insertion at the 3′ end of box D that is complementary to the substituted sequence at box C (insert I), (iii) a mutant containing both the box C substitution sequence and 3′ complementary insertion I sequence, and (iv) a variant containing the box C substitution and a nine-base insertion-substitution sequence at the 3′ end of box D. This last mutant includes the complementary seven-base insert I and alters the two last essential bases of box D (insert II). The two combination substitution-insertion mutants have the potential to form a terminal stem with 10 contiguous base pairs.

The mutant RNAs were assayed for accumulation and for the ability to support growth and associate with fibrillarin (Fig. 6). Fibrillarin association was evaluated with fibrillarin antibody. The content of snR190 was also estimated (U14 and snR190 are transcribed from DNA sequences separated by only 67 bases [42]). The results show that U14 RNAs with

the box C substitution or 3′ distal insertion did not accumulate and thus could not support growth or associate with fibrillarin (Fig. 6B). However, RNAs containing the box C substitution sequence plus either of the complementary 3′ insertions did accumulate and were functional. These variants accumulated to levels about 20 to 30% of the wild-type level. Growth in rich medium (YPD) was normal with mutant RNA containing box D but slower at 25 and 37°C for cells containing the defective box D variant; growth was slow in minimal medium (SD) for the box D-containing variant. Levels of snR190 were not affected by either 5′ or 3′ mutation alone, but curiously, snR190 was not detected in either double mutant (Fig. 6B). The basis of snR190 loss is not known and could result from a differential effect on synthesis or stability of cotranscripts or from assembly of RNPs containing snR190.

Immunoprecipitation analysis of the extracts showed that U14 association with fibrillarin does not require box C but may require box D, as little or no interaction was observed in its absence (Fig. 6C). These results also suggest that fibrillarin binding may not be the basis of the stability provided by box C. Precipitation of the fibrillarin-positive RNP complexes containing the extended stem variant exhibited the same sensitivity to salt as did precipitation of wild-type U14 RNP over a range of salt concentrations from 150 to 405 mM (NaCl plus KCl; data not shown). This similarity could reflect similar binding of fibrillarin with mutant and normal U14 RNPs or, less interesting, a basic feature of the antibody-fibrillarin reaction.

**Intragenic suppression of a lethal box C mutation.** A search for suppressor mutations that correct lethal box C and box D mutations is in progress, with the aim of identifying gene products that might interact with these elements. Initial efforts have been concerned with suppressors to the lethal A9→U mutation in box C (Table 1). Fifty-six Gal$^{l}$ isolates were screened by PCR sequencing of plasmid DNA to distinguish between intragenic and extragenic mutations. These data identified (i) 29 mutants with a base change at a second site in U14 DNA, (ii) one revertant at the original A9 position, and (iii) 26 isolates with extragenic mutations; the extragenic mutants are being characterized and will be described elsewhere. The 29 novel intragenic revertants were identical, with a G→A change at position 6 (Fig. 7A). Transformation of fresh, unmutagenized test cells with purified plasmid verified that the G6→A change was the basis of suppression.

On solid glucose medium, the transformants exhibited wild-type-like growth at 32 and 37°C but slow growth at 25°C. Growth in liquid glucose medium at 32°C was indistinguishable from that of wild-type cells. Analysis of U14 levels showed that the double mutant accumulated to about 10 to 20% of the wild-type RNA level (Fig. 7B). The mutant RNA precipitated with fibrillarin antibody, with the same salt and pH dependence as for wild-type U14 RNA (data not shown). This last finding suggests that the affinity for fibrillarin may not be altered by the double mutation.

Interestingly, inspection of the sequence for the double mutant shows that the second mutation creates a new box C-like sequence, PyGAUG, two bases upstream from the original box C element (Fig. 7A). The fact that the original box C mutation can be rescued by creation of a new box C element confirms the importance of the sequence revealed in the mapping analysis. The close proximity of the old and new box C elements is consistent with the possibility that box C serves as a recognition site for a protein, perhaps one with stabilizing function.

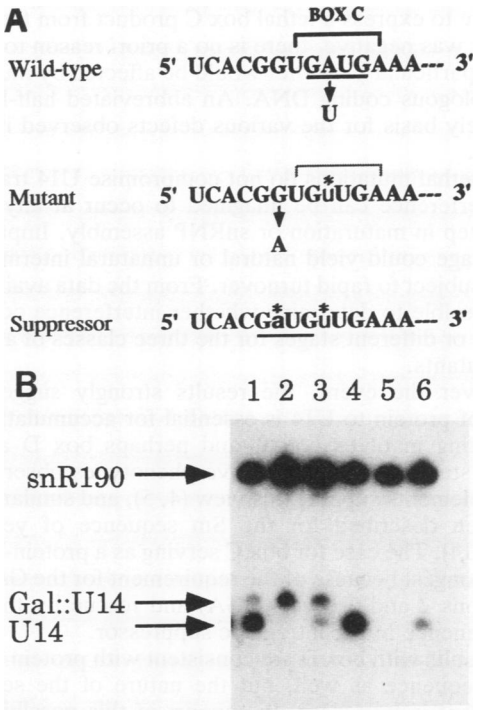



FIG. 7. Effect of an intragenic suppressor mutation on U14 accumulation and fibrillarin association. (A) Sequence of the box C suppressor. A lethal mutation at position A9 in the box C region is suppressed by a G→A mutation at position G6. The suppressor RNA contains a new box C-like PyGAUG sequence. The common GAUG sequences are underlined. (B) Patterns of U14 and snR190 production, determined for cells containing the box C mutant A9→U and intragenic suppressing mutation G6→A. Total RNA from test transformants grown at 32°C in galactose (lanes 1 to 3) or glucose (lanes 4 to 6) was fractionated in an 8% polyacrylamide gel and analyzed by Northern hybridization. RNAs were detected with $^{32}$P-labeled oligonucleotides. The samples were plasmid-encoded wild-type U14 (lanes 1 and 4), box C mutant A9→U (lanes 2 and 5), and intragenic suppressor mutant A9→U/G6→A (lanes 3 and 6). Probing for snR190 was done to normalize for possible differences in plasmid copy number. Cells harboring the A9→U variant were grown first in galactose medium; 1% of this culture was used to inoculate glucose-containing medium, and RNA was prepared after 24 h.

## DISCUSSION

The results presented above demonstrate that box C, box D, and the terminal stem are required for accumulation. The basis of U14 loss is not clear, but metabolic instability seems more likely than interference with transcription. This view is supported by the finding that the accumulation function of boxes C and D can be provided by extension of the terminal stem. Considered alone, the lethal effects of box C and even those of box D can be imagined to reflect defects in transcription, in particular binding of transcription factors. Consistent with this possibility is the demonstrated importance of factor binding to promoter proximal and distal elements in expression of the U1 and U2 snRNAs (reviewed in references 6 and 26). It seems unlikely, however, that a requirement for transcription factor binding at either or both sites could be satisfied by incorporation of complementary sequences at both ends of the coding sequence. The new segments do not contain the essential features of the box elements, and the box C substitution alone is lethal. While

the failure to express a lethal box C product from the *GAL1* promoter was negative, there is no a priori reason to expect that this particular promoter would be affected by mutations in heterologous coding DNA. An abbreviated half-life is a more likely basis for the various defects observed in accumulation.

If the lethal mutations do not compromise U14 transcription, interference can be imagined to occur at any downstream step in maturation or snRNP assembly. Impairment at any stage could yield natural or unnatural intermediates that are subject to rapid turnover. From the data available, it is not possible to determine whether interference occurs at the same or different stages for the three classes of accumulation mutants.

Whatever the event, the results strongly suggest that binding of protein to U14 is essential for accumulation and that binding involves box C and perhaps box D and the terminal stem. The cold-sensitive phenotypes observed for the box elements support this view (4, 5), and similar effects have been described for the Sm sequence of yeast U5 snRNA (14). The case for box C serving as a protein-binding site is strongest because of the requirement for the GA bases at positions 2 and 3 (PyGAUGA) and rescue by the PyGAUG sequence in the intragenic suppressor.

The results with box D are consistent with protein binding to this sequence as well, but the nature of the sequence requirement is less clear. In support of this possibility are preliminary data for one extragenic suppressor of box C that also suppresses a lethal box D mutation (35b). This situation suggests that U14 stability might be provided by a single protein binding to one or both sites. The UGA motif is common to both box elements and is attractive as a candidate binding signal. Interestingly, fibrillarin binds to the multibase box C substitution mutant with an intact box D but poorly or not at all to the variant with the incomplete box D (Fig. 6). We note that these variants contain a GA doublet in the former box C region opposite box D; this is not the situation for the initial box C mutant. Although we do not know whether GA alone can provide box C function, fibrillarin binding to the extended stem mutant could, in theory, depend on recognition of a box C-like GA opposite box D.

The stabilizing function of the terminal stem might reflect a role in protein binding, but its role could simply be to protect against attack by exonuclease. Similarly, the ability of the extended stem to complement the lethal box C and D mutations could reflect increased resistance of an unstable U14 to nucleolytic attack. At the same time, the extension may allow otherwise unfavored protein binding to occur in the regions encoding the essential box elements. The box C element occurs near the 5′ terminus in both yeast and mammalian U14 RNAs. Secondary folding projections for both RNA types show box C to be adjacent to similar 5′, 3′ terminal stem domains, with the box D sequences in close proximity to box C (2a, 7, 34, 39, 44). As noted above, box D is situated near the 3′ end in the nucleolar snRNAs in which it occurs, whereas box C is not always at the 5′ end. Box C is in the interior of the various U3 RNAs and human U8 and U13. Folding models predict that box C and box D occur in single-stranded regions of U3, U8, and U13 and that these elements are in close proximity in U8 and U13 (13, 24, 39). Boxes C and D are further apart in the models of U3 but could be close in the native RNP. Interestingly, a UGA triplet reminiscent of box D occurs opposite box C in the models for U3, and the nearby box B includes a box C-like GA.

It will be important to determine whether the effects observed in this study for the box elements of yeast U14 apply to the box sequences of other snRNAs as well, from *S. cerevisiae* and other sources. Initial results from mutagenesis of box C of the snR190 species of *S. cerevisiae* have shown that a change of A→U at the third position in box C (PyGUUGA) also blocks accumulation of this RNA (35a). The box C sequence of human U3 has been shown to be required for formation of fibrillarin-containing RNP complexes in vitro, as measured with fibrillarin antibody. Mutations at several positions in box C compromised fibrillarin binding, but the effect was not limited to the GA bases shown to be important for box C of U14 (3). Binding with fibrillarin did not require box D. It will be interesting to determine whether accumulation of mammalian or yeast U3 also depends on box C. To gain insight into this matter, we are currently conducting a study parallel to this one for yeast U3.

While the U14 and U3 mutational results imply that similar effects can be anticipated with other nucleolar snRNAs, the potential for differences also exists, especially for box C, which is not positionally conserved. The compilation of box C sequences in Table 1 might include elements that specify different functions, more than one function, and even no function at all, despite the statistical rarity of the conserved sequence. The possibility of overlapping or different functions must be kept in mind, especially when a vital PyGA UGA sequence is close to the 5′ end. We note too that lethal phenotypes should not be expected for every case of snRNA loss, as several species (including snR190) are known to be dispensable (25, 42).

## ACKNOWLEDGMENTS

The contributions of the first three authors were approximately equal. We are grateful to K. M. Pollard and E. M. Tan for supplying fibrillarin antibody and thank R. Lempicki for advice on the immunological assays. We also thank our colleagues R. Lempicki, H. V. Li, and A. Balakin for helpful discussions and I. Mattaj for reviewing the manuscript.

This work was supported by NIH grant GM19351.

## REFERENCES

1. **Adams, D. S., R. J. Herrera, R. Luhrmann, and P. M. Lizardi.** 1985. Isolation and partial characterization of U1-U6 small RNAs from *Bombyx mori*. Biochemistry **24**:117–125.
2. **Aris, J. P., and G. Blobel.** 1988. Identification and characterization of a yeast nucleolar protein that is similar to a rat liver nucleolar protein. J. Cell Biol. **107**:17–31.

2a. **Balakin, A., R. Lempicki, and G. Huang.** Unpublished data.

3. **Baserga, S. J., W. Yang, and J. A. Steitz.** 1991. An intact box C sequence in the U3 snRNA is required for binding of fibrillarin, the protein common to the major family of nucleolar snRNPs. EMBO J. **10**:2645–2651.
4. **Bayliss, F. T., and J. L. Ingraham.** 1974. Mutation in *Saccharomyces cerevisiae* conferring streptomycin and cold sensitivity by affecting ribosome formation and function. J. Bacteriol. **118**:319–328.
5. **Carter, C. J., and M. Cannon.** 1980. Maturation of ribosomal precursor RNA in *Saccharomyces cerevisiae*. J. Mol. Biol. **143**:179–199.
6. **Dahlberg, J. E., E. Lund, and M. Birnstiel (ed.).** 1988. Structure and function of major and minor small nuclear ribonucleoprotein particles, p. 38–70. Springer-Verlag, Berlin.
7. **Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**:9373–9377.
8. **Guarente, L., and T. L. Mason.** 1983. Heme regulates transcription of the *CYC1* gene of *Saccharomyces cerevisiae* via an upstream activation site. Cell **32**:1279–1286.

9. **Hill, D. E., A. R. Oliphant, and K. Struhl.** 1987. Mutagenesis with degenerate oligonucleotides: an efficient method for saturating a defined DNA region with base pair substitutions. Methods Enzymol. **155**:558–568.
10. **Hughes, J. M. X., and M. Ares, Jr.** 1991. Depletion of U3 small nucleolar RNA inhibits cleavage in the 5′ external transcribed spacer of yeast pre-ribosomal RNA and impairs formation of 18S ribosomal RNA. EMBO J. **10**:4231–4239.
11. **Hughes, J. M. X., D. A. M. Konings, and G. Cesareni.** 1987. The yeast homologue of U3 snRNA. EMBO J. **6**:2145–2155.
12. **Jarmolowski, A., J. Zagorski, H. V. Li, and M. J. Fournier.** 1990. Identification of essential elements in U14 RNA of *Saccharomyces cerevisiae*. EMBO J. **9**:4503–4509.
13. **Jeppesen, C., B. Stebbins-Boaz, and S. Gerbi.** 1988. Nucleotide sequence determination and secondary structure of *Xenopus* U3 snRNA. Nucleic Acids Res. **16**:2127–2148.
14. **Jones, M. H., and C. Guthrie.** 1990. Unexpected flexibility in an evolutionarily conserved protein-RNA interaction: genetic analysis of the Sm binding site. EMBO J. **9**:2555–2561.
15. **Kiss, T., and F. Solymosy.** 1990. Molecular analysis of a U3 RNA gene locus in tomato: transcription signals, the coding region, expression in transgenic tobacco plants and tandemly repeated pseudogenes. Nucleic Acids Res. **18**:1941–1949.
16. **Kiss, T., M. Toth, and F. Solymosy.** 1985. Plant small nuclear RNAs—nucleolar U3 snRNA is present in plants: partial characterization. Eur. J. Biochem. **152**:259–266.
17. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Acad. Sci. USA **82**:488–492.
18. **Li, H. V., and M. J. Fournier.** 1992. U14 function in *Saccharomyces cerevisiae* can be provided by large deletion variants of yeast U14 and hybrid mouse-yeast U14 RNAs. EMBO J. **11**:683–689.
19. **Li, H. V., J. Zagorski, and M. J. Fournier.** 1990. Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **10**:1145–1152.
20. **Lischwe, M. A., R. L. Ochs, R. Reddy, R. G. Cook, L. C. Yeoman, E. M. Tan, M. Reichlin, and H. Busch.** 1985. Purification and partial characterization of a nucleolar scleroderma antigen (Mr34,000; pI8.5) rich in N6, N6-dimethylarginine. J. Biol. Chem. **260**:14304–14310.
21. **Liu, J., and E. S. Maxwell.** 1990. Mouse U14 snRNA is encoded in an intron of the mouse cognate hsc70 heat shock gene. Nucleic Acids Res. **18**:6565–6571.
22. **Lohr, D.** 1988. Isolation of yeast nuclei and chromatin for studies of transcription-related processes, p. 125–145. *In* I. Campbell, and J. H. Duffus (ed.), Yeast: a practical approach. IRL Press, Washington, D.C.
23. **Miyajima, A., N. Nakayama, I. Miyajima, N. Arai, H. Okayama, and K. Arai.** 1984. Analysis of full-length cDNA clones carrying *GAL1* of *Saccharomyces cerevisiae*: a model system for cDNA expression. Nucleic Acids Res. **12**:6397–6414.
24. **Parker, K. A., and J. A. Steitz.** 1987. Structural analyses of the human U3 ribonucleoprotein particle reveal a conserved sequence available for base pairing with pre-rRNA. Mol. Cell. Biol. **7**:2899–2913.
25. **Parker, R., E. O. Shuster, T. Simmons, P. G. Siliciano, and C. Guthrie.** 1988. Genetic analysis of small nuclear RNAs in *Saccharomyces cerevisiae*: viable sextuple mutant. Mol. Cell. Biol. **8**:3150–3159.
26. **Parry, H. D., D. Scherly, and I. W. Mattaj.** 1989. Snurpogenesis: the transcription and assembly of U snRNP components. Trends Biochem Sci. **14**:15–20.
27. **Porter, G. L., P. J. Brennwald, K. A. Holm, and J. A. Wise.** 1988. The sequence of U3 from *Schizosaccharomyces pombe* suggests structural divergence of this snRNA between metazoans and unicellular eukaryotes. Nucleic Acids Res. **16**:10131–10152.
28. **Reddy, R., D. Henning, and H. Busch.** 1979. Nucleotide sequence of nucleolar U3B RNA. J. Biol. Chem. **254**:11097–11105.
29. **Reddy, R., D. Henning, and H. Busch.** 1985. Primary and secondary structure of U8 small nuclear RNA. J. Biol. Chem. **260**:10930–10935.
30. **Reimer, G., K. M. Pollard, C. A. Penning, R. L. Ochs, M. A. Lischwe, H. Busch, and E. M. Tan.** 1987. Monoclonal autoantibodyfroma (New Zealand black × New Zealand white) F1 mouse and some human scleroderma sera target an Mr 34,000 nucleolar protein of the U3 RNP particle. Arthritis Rheum. **30**:793–800.
31. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Methods in yeast genetics: a laboratory course manual, p. 181–182. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
32. **Schimmang, T., D. Tollervey, H. Kern, R. Frank, and E. C. Hurt.** 1989. A yeast nucleolar protein related to mammalian fibrillarin is associated with small nuclear RNA and is essential for viability. EMBO J. **8**:4015–4024.
33. **Schneider, J. C., and L. Guarente.** 1991. Vectors for expression of cloned genes in yeast: regulation, overproduction, and underproduction. Methods Enzymol. **194**:379–385.
34. **Shanab, G. M., and E. S. Maxwell.** 1991. Proposed secondary structure of eukaryotic U14 snRNA. Nucleic Acids Res. **19**:4891–4894.
35. **Steitz, J. A.** 1989. Immunoprecipitation of ribonucleoproteins using autoantibodies. Methods Enzymol. **180**:468–481.

35a. **Struck, J.** Unpublished data.

35b. **Struck, J., and R. Lempicki.** Unpublished data.

36. **Tollervey, D.** 1987. A yeast small nuclear RNA is required for normal processing of pre-ribosomal RNA. EMBO J. **6**:4169–4175.
37. **Tollervey, D., and E. C. Hurt.** 1990. The role of small nucleolar ribonucleoproteins in ribosome synthesis. Mol. Biol. Rep. **14**:103–106.
38. **Trinh-Rohlik, Q., and E. S. Maxwell.** 1988. Homologous genes for mouse 4.5S hybRNA are found in all eukaryotes and their low molecular weight RNA transcripts intermolecularly hybridize with eukaryotic 18S ribosomal RNAs. Nucleic Acids Res. **16**:6041–6056.
39. **Tyc, K., and J. A. Steitz.** 1989. U3, U8 and U13 comprise a new class of mammalian snRNPs localized in the cell nucleolus. EMBO J. **8**:3113–3119.
40. **Warner, J. R.** 1990. The nucleolus and ribosome formation. Curr. Opin. Cell Biol. **2**:521–527.
41. **Wise, J. A., and A. M. Weiner.** 1980. *Dictyostelium* small nuclear RNA D2 is homologous to rat nucleolar RNA U3 and is encoded by a dispersed multigene family. Cell **22**:109–118.
42. **Zagorski, J., D. Tollervey, and M. J. Fournier.** 1988. Characterization of an *SNR* gene locus in *Saccharomyces cerevisiae* that specifies both dispensible and essential small nuclear RNAs. Mol. Cell. Biol. **8**:3282–3290.
43. **Zubenko, G. S., A. P. Mitchell, and E. W. Jones.** 1980. Mapping of the proteinase b structural gene *PRB1* in *Saccharomyces cerevisiae* and identification of nonsense alleles within the locus. Genetics **96**:137–146.
44. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamic and auxiliary information. Nucleic Acids Res. **9**:133–148.