MOLECULAR AND CELLULAR BIOLOGY, Dec. 1996, p. 6993–7003 Vol. 16, No. 12
0270-7306/96/$04.00+0

Repression of Gene Expression by an Exogenous Sequence Element Acting in Concert with a Heterogeneous Nuclear Ribonucleoprotein-Like Protein, Nrd1, and the Putative Helicase Sen1

ERIC J. STEINMETZ AND DAVID A. BROW*

Department of Biomolecular Chemistry, University of Wisconsin Medical School, Madison, Wisconsin 53706-1532

Received 3 July 1996/Returned for modification 27 August 1996/Accepted 11 September 1996

We have fortuitously identified a nucleotide sequence that decreases expression of a reporter gene in the yeast *Saccharomyces cerevisiae* 20-fold when inserted into an intron. The primary effect of the insertion is a decrease in pre-mRNA abundance accompanied by the appearance of 3′-truncated transcripts, consistent with premature transcriptional termination and/or pre-mRNA degradation. Point mutations in the *cis* element relieve the negative effect, demonstrating its sequence specificity. A novel yeast protein, named Nrd1, and a previously identified putative helicase, Sen1, help mediate the negative effect of the *cis* element. Sen1 is an essential nuclear protein that has been implicated in a variety of nuclear functions. Nrd1 has hallmarks of a heterogeneous nuclear ribonucleoprotein, including an RNA recognition motif, a region rich in RE and RS dipeptides, and a proline- and glutamine-rich domain. An N-terminal domain of Nrd1 may facilitate direct interaction with RNA polymerase II. Disruption of the *NRD1* gene is lethal, yet C-terminal truncations that delete the RNA recognition motif and abrogate the negative effect of the *cis* element nevertheless support cell growth. Thus, expression of a gene containing the *cis* element could be regulated through modulation of the activity of Nrd1. The recent identification of Nrd1-related proteins in mammalian cells suggests that this potential regulatory pathway is widespread among eukaryotes.

A critical phase in the expression of a eukaryotic protein-coding gene occurs between the onset of synthesis of its pre-mRNA transcript in the nucleus and the transport of the mature mRNA product to the cytoplasm. During this period, the 5′ end of the nascent transcript acquires a 7-methylguanosine cap, the 3′ end of the transcript is defined by endonucleolytic cleavage and addition of a polyadenylate tail, and introns are recognized and accurately excised (13, 44). These processes are mediated by a remarkably complex machinery that is largely conserved from yeasts to vertebrates. Only after these processing steps are complete does mRNA normally become a substrate for transport through the nuclear pores and eventual translation (35).

Despite the increasingly detailed picture of the mechanisms of the transcription and pre-mRNA processing reactions, many aspects of the nuclear metabolism of pre-mRNA and mRNA remain obscure. Mechanisms coordinating the various processing reactions, and the transport of mRNA from the sites of synthesis and processing to the nuclear pores, are only beginning to be elucidated (18, 56). Proteins that associate with pre-mRNA as it is synthesized, known collectively as heterogeneous nuclear ribonucleoproteins (hnRNPs) (16), presumably play a central role in directing these processes. For example, manipulation of the ratios of the hnRNP A1 protein and the splicing factor ASF/SF2 in vitro and in vivo can influence 5′ splice site choice (9, 40, 61, 67). Some hnRNPs, including A1, are thought to function in nucleocytoplasmic transport (41, 48). hnRNP L can stimulate the maturation and cytoplasmic accumulation of intronless mRNAs containing appropriate sequence elements (38). An additional role for hnRNPs in determining the stability of nuclear pre-mRNA seems plausible, although mechanisms for the nuclear turnover of pre-mRNA are not well defined. The lack of genetic approaches to the study of hnRNP function has hampered progress in understanding the roles of these proteins in vivo, but the recent identification of hnRNP proteins in *Saccharomyces cerevisiae* (3, 4, 39, 66) holds promise.

We have identified a sequence element that directs underexpression of a pre-mRNA, and consequently its mRNA and protein products, when inserted into an intron of a yeast reporter gene. The element is an artificial construct about 140 nucleotides (nt) long and includes antisense U6 spliceosomal RNA. Response to the *cis* element occurs after transcription initiation but before pre-mRNA splicing and results in the generation of truncated pre-mRNA fragments containing partial intron sequences, consistent with premature transcription termination and/or pre-mRNA degradation. A 20-fold decrease in the accumulation of pre-mRNA and mRNA is observed. Spontaneous point mutations arising within a 14-nt region of the antisense U6 sequence relieve the pre-mRNA accumulation defect but do not act by disrupting complementarity to U6 RNA. Rather, they may define a recognition site for a binding factor.

Mutations in a novel gene, which we have named *NRD1* (for nuclear pre-mRNA down-regulation), and in the gene encoding Sen1 (15), a putative ATP-dependent helicase, greatly diminish the *cis*-element-directed effect. The Nrd1 protein has several hnRNP-like features, including a single consensus RNA recognition motif (RRM), a short arginine-, serine-, and glutamate-rich segment similar to the regions rich in RE and RS dipeptides (RE/RS domains) found in many metazoan splicing factors, and a proline- and glutamine-rich C-terminal domain (P+Q domain) similar to domains found in several yeast hnRNPs (66). Disruption of *NRD1* is lethal; however, analysis of several mutant alleles indicates that the RRM, RE/RS, and P+Q domains are dispensable, and an N-terminal domain which may facilitate interactions with RNA polymer-

* Corresponding author. Mailing address: Department of Biomolecular Chemistry, University of Wisconsin Medical School, 1300 University Ave., Madison, WI 53706-1532. Phone: (608) 262-1475. Fax: (608) 262-5253. Electronic mail address: dabrow@facstaff.wisc.edu.

ase II (68) is sufficient for viability. We propose that Nrd1 binds to *cis*-element-containing pre-mRNA as it is synthesized and that this is a key step in specifying the pre-mRNA for down-regulation. Modulation of Nrd1's activity may thus provide a mechanism for regulation of gene expression at the level of pre-mRNA accumulation.

MATERIALS AND METHODS

Yeast strains and methods. Yeast procedures were performed according to standard methods (25). Copper plates were made and assays were performed as described previously (36, 37). *S. cerevisiae* strain 46α (*MAT*α *cup1*Δ *ura3 his3 trp1 lys2 ade2 leu2*) was obtained from C. Lesser. An isogenic *MAT***a** derivative of 46α was generated by introducing the HO endonuclease gene on a *URA3*-marked centromere plasmid (YCp50-HO) to induce mating type switching (25). Transformants were then streaked on 5-fluoroorotic acid (5-FOA) plates to select for loss of the *URA3*-marked plasmid, and individual colonies were tested for the ability to mate with 46α.

Plasmids were maintained in *Escherichia coli* DH5α and were transformed into *S. cerevisiae* by the lithium acetate procedure (55). Plasmids were recovered from yeast cells by the method of Hoffman and Winston (28). *NRD1* and *NRD2* were cloned from a yeast DNA library containing fragments from a partial *Sau*3A digest inserted in the *Bam*HI site of YCp50 (52). pRS316NRD1 has the *Hin*dIII fragment (see Fig. 6) inserted into the *Hin*dIII site of pRS316 (57), which harbors the *URA3* gene.

Mutant alleles of *NRD1* and *NRD2* were recovered on plasmids by gap repair (53). The one-step gene disruption procedure (53) was used to create the heterozygous Nrd1 disruption strain EJS101. The disruption of *NRD1* was created by inserting a 1.76-kb *Bam*HI fragment containing the *HIS3* gene between the *Bcl*I and *Bgl*II sites in the *NRD1* gene of pRS316NRD1, yielding pRS316nrd1Δ::HIS3 (see Fig. 6). An *Eco*RV-*Sal*I fragment containing the disruption construct was isolated and used to transform diploid strain EJS100. EJS100 was created by mating strains PJ43-2b (*MAT*α *his3-11,15 leu2-3,112 trp1-1 ura3-52 met2-*Δ*1 lys2-*Δ*2 can1-100 ade2-1*) and PJ51-3a, which has the same genotype as PJ43-2b except that it is *MAT***a**. PJ43-2b and PJ51-3a were obtained from Phil James (University of Wisconsin). Genomic DNA was isolated from several *HIS*$^+$ transformants, digested with *Hin*dIII, and analyzed by Southern blotting to confirm the disruption of one copy of the genomic *NRD1* locus.

ACT-CUP plasmid constructions. pGAC24, a *LEU2*-marked 2μm plasmid containing the ACT-CUP fusion gene under the control of the *GPD* promoter and *PGK* terminator, was described previously (36). Many of the ACT-CUP constructs described below were also moved into pRS426, a *URA3*-marked 2μm plasmid (10), by subcloning a 4.2-kb *Hin*dIII fragment that contained the *LEU2* gene as well as the ACT-CUP gene from pGAC24 to create pRS426AC and derivatives.

The branch point consensus sequence mutation in which adenine is substituted for cytosine at position 256 of the intron (C256A) was created by site-directed mutagenesis (32) of a *Bam*HI-*Sal*I fragment containing the ACT-CUP coding region cloned into pRS313 (57). Standard cloning procedures were then used to recombine the branch point mutation with the U6 insertion constructs.

To introduce the U6 sequence into the *Xho*I site in the *ACT1* intron in pGAC24, the transcribed portion of the yeast U6 gene (7) was amplified using the U6 5′*Afl*III and U6 3′*Bam*HI oligonucleotide primers. The product was digested with *Bam*HI and *Afl*III, incubated with *E. coli* DNA polymerase I Klenow fragment and deoxynucleoside triphosphates to fill in the sticky ends, and blunt-end ligated to *Xho*I-digested and Klenow-treated pGAC24. This yielded pGAC24-U6F, which contains the U6 sequence in the forward orientation, and pGAC24-U6R*, which contains the U6 sequence in the reverse orientation as well as 5′ and 3′ flanking sequences derived indirectly from the primer oligonucleotides. The sequence of the U6R* insert is as follows: ctcgaCATGTAGTTGCGACATGTAGTTGCGAGATCCGT**AAAACGAA**(94 nt) **CTTCGCGAACATGT**AGTTGCGAAAGGGTTACTTCGCGAACATGtcgagagattt, where intron sequence is represented by lowercase letters, inserted sequence is represented by capital letters, and the antisense U6 sequence is represented by bold-faced type. *Nru*I restriction sites are underlined, and *Afl*III sites are doubly underlined.

Deletion variants of the U6R* construct were created using pRS426AC-U6R* as the parent vector. pRS426AC-U6R*ΔNru was created by digestion with *Nru*I and religation, removing 30 nt of flanking sequence from the 3′ end of the U6R* insertion. The resulting 3′ junction sequence, **CTTCGCGAAC**ATGtcgaga (in which the antisense U6 sequence is in bold-faced capital letters, the recreated *Nru*I site is underlined, and the intron sequence is in lowercase letters) is identical to that in the U6R construct (see below). U6R*Δ13 was created by deleting sequence between the two *Afl*III sites in the 5′ flanking portion of U6R*ΔNru. The R*ΔU6 construct was created by site-directed mutagenesis of the *Bam*HI-*Sal*I fragment of U6R*ΔNru cloned into pRS313, and the mutated fragment was then subcloned back into pRS426AC or pGAC24. U6R*ΔNE was created by replacing a *Bam*HI-*Nru*I fragment of pRS426AC-U6R*ΔNru with a *Bam*HI-*Eco*NI fragment that had its *Eco*NI end filled in with Klenow fragment, thereby deleting 67 nt of antisense U6.

To generate pGAC24-U6R, which contains the U6 sequence in the reverse orientation but without the flanking sequences of U6R*, the U6 sequence was amplified with U6 5′*Xho*I and U6 3′*Xho*I primers, and the product was then digested with *Xho*I and ligated directly into pGAC24 digested with *Xho*I. A product of this ligation containing the U6 sequence in the reverse orientation was then digested with *Nru*I, which cuts once in the vector upstream of the ACT-CUP gene and once at the 3′ end of the U6R sequence. The resulting 1-kb fragment was then cloned into *Nru*I-digested pGAC24-U6R* to generate pGAC24-U6R, in which the 5′ end of antisense U6 directly abuts the intron *Xho*I site as follows: ttgctcgag**AAAACGAAA** (where the *Xho*I site is underlined and the antisense U6 sequence is in bold-faced capital letters). The sequence of U6R at the 3′ antisense U6-intron junction is identical to that in the U6R*ΔNru construct. Thus, U6R and U6R*ΔNru differ only in the 33 nt of 5′ flanking sequences present in the latter in place of the last G residue of the *Xho*I site.

Oligonucleotides. The oligonucleotides used in this study are as follows: U6 5′*Afl*III, 5′-TCGCAACTACATGTTCGCGAAGTAACCCTT; U6 3′*Bam*HI, 5′-CGGGATCCGTAAAACGAAATAAATCTCTTTGT; U6 5′*Xho*I, 5′-TCCGCTCGAGGTCGCGAAGTAACCCTT; U6 3′*Xho*I, 5′-TCCGCTCGAGAAAACGAAATAAATCTCT; ACT1-C256A, 5′-ATTTGCTACTGTGTCTCATGTAATAACATCG; R*ΔU6, 5′-TGCGAGATCCGTATGTCGAGAGAT; 3′-CUP, 5′-CTTCATTTTGGAAGTTAATTAATT; U5B, 5′-AAGTTCCAAAAAATATGGCAAGC; Sen1 OLI2, 5′-GGGAGCTAATGGCTCAAGCTGG; and Sen1 20397, 5′-TAAACGTTGCAATTAAGG.

RNA analysis. Samples for RNA preparation were harvested from cultures growing at log phase in liquid media at 30°C unless otherwise indicated. Selection for the ACT-CUP plasmid was maintained by growth in dropout medium lacking either leucine (for pGAC24 derivatives) or uracil (for pRS426 derivatives), but never under conditions (i.e., in the presence of added copper) that selected for ACT-CUP expression, to eliminate selection for an increased copy number of plasmids containing poorly expressed ACT-CUP genes. Total yeast RNA was prepared by the glass bead-phenol method (63). Alternatively, some samples were prepared by the guanidinium thiocyanate method (25). RNA extracted from identical cultures by the two methods gave very similar results, although the second procedure gave a slightly higher recovery of RNA per microgram of nucleic acid (determined by A_{260} readings). The first method was adopted for most experiments because of its simplicity.

For primer extension analysis, a procedure based on that of Frank and Guthrie (19) was used. The 3′-CUP oligonucleotide used for analysis of ACT-CUP mRNA is complementary to *CUP1* sequence just downstream of the ACT-CUP fusion junction and is identical to that used previously (36). As an internal control, an oligonucleotide (U5B) complementary to the spliceosomal RNA U5 was used. Oligonucleotides were 5′ end labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase and separated from unincorporated label on G25 D-select spin columns (5′-3′ Inc.). Labeled oligonucleotide (50 to 100 fmol) was annealed to 10 μg of dried RNA in 6 μl of annealing buffer (70 mM NaCl, 50 mM Tris-HCl [pH 8.0], 10 mM dithiothreitol) by incubating the reaction mixture at 95°C for 1 min and 68°C for 3 min and then freezing it immediately on dry ice. Samples were slowly thawed, 9 μl of extension mix (20 mM NaCl, 17 mM Tris-HCl [pH 8.0], 3.4 mM dithiothreitol, 10.4 mM $MgCl_2$) containing 5 to 10 U of avian myeloblastosis virus reverse transcriptase (United States Biochemical) was added, and reaction mixtures were incubated at 37°C for 5 min and then at 42°C for 20 min. Formamide sample loading buffer (7.5 μl) was then added, and samples were heated to 95°C for 3 min before being electrophoresed on 6% polyacrylamide (19:1)–8 M urea sequencing gels. Gels were fixed in 5% methanol—5% acetic acid and dried before exposure to X-Omat AR5 film (Kodak) with an intensifying screen. Gels were also exposed to a phosphor storage screen for quantitation on a Molecular Dynamics PhosphorImager SI using ImageQuant software.

For Northern (RNA) blot analysis, 5 μg of total yeast RNA per lane was run on 5% acrylamide (19:1)–8 M urea gels and electroblotted and UV cross-linked onto Zetaprobe membranes (Bio-Rad). Antisense RNA probes were generated by T7 RNA polymerase (United States Biochemical) transcription of a *Bam*HI-*Cla*I fragment of the U6R*ΔNru ACT-CUP construct in pRS426, linearized at the *Bam*HI site, using [α-^{32}P]UTP. The resulting probe is complementary to the 5′ exon and most of the intron of the U6R*ΔNru pre-mRNA (see Fig. 3). Blots were prehybridized for 1 to 2 h at 65°C in a solution containing 0.9 M NaCl, 1% sodium dodecyl sulfate (SDS), 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA and probed in the same solution at 65°C overnight. Blots were washed twice in 2× SSC–1% SDS (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 65°C and twice in 0.1× SSC–1% SDS at 65°C and then exposed to X-ray film and the phosphor screen.

Nucleotide sequence accession number. The GenBank accession number for the sequence presented in this paper is U28487.

RESULTS

An intronic element that contains antisense U6 RNA inhibits gene expression. The experiments described below utilize the *ACT1-CUP1* fusion reporter gene developed by Lesser and Guthrie (36). This gene contains the first 7 codons of the actin coding region (interrupted in codon 4 by a 303-nt intron) fused

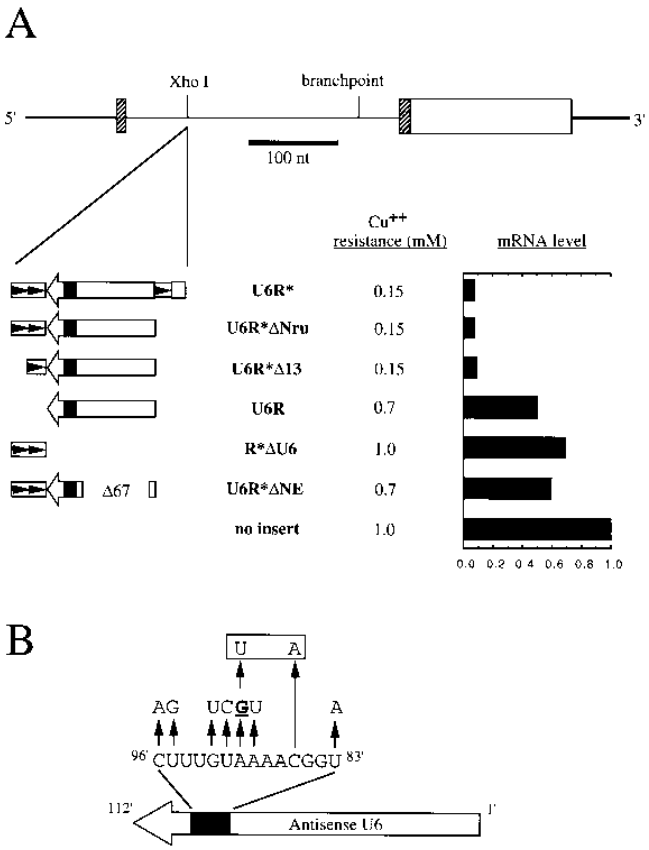

FIG. 1. Altered expression of ACT-CUP fusion genes with antisense U6 RNA sequences in the intron. (A) Schematic diagram of the ACT-CUP pre-mRNA, showing the location and structures of sequences inserted into the intron. The actin coding region is represented by the hatched boxes, and the Cup1 coding region is indicated by the open box. For sequences inserted into the *Xho*I site of the actin intron, the leftward-pointing open arrow represents antisense U6 RNA sequence, with the filled box designating the region of suppressor point mutations (see panel B). The boxed arrowheads represent direct repeats present in the sequences flanking antisense U6. The maximum copper concentrations tolerated by yeast strain 46α containing the indicated constructs in pGAC24 are listed. Copper concentrations tested were 0.05, 0.15, 0.25, 0.4, 0.7, and 1.0 mM. The histogram indicates the steady-state level of mature ACT-CUP mRNA in each strain, determined by primer extension analysis (see Fig. 2) and presented as the ratio of mRNA to U5 RNA, normalized to the wild-type ACT-CUP fusion (no insert). (B) The sequence of the 14-nt region of antisense U6 defined by *cis* suppressor point mutations. Arrows indicate the seven single point mutant suppressors and one double point mutant (boxed). The A90′G mutation that was recovered independently four times is underlined. Numbering is relative to the 5′ end of sense U6 RNA; for example, nt 1′ is complementary to nt 1 of U6 RNA.

in frame to the Cup1 coding region. Cup1 is a copper-chelating protein that is required by yeast for growth in the presence of Cu^{2+} (26). When a plasmid-borne *ACT1-CUP1* fusion gene is provided in a yeast strain from which the chromosomal *CUP1* locus has been deleted, growth of cells on media containing various concentrations of copper is a good indicator of fusion protein expression levels over a broad range and consequently is a sensitive reporter of splicing efficiency as well as other aspects of gene expression (36, 37).

As part of our studies on the function of the U6 spliceosomal RNA, we introduced the 112-nt sequence of the yeast U6 RNA (7) into the *Xho*I site in the intron of the *ACT1-CUP1* fusion gene (Fig. 1A) in both orientations. In the forward orientation, the U6 sequence exerted no effect on the expression of the fusion gene, as assessed by growth on copper-containing media (data not shown). However, insertion of the U6 sequence in the reverse orientation (U6R) resulted in moderate sensitivity to copper (no growth at concentrations of >0.7 mM). A construct containing U6R as well as 5′ and 3′ flanking sequences (generated fortuitously during cloning) at both ends of the U6 sequence (U6R*; see Materials and Methods) resulted in severe sensitivity to copper at concentrations of >0.15 mM and a 20-fold decrease in the ACT-CUP mRNA level (Fig. 1A).

The U6R* insertion differs from the precise U6R insertion in the presence of 5′ and 3′ flanking sequences of 33 and 34 nt in length, respectively. Since U6R results in only modest copper sensitivity, one or both of the flanking sequences must be required for the expression defect of U6R*. To identify the minimum required flanking sequences, the effect of partial deletions of the U6R* element on copper sensitivity and ACT-CUP mRNA level was determined. Deletion of the 3′ flanking sequence (yielding U6R*ΔNru) did not relieve copper sensitivity or the RNA accumulation defect (Fig. 1A), indicating that the 5′ flanking sequence in combination with antisense U6 is sufficient. The 5′ flanking sequence contains two overlapping copies of a 16-nt sequence. Deletion of one of the repeats (yielding U6R*Δ13) did not alleviate the RNA accumulation defect (Fig. 1A); thus, only a single copy is required for the full effect on expression. In the absence of antisense U6 (R*ΔU6), the 5′ flanking sequence did not significantly inhibit ACT-CUP expression, and thus it is not sufficient for the effect.

A 14-nt region of the antisense U6 RNA sequence required for the expression defect is defined by spontaneous *cis* suppressor point mutations (Fig. 1B) (see below). To determine if other regions of the antisense U6 sequence are also required, we deleted 67 nt of antisense U6 distal to the *cis* suppressor region. This deletion (U6R*ΔNE) restores mRNA expression to 60% of the wild-type level (Fig. 1A), indicating that important features are contained in the distal portion of antisense U6. This region appears to contain redundant features, since inclusion of either 5′ or 3′ portions of the 67 nt resulted in restoration of much of the expression defect (data not shown).

***cis* suppressor mutations define a central sequence-specific portion of the element.** Although the U6R* insertion in the *ACT1-CUP1* fusion gene intron prevents growth on plates containing >0.15 mM copper, exceptional colonies resistant to copper concentrations of 0.4 mM or higher arise spontaneously at a frequency of $\sim 10^{-7}$. We have analyzed approximately 30 of these copper-resistant revertants. Seventeen contained plasmid-linked suppressor mutations allowing increased expression of the fusion gene, since recovery of their ACT-CUP plasmids and transformation into the parental strain resulted in copper-resistant growth. Sequencing of the ACT-CUP intron in each of these plasmids revealed that the *cis*-acting suppressor mutations fall into two categories: (i) complete deletion of the reverse U6 sequence, facilitated by direct repeats of 14 bp present in the 5′ and 3′ flanking sequences of U6R* (six independent isolates); and (ii) 10 single point mutants and 1 double point mutant clustered in a 14-nt region near the 5′ end of the reverse U6 sequence (Fig. 1B). One of these mutations (A90′G) was recovered four times independently. For all four of the point mutations tested, primer extension analysis of RNA revealed significantly increased levels of the spliced ACT-CUP mRNA (data not shown).

To test whether the point mutant suppressors act by disrupting complementarity to U6 RNA, individual point mutations complementary to one of two suppressor mutations (A90′G or G92′U) were introduced into the U6 gene (yielding U6-U90C or U6-C92A). Both mutations in U6 RNA were found to be viable, but neither suppressed the accumulation defect of U6R* pre-mRNA (data not shown). Furthermore, the sup-

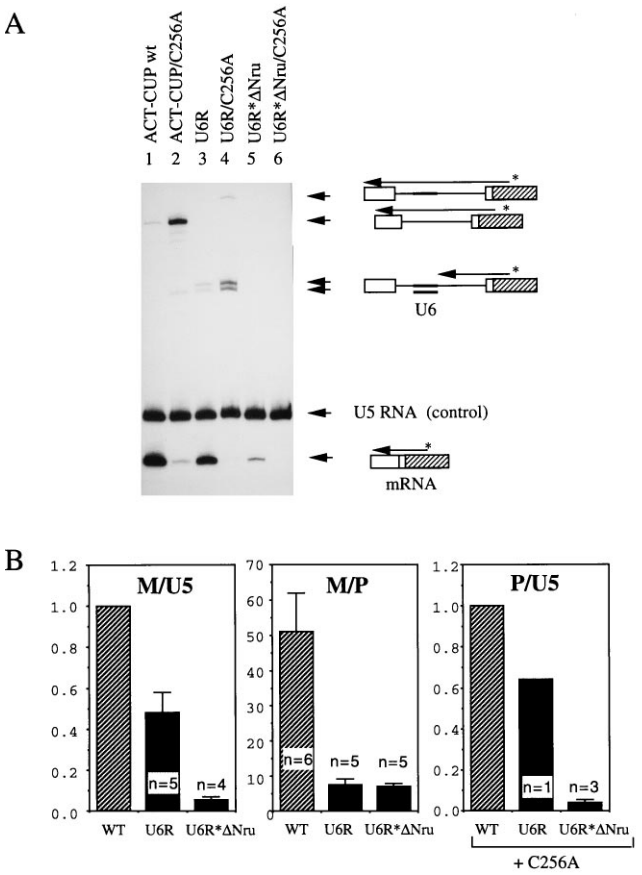

FIG. 2. Steady-state levels of ACT-CUP mRNA and pre-mRNA with antisense U6 RNA sequences in the intron. (A) Total yeast RNA preparations (10 μg) from the *cup1*Δ strain, 46α, containing the indicated ACT-CUP fusion genes were analyzed by primer extension using a probe complementary to *CUP1* sequence in the second exon (see Materials and Methods). A low-specific-activity probe for U5 RNA provides an internal control for total RNA levels. Extension products derived from spliced mRNA and pre-mRNA are indicated. Symbols to the right indicate the identities of the primer extension products, with hatched boxes symbolizing *CUP1* sequence, open boxes indicating *ACT1* exons, arrows indicating cDNA products, and asterisks designating ^{32}P-labeled primer. Lane 1, wild-type ACT-CUP fusion; lane 2, C256A branch point mutation in an otherwise wild-type ACT-CUP fusion; lane 3, ACT-CUP gene with U6R insertion; lane 4, C256A branch point mutation in the U6R construct; lane 5, ACT-CUP gene with a U6R*ΔNru insertion; lane 6, C256A branch point mutation in the U6R*ΔNru construct. (B) Quantitation of primer extension products from gels like that in panel A, including mRNA levels (relative to U5 RNA and normalized to wild-type [WT] ACT-CUP fusion) (M/U5), ratios of mRNA to pre-mRNA (M/P), and levels of pre-mRNA in the context of the branch point mutation C256A (P/U5).

pression afforded by the *cis* mutation, A90′G, was not reverted by the complementary mutation, U90C, in U6 RNA. The failure of the U90C and C92A mutations in U6 RNA to act as *trans* suppressors allows us to conclude that the complementary nucleotides in antisense U6 are not important because of their complementarity to U6 per se.

The U6R* element causes decreased accumulation of pre-mRNA. To examine whether decreased expression of the ACT-CUP fusion protein from the insertion constructs is due to inhibition of splicing, we determined the steady-state ratio of mature spliced ACT-CUP mRNA (M) to unspliced pre-mRNA (P) by primer extension with a second-exon probe (Fig. 2). Changes in splicing efficiency are most accurately represented by changes in the M/P ratio, while any influence of mutations on pre-mRNA stability is reflected proportionally in both the mature and precursor RNA pools (47).

The wild-type actin intron in the ACT-CUP fusion is spliced very efficiently, with abundant spliced mRNA and very little accumulation of pre-mRNA (Fig. 2A, lane 1). Primer extension of RNA from the U6R insertion construct shows somewhat reduced levels of spliced mRNA, as well as a faint novel doublet of intermediate mobility (lane 3). Northern blot analysis (see below) shows that full-length pre-mRNA and spliced mRNA are the only major ACT-CUP products in this RNA preparation, suggesting that the truncated cDNA products are derived from full-length pre-mRNA as the result of premature reverse transcriptase stops. The size of the aberrant cDNAs is consistent with reverse transcriptase stops near the 3′ end of the antisense U6 sequence, suggesting that extension to the 5′ end of the pre-mRNA may be prevented by hybridization of the antisense U6 sequence to U6 RNA present in the preparation. This was confirmed by adding a cocktail of oligonucleotides complementary to U6 RNA to prevent its base pairing to antisense U6 during the annealing step of the primer extension assay, which increased the yield of cDNA extended to the 5′ end of the pre-mRNA (data not shown). Using these truncated products as well as the small amount of full-length cDNA product to represent the abundance of full-length pre-mRNA, quantitation of the M/P ratio revealed that splicing of the U6R intron is inhibited seven- to eightfold compared with that of the wild-type intron, although the steady-state level of mature mRNA is decreased only approximately twofold (Fig. 2B).

Also shown in Fig. 2 are the results of analysis of RNA from the U6R*ΔNru construct. The most striking result is the very low abundance of both spliced mRNA and unspliced U6R*ΔNru precursor (lane 5). As with the U6R RNA, primer extension of the full-length pre-mRNA appears to be blocked, resulting in truncated cDNA products (visible on longer exposure). Quantitation of the low levels of mRNA and pre-mRNA represented by the primer extension products indicates that they occur in the same ratio as in the U6R RNA sample, although the steady-state level of mature mRNA is 20-fold less than is produced from the wild-type pre-mRNA and 10-fold less than is produced from the U6R pre-mRNA (Fig. 2B). Thus, in addition to the splicing defect seen with the precise antisense U6 sequence, the U6R*ΔNru insert inhibits the accumulation of both mRNA and pre-mRNA.

To examine the pre-mRNA accumulation deficit independently of the splicing inhibition phenotype, we introduced a mutation in the intron branch point sequence (UA**A**UAAC, or C256A) that strongly inhibits splicing of the actin intron, leading to the accumulation of pre-mRNA (46, 65). Quantitation of pre-mRNA levels by primer extension showed that the U6R/C256A pre-mRNA (Fig. 2A, lane 4) accumulates to approximately 70% of the level of wild-type ACT-CUP/C256A pre-mRNA (lane 2), but the U6R*ΔNru/C256A pre-mRNA (lane 6) accumulates to only about 5% of the level of wild type/C256A or 7% of that of U6R/C256A pre-mRNA (Fig. 2B). This indicates that a severe defect in expression of the U6R* constructs occurs at the level of pre-mRNA accumulation, independently of the catalysis of splicing.

Appearance of 3′-truncated pre-mRNA products coincides with the accumulation defect. The decreased accumulation of the U6R*-containing pre-mRNA could result from a block in transcript initiation or elongation or from enhancement of posttranscriptional turnover of pre-mRNA. The primer extension analysis presented above employed a single oligonucleotide probe complementary to sequences in the 3′ exon and thus was not well suited for the detection of products of stalled transcription or possible degradation intermediates. To test for the presence of such products, Northern blot analysis was performed with a probe complementary to the 5′ exon and

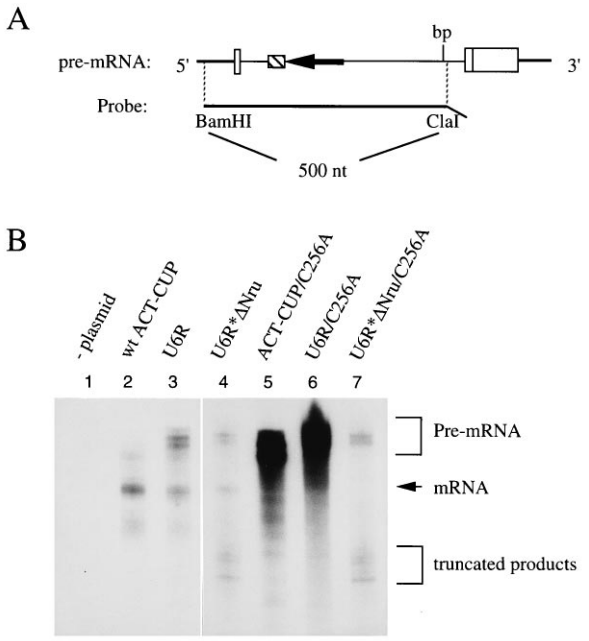

FIG. 3. Northern blot analysis of ACT-CUP pre-mRNA with the U6R* sequence in the intron reveals 3′-truncated products. (A) The probe is an antisense RNA complementary to most of exon 1 and the intron to just downstream of the branch site. Open boxes, protein-coding region; thick lines, 5′ and 3′ untranslated regions; thin lines, intron sequences; arrow; antisense U6 sequences; hatched box, remainder of UGR*ΔNru element; bp, branch point. (B) Northern blot showing truncated pre-mRNA fragments present in the U6R*ΔNru and U6R*ΔNru/C256A samples. Lane 1, strain 46α with no ACT-CUP plasmid; lane 2, wild-type (wt) ACT-CUP fusion with no insertion in intron; lane 3, U6R insertion in ACT-CUP intron; lane 4, U6R*ΔNru insertion in intron; lane 5, C256A branch point mutation in an otherwise wild-type ACT-CUP fusion; lane 6, C256A mutation in the U6R construct; lane 7, C256A mutation in the U6R*ΔNru construct. The probe hybridizes more efficiently to pre-mRNA than to mRNA, as judged by comparison to primer extension results (Fig. 2).

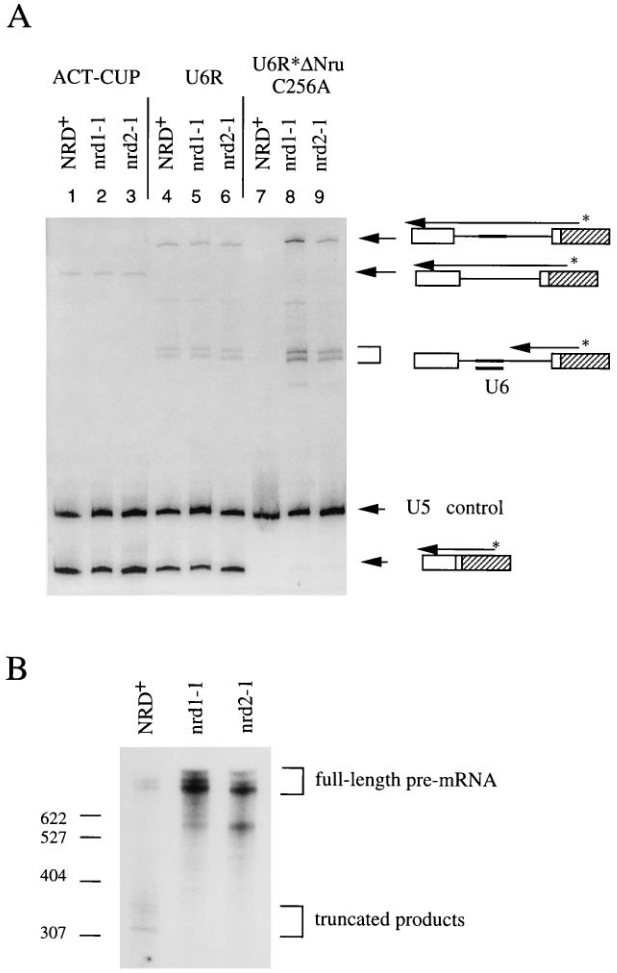

FIG. 4. The *nrd1-1* and *nrd2-1* mutations relieve the pre-mRNA accumulation defect and prevent the generation of truncated products from the U6R*-containing pre-mRNA. (A) Primer extension analysis of wild-type ACT-CUP, U6R, and U6R*ΔNru/C256A mRNA and pre-mRNA in *NRD*$^+$, *nrd1-1*, and *nrd2-1* strains at the permissive temperature (23°C). Lanes 1 to 3, RNA from strains containing the wild-type ACT-CUP fusion; lanes 4 to 6, RNA from strains containing the U6R insertion in the ACT-CUP fusion; lanes 7 to 9, RNA from strains containing the U6R*ΔNru insertion in the ACT-CUP fusion gene that also contains the C256A branch point sequence mutation. Lanes 1, 4, and 7, *NRD*$^+$ parent strain (46α); lanes 2, 5, and 8; *nrd1-1* strain; lanes 3, 6, and 9, *nrd2-1* strain. Positions of primer extension products and the U5 RNA primer extension control are indicated. Symbols are defined in the legend to Fig. 2A. (B) Northern blot analysis of U6R*ΔNru/C256A pre-mRNA in *NRD*$^+$, *nrd1-1*, and *nrd2-1* strains. The probe was the same as that used in the experiment whose results are shown in Fig. 3. Molecular weight markers were single-stranded *Msp*I fragments of pBR322, and their sizes (in nucleotides) are indicated on the left.

most of the intron of the U6R*ΔNru pre-mRNA (Fig. 3A). Several discrete minor products were detected in samples of total RNA from either the U6R*ΔNru or U6R*ΔNru/C256A constructs, but not in samples from the wild-type ACT-CUP or U6R constructs or their C256A derivatives (Fig. 3B).

These RNA fragments are derived from the 5′ portion of the gene, since they were also detected with a smaller probe complementary to only exon 1 and the 5′ portion of the intron (data not shown), and must include intron sequences, since identical products are present in samples from the wild-type and mutant branch point samples. Assuming that these RNA molecules retain intact 5′ ends, the size of the smallest of these fragments (approximately 320 nt) places its 3′ end very close to the 3′ junction of the U6R*ΔNru insertion with intronic sequences. This coincides with a potential RNA hairpin that can form in the antisense U6 sequence, which may form a barrier to 3′-5′ exonucleolytic digestion or may be a determinant of transcriptional elongation arrest (see Discussion). The doublet of more slowly migrating species at 350 to 360 nt corresponds to 3′ ends occurring within the intron sequence, about 30 to 40 nt downstream of the insertion site, and does not correspond to any obvious RNA secondary structure, but it may fall within a long pyrimidine-rich stretch containing several runs of U residues.

Genomic loci involved in U6R*-dependent control of pre-mRNA accumulation. In addition to the *cis*-acting suppressor mutations described above, several spontaneously arising mutations unlinked to the ACT-CUP plasmid allowed growth of the U6R*-containing strain on media with copper concentrations of >0.4 mM. Preliminary analysis of RNA from several of these putative *trans* suppressor strains revealed the presence of increased levels of both mature ACT-CUP mRNA and pre-mRNA. We chose one of these suppressor strains, which had a tight temperature-sensitive growth defect that was recessive in a backcross to an otherwise isogenic strain, for more detailed study. Upon sporulation of this heterozygous diploid, the temperature sensitivity and suppression phenotypes cosegregated as a single Mendelian determinant in each of 13 tetrads analyzed, indicating tight linkage of the two phenotypes.

A haploid *MAT***a** strain derived from this backcross that

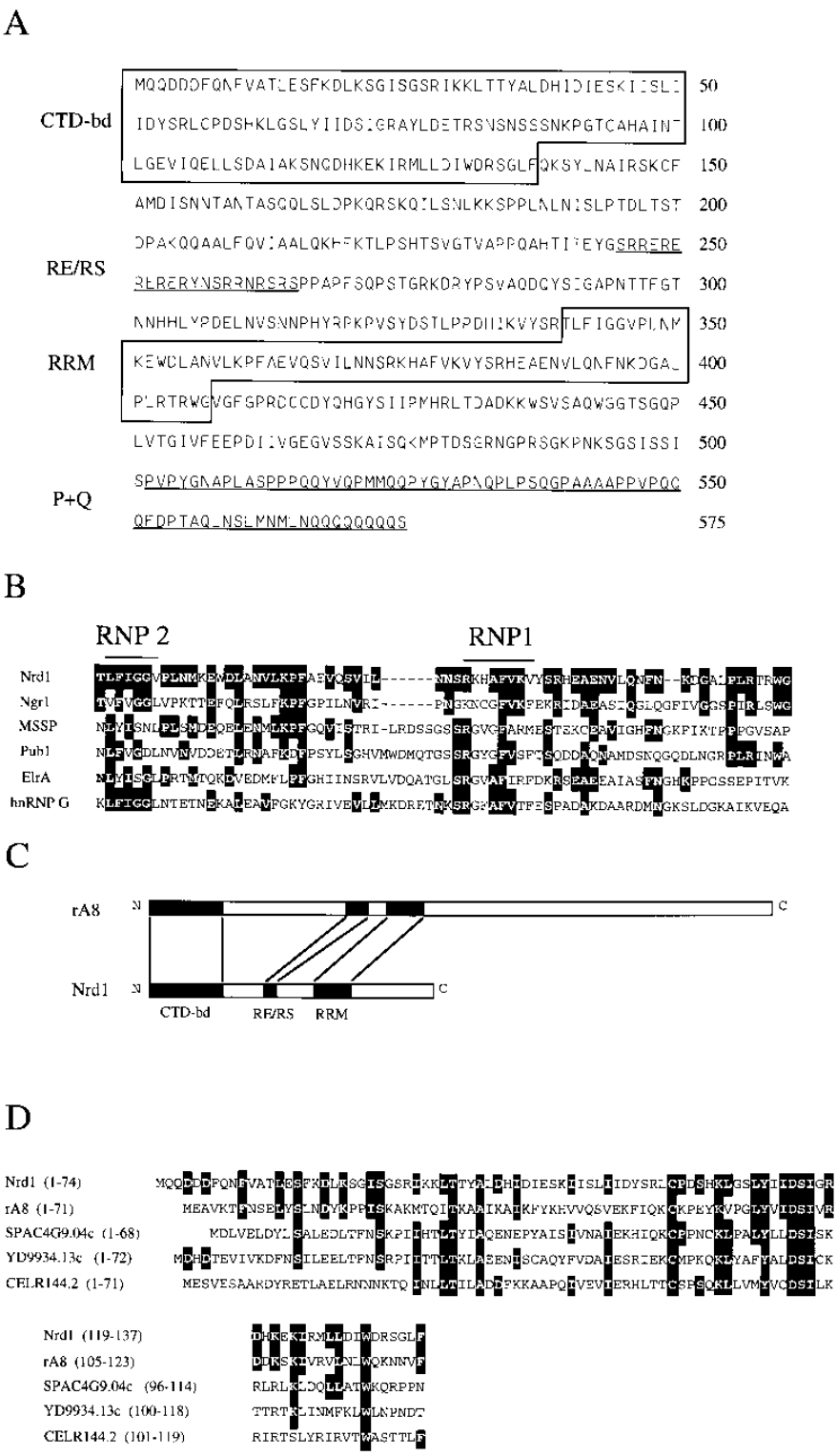

FIG. 5. Nrd1 is a novel hnRNP-like protein. (A) Predicted amino acid sequence of Nrd1. The RE/RS segment and P- and Q-rich region are underlined, and the RNA polymerase II CTD-binding (CTD-bd) and RRM domains are boxed. (B) Alignment of the Nrd1 RRM to several of the most closely related RRMs identified by the BLASTP program (2). Residues showing identity to the Nrd1 protein sequence are boxed. Nrd1 protein, residues 340 to 407; yeast Ngr1 (Rbp1) protein, residues 361 to 430 (33); human *myc* single-stranded-DNA-binding protein (MSSP), residues 109 to 183 (43); yeast Pub1 protein, residues 163 to 238 (3, 39, 50); *Xenopus* elav-like ribonucleoprotein (elrA), residues 107 to 182 (22); human hnRNP G protein, residues 10 to 84 (58). (C) Schematic illustration of primary structure of yeast Nrd1 and rat rA8 (68) proteins showing their similar domain organizations. (D) Alignment of the N-terminal CTD-binding domains of Nrd1 (residues 1 to 74 and 119 to 137) and rat rA8 (residues 1 to 71 and 105 to 123) proteins. Also shown are the N-terminal sequences of putative gene products from *Schizosaccharomyces pombe* (SPAC4G9.04c gene product; GenBank accession number Z69727; residues 1 to 68 and 96 to 114), *S. cerevisiae* (YD9934.13c gene product; GenBank accession number Z48612; residues 1 to 72 and 100 to 118), and *C. elegans* (CELR144.2 gene product; accession number U23515; residues 1 to 71 and 101 to 119). Residues identical to those in the Nrd1 protein are boxed.

contained the suppressor mutation causing temperature sensitivity was then used in genetic complementation tests against a series of *MAT*α suppressor strains that also had temperature-sensitive phenotypes scoring as recessive or partially recessive in the parental backcross. These crosses defined two complementation groups, one represented by six independent alleles and a second defined by a single suppressor mutation. We have designated these complementation groups *nrd1* and *nrd2*, respectively (for nuclear pre-mRNA down-regulation).

To allow quantitation of U6R* pre-mRNA levels in these strains without the additional complication of possible effects on splicing efficiency, we again used the C256A branch point mutation in the context of the U6R*ΔNru construct. Primer extension analysis of RNA from the *nrd1-1* and *nrd2-1* suppressor strains showed 30- and 10-fold-increased abundance of U6R*ΔNru/C256A pre-mRNA relative to the parent strain, respectively (Fig. 4A, lanes 7 to 9). However, mRNA levels and pre-mRNA/mRNA ratios produced from the wild-type ACT-CUP fusion (lanes 1 to 3) and from the U6R construct (lanes 4 to 6) are indistinguishable in the *nrd1-1*, *nrd2-1*, and wild-type parent strains. Thus, the *trans* suppressor mutations specifically relieve the pre-mRNA accumulation defect due to the U6R* insertion but not the moderate splicing defect due to the precise U6R insertion. Furthermore, *nrd1-1* and *nrd2-1* do not act as general enhancers of ACT-CUP fusion gene expression.

We next examined the RNA products of the U6R*ΔNru/C256A construct in the *nrd1-1* and *nrd2-1* strains by Northern blot analysis. As shown in Fig. 4B, both mutant strains prevent the accumulation of truncated fragments and greatly increase the accumulation of full-length pre-mRNA in comparison to the wild-type parent. This strongly suggests that the underexpression of full-length pre-mRNA and the accumulation of truncated transcripts are causally related.

***NRD1* encodes a novel hnRNP-like protein.** We cloned the *NRD1* gene by complementation of the temperature sensitivity of the *nrd1-1* strain by using a YCp50-based yeast genomic library (52). Two independent transformants selected at 37°C contained apparently identical complementing plasmids with an insert of approximately 10 kb. Deletion analysis localized the complementing gene to a 4.8-kb *Hin*dIII-*Sal*I fragment, and an internal *Eco*RI site further divides this fragment into a weakly complementing 2.7-kb *Hin*dIII-*Eco*RI fragment and a noncomplementing *Eco*RI-*Sal*I fragment (see Fig. 6A). DNA sequencing revealed a single large open reading frame (ORF) between the *Hin*dIII and *Eco*RI sites, ending just downstream from the *Eco*RI site and encoding a predicted polypeptide of 575 amino acids (Fig. 5A). The sequence upstream of the ORF overlaps with the reported sequence of a clone containing the *RAD50* gene (1) and thus localizes the cloned DNA to the left arm of chromosome XIV, with the newly identified ORF and the *RAD50* gene divergently transcribed and their coding regions separated by 1.1 kbp. That this ORF represents the *NRD1* gene was confirmed by the identification of mutations in several independently selected, spontaneously arising *nrd1* mutant strains (see below).

The sequence of the predicted Nrd1 protein contains several features relevant to its potential role in pre-mRNA metabolism (Fig. 5). The region between residues 340 and 410 contains sequence matching the RNA recognition motif (RRM), or a ribonucleoprotein consensus sequence derived from a large number of RNA-binding proteins (5, 8, 29). The RNP2 hexamer of Nrd1, LFIGGV (residues 341 to 346), is very similar to those of mammalian hnRNP A, B, and G proteins. The RNP1 octamer, RKHAFVKV (residues 374 to 381), is more atypical, particularly in the presence of the lysine at the second position of the octamer, shared only with RRM 1 of nucleolins, and the histidine residue at position 3, unique among reported RRM protein sequences.

Another segment of Nrd1 with potential for a role in pre-mRNA metabolism is the region between residues 245 and 265, which contains a high density of arginine, serine, and glutamate residues, with arginine often alternating with glutamate or serine (RE/RS). Similar but often more extensive RS or RE/RS domains occur in a number of metazoan RRM-containing splicing factors, including members of the SR protein family and related proteins (6, 21).

The C-terminal 75 residues of Nrd1 are rich in glutamine (Q) and proline (P) residues (25 and 21%, respectively), including a stretch of 8 consecutive Q residues just prior to the C-terminal residue. Several yeast hnRNP-like proteins, including Nab1, Nab2, and Nab3, contain similar P- and Q-rich domains (66). The Nrd1 P+Q domain is most similar in amino acid composition and in its carboxy-terminal location to that found in Nab3. Similar P- and Q-rich domains are also found in a number of metazoan RRM proteins, including the *Drosophila* elav (51) and the mammalian TIA-1 (62) proteins. In none of these cases has a function been ascribed to the P- and Q-rich segment.

The amino-terminal 140 residues of Nrd1 exhibit similarity to the amino-terminal domain of a recently identified mammalian SR-like protein called rA8 (68). This domain of rA8 and the corresponding region of Nrd1 have been shown to bind to the C-terminal domain (CTD) of the largest subunit of mouse RNA polymerase II in a yeast two-hybrid assay (68). In addition to sequence similarity in the N-terminal domain, rA8 has a domain organization similar to that of Nrd1, with an RE/RS dipeptide-rich region and a single RRM domain as well as several proline and/or glutamine-rich segments (Fig. 5C). Shown in Fig. 5D is the alignment of the N-terminal CTD-binding domains of Nrd1 and rA8, along with N-terminal domains of putative gene products from *S. cerevisiae*, *Schizosaccharomyces pombe*, and *Caenorhabditis elegans* identified in BLAST searches. Apart from their N termini, the last three proteins have a domain organization that differs from that of the Nrd1 and rA8 proteins but are highly related to each other through much of their lengths, including potential Zn^{2+} finger motifs. The function of the rA8 protein is not known, but it and related mammalian proteins have been proposed to couple transcription with pre-mRNA splicing (68).

Disruption of *NRD1* is lethal. To determine if *NRD1* is essential for growth, the gene was disrupted by substitution of a DNA fragment containing the *HIS3* gene for all but the N-terminal 38 codons of the *NRD1* coding region (Fig. 6A). A diploid strain heterozygous for the chromosomal disruption of *NRD1* was created by one-step gene disruption (53), and integration of the *nrd1*Δ::*HIS3* allele into the *NRD1* locus was confirmed by Southern blot analysis (data not shown). Sporulation of this strain gave rise to tetrads that segregated 2:0 for viability (Fig. 6B), with no viable *HIS*$^+$ spores recovered from dissection of 42 tetrads, suggesting that disruption of *NRD1* is lethal. Upon microscopic examination, many of the nonviable spores were found to undergo germination and several rounds of cell division before ceasing to grow.

To demonstrate that the disruption could be complemented by the cloned Nrd1 gene, the diploid strain heterozygous for the *nrd1*Δ::*HIS3* disruption was sporulated after introduction of *NRD1* on a plasmid marked with the *URA3* gene. *HIS*$^+$ *URA*$^+$ progeny but no *HIS*$^+$ *ura*$^-$ progeny were recovered from this dissection, whereas both *his*$^-$ *ura*$^-$ and *his*$^-$ *URA*$^+$ progeny were recovered. The resultant strains were tested for dependence on the plasmid-borne *NRD1* gene by plating them on medium containing 5-FOA to select for loss of the *URA3*-

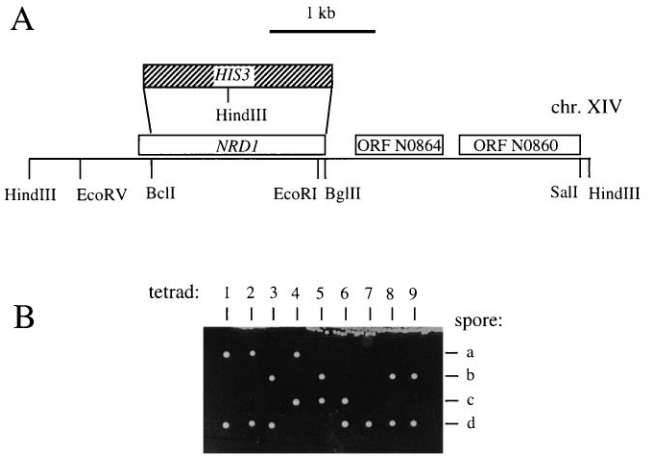

FIG. 6. Disruption of *NRD1* is lethal. (A) Diagram of the *NRD1* region of chromosome XIV showing the structure of the *nrd1*Δ::*HIS3* disruption construct. An *Eco*RV-*Sal*I fragment containing the disrupted Nrd1 allele was integrated into diploid strain EJS100, and integration into the *NRD1* locus on one copy of chromosome (chr.) XIV was verified by Southern analysis of *Hin*dIII-digested genomic DNA (data not shown). (B) An example of the 2:0 segregation pattern of dissected tetrads from the heterozygous disruption strain EJS101.

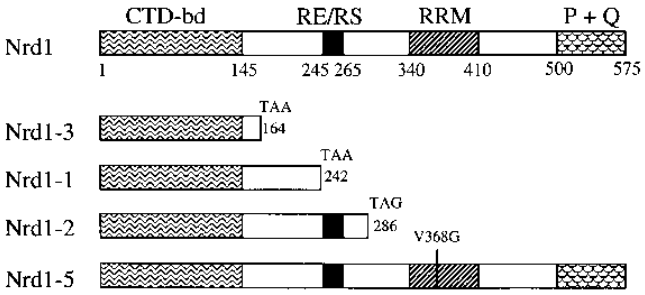

FIG. 7. Spontaneous suppressor alleles of *NRD1*. A schematic diagram of the Nrd1 protein is presented, with RNA polymerase II CTD-binding domains (CTD-bd), arginine-, serine-, and glutamate-rich segments (RE/RS), RRM, and proline- and glutamine-rich (P+Q) regions indicated by the patterned boxes. The predicted primary structures of the products of the *nrd1-1*, *-2*, *-3*, and *-5* alleles are shown below.

marked plasmid. The *HIS*$^+$ *URA*$^+$ progeny (containing the chromosomal *nrd1*Δ::*HIS3* allele) were unable to grow on plates containing 5-FOA at 18, 23, 30, or 37°C, but *his*$^-$ *URA*$^+$ colonies were viable on 5-FOA plates at all temperatures. We conclude that disruption of *NRD1* is lethal at all temperatures.

To identify the *NRD1* mutations responsible for temperature sensitivity and suppression of the pre-mRNA accumulation defect, the *nrd1-1*, *-2*, *-3*, and *-5* alleles were isolated by gap repair (53). Mutations were identified by sequencing and are summarized in Fig. 7. The *nrd1-1*, *-2*, and *-3* alleles each contain a single point mutation that creates a premature stop codon early in the coding region, such that their predicted products lack the RRM and P+Q domains. The truncated proteins encoded by *nrd1-1* and *nrd1-3* additionally lack the RE/RS domain. The *nrd1-5* allele results in the substitution of a glycine residue for valine at residue 368, within the RRM, at a site predicted to fall within β-strand 2 of the RRM structure based on alignment with RRM domains whose structures have been determined (42). The *nrd1-3* strain, although temperature sensitive, is viable at 18 to 30°C, suggesting that the N-terminal 163 amino acids of Nrd1 are sufficient for viability of yeast cells. However, a function(s) mediated by at least the RRM domain is clearly required for the U6R*-dependent pre-mRNA accumulation defect.

***NRD2* is identical to *SEN1*, a putative helicase.** The *NRD2* gene was cloned by complementation of the temperature sensitivity of the *nrd2-1* strain with the yeast genomic library. Five independent complementing clones were recovered from ~50,000 transformants of the *nrd2-1* strain. Preliminary restriction and sequence analyses revealed that the clones fall into two related classes containing overlapping fragments of chromosome XII. The 13.3-kb overlap shared by the two clone types contains three potential ORFs, one of which encodes Sen1 (15). Deletion of *SEN1* resulted in loss of complementation activity, and a clone containing only the Sen1 ORF was able to complement the temperature sensitivity of *nrd2-1*, suggesting that the *nrd2-1* mutation is in the *SEN1* gene.

Sen1 is an essential nuclear protein (15, 64). Although it is 2,231 amino acids in length, a previous study demonstrated that the carboxy-terminal 1,214 residues, containing a putative ATP-dependent helicase domain, are sufficient for viability (15). The similarity of the Sen1 helicase domain to that in the yeast Upf1 protein, which is involved in nonsense-mediated mRNA decay, and to a mouse protein, Mov-10, of unknown function has been noted previously (31, 34). Upf1 was recently demonstrated to have ATP-dependent helicase activity in vitro (11). The BLAST program (2) identifies a number of additional putative helicases from *S. cerevisiae*, *S. pombe*, *C. elegans*, and mammals which have a high degree of similarity to this domain of Sen1.

The chromosomal *nrd2-1* allele was isolated by gap repair (53). Sequencing of approximately 1.2 kb in the helicase domain identified a point mutation converting codon 1597 (1478 in reference 15) from GAA (coding for glutamate) to AAA (coding for lysine) (Fig. 8). This mutation, E1597K, alters a glutamate residue that is conserved in Upf1, Mov-10, and many of the other related putative helicases. It is located 7 residues C terminal to an aspartate residue (marked by an asterisk in Fig. 8) that is widely conserved among ATPases and ATP-dependent helicases (23, 27, 31) and has been implicated in positioning a Mg^{2+} ion thought to be involved in ATP hydrolysis (20, 60). The *nrd2-1* mutation thus is likely to affect the putative ATPase activity and helicase function of Sen1.

To demonstrate that the E1597K mutation is solely responsible for the temperature sensitivity and suppression phenotypes of the *nrd2-1* strain, linear DNA fragments encompassing codons 1457 to 1696 were generated by PCR amplification of the *SEN1* wild-type or *nrd2-1* mutant clones. These DNAs were used to transform the *nrd2-1* strain, and temperature-resistant revertants were selected at 37°C. Transformation with the PCR product from the wild-type clone yielded temperature-resistant revertant colonies that showed coreversion of the copper resistance phenotype. Transformation with the *nrd2-1* mutant PCR product failed to yield temperature-resistant revertants. We conclude that the E1597K mutation in Sen1 is responsible for both the temperature sensitivity and the suppression of the U6R*-dependent pre-mRNA accumulation defect.

DISCUSSION

A novel nuclear pathway for repression of gene expression in *S. cerevisiae*. We have identified an exogenous sequence element that causes underaccumulation of a reporter pre-mRNA when inserted into its intron and have defined two *trans*-acting factors that participate in this process. One of these factors, Nrd1, is a novel yeast hnRNP-like protein, and

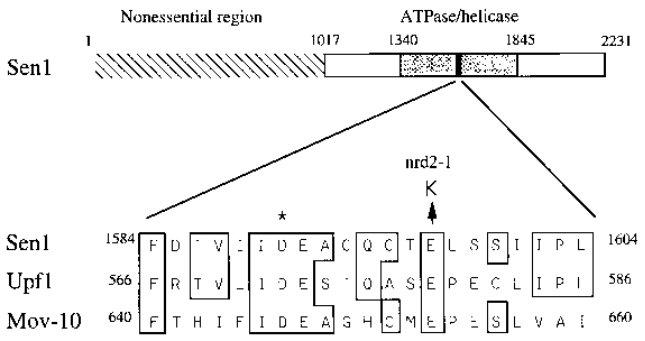

FIG. 8. The *nrd2-1* mutation alters a conserved residue near the predicted ATP hydrolysis site of Sen1. A linear diagram of the Sen1 protein is presented, with the nonessential N-terminal portion and the putative helicase domain indicated. The expanded region shows the sequences of Sen1, Upf1, and Mov-10 in the region of the *nrd2-1* mutation (E1597K). Residues in Upf1 and Mov-10 that are identical to those in Sen1 are boxed. The asterisk identifies an aspartate residue that is highly conserved among ATPases and helicases and is thought to participate in ATP binding and hydrolysis by positioning a Mg^{2+} ion that is coordinated to the β and γ phosphates of ATP (20, 60). The numbering of Sen1 amino acids corresponds to the ORF predicted by the sequence of chromosome XII as determined by systematic sequencing of the yeast genome (GenBank accession number 20939). This sequence differs from that previously reported (15) by the inclusion of an additional A residue in codon 130 of the current numbering system, shifting the frame of the Sen1 ORF such that the N-terminal 11 residues of the sequence of DeMarini et al. (15) are replaced by an additional 130 residues.

the second, Sen1, is a previously identified nuclear protein with putative helicase activity. Several observations indicate that control of pre-mRNA accumulation via this pathway occurs in the nucleus at a stage subsequent to the initiation of transcription but prior to splicing: (i) the *cis* element functions within an intron, more than 170 nt downstream of the transcription initiation site; (ii) Nrd1 contains a consensus RRM and other auxiliary domains similar to those found in a number of hnRNP proteins; and (iii) underexpression of *cis*-element-containing pre-mRNA is accompanied by the appearance of 3′-truncated pre-mRNA fragments containing partial intron sequences. Furthermore, the underexpression of pre-mRNA is independent of the cytoplasmic nonsense-mediated mRNA decay pathway (59a).

We hypothesize that Nrd1 binds via its RRM to the U6R* element in the nascent pre-mRNA and that this binding is a key step in specifying the pre-mRNA for down-regulation. The 14-nt sequence-specific subelement defined by *cis* suppressor mutations is an excellent candidate target site for recognition by Nrd1. Considering that pre-mRNA splicing is likely to occur cotranscriptionally, the problem of recognizing an intronic sequence element before it is spliced out implies that Nrd1 associates with its target site very soon after that site is synthesized. This suggests that Nrd1 may be associated with the transcription apparatus. It is therefore intriguing that a recently identified mammalian homolog of Nrd1 interacts with the repetitive CTD of the largest subunit of mouse RNA polymerase II in a two-hybrid screen (68). Indeed, the putative CTD-binding domain of Nrd1 also interacts with the mouse CTD in the two-hybrid assay, albeit weakly (68). Importantly, the mouse CTD can functionally replace the yeast CTD and so must be capable of mediating interactions essential for CTD functions in *S. cerevisiae* (68). It will be interesting to determine if the mammalian Nrd1 homolog can function in yeast cells.

Mutations in Sen1 affect a variety of essential nuclear functions. The *sen1-1* mutation, first identified in a screen for mutants defective in pre-tRNA splicing (15), was subsequently found to result additionally in mislocalization of the nucleolar proteins Ssb1 and Nop1 and in pre-rRNA processing defects (64). A *SEN1* allele (named *cik3-1*) has also been identified in a screen for mutations that cause chromosomal instability and defects in nuclear fusion (45, 64). Results presented here represent the first evidence of a role for the putative helicase activity of Sen1 in pre-mRNA metabolism. The involvement of Sen1 in Nrd1-dependent control of pre-mRNA accumulation provides further evidence that this pathway is localized to the nucleus.

Potential mechanisms for repression of pre-mRNA accumulation by the Nrd pathway. Physical association of Nrd1 with the CTD of RNA polymerase II could provide a mechanism for controlling pre-mRNA accumulation at the level of transcript elongation. Transcription of several proto-oncogenes is regulated at the level of elongation (59), and the importance of such control is further underscored by the recent finding that the product of the von Hippel-Lindau tumor suppressor gene negatively controls the elongation-enhancing activity of an RNA polymerase II auxiliary factor known as SIII or elongin (17, 30). Perhaps the specificity of elongation control in eukaryotes is endowed in part by factors like Nrd1, just as RNA-binding factors regulate elongation in *E. coli* (14, 24, 49). Interaction of Nrd1 with the CTD, and any consequent effects on transcript elongation, may be modulated by the phosphorylation state of the CTD, which is correlated with the elongation properties of the polymerase (12).

A number of parallels exist between our observations and the paradigm of rho-dependent termination in *E. coli* (for a review, see reference 49). Rho factor is an RNA-dependent ATPase and helicase that is activated by specific elements in nascent RNA. As a consequence of rho's action, RNA polymerase terminates transcription at sites downstream of rho's initial binding site. Perhaps the pre-mRNA fragments generated in response to the U6R* element are produced by a similar mechanism, given the involvement of the putative RNA-binding factor Nrd1 and the putative helicase Sen1.

Alternatively, binding to the CTD may serve simply to localize Nrd1 in the vicinity of the nascent transcript and may not directly control transcript elongation activity per se. In this case, truncation of U6R*-containing pre-mRNA may occur instead through recruitment of a degradative pathway. The similarity of Sen1 to Upf1, which participates in nonsense-mediated mRNA degradation (34), suggests the possibility that a nuclear pre-mRNA surveillance system responsible for degradation of abnormal transcripts is recruited by the U6R* element. Distinction between mechanisms invoking control of transcript elongation and those involving pre-mRNA degradation may prove difficult to establish experimentally. Indeed, transcription termination may actually be coupled to endonucleolytic cleavage of the transcript (54).

Cellular functions of Nrd1. Disruption of the *NRD1* gene is lethal to yeast cells. Our characterization of spontaneously arising mutant alleles of *NRD1* suggests that the N-terminal 163 amino acid residues of the protein, roughly coinciding with the N-terminal CTD-binding domain (68), are sufficient for viability except at elevated temperatures. Thus, the essential role of Nrd1 may be as an auxiliary factor for RNA polymerase II. However, *cis*-element-dependent control of pre-mRNA accumulation requires functions mediated by the remainder of Nrd1 and, in particular, by the putative RNA binding domain. Mutant alleles of Nrd1 containing premature stop codons or a missense mutation in the RRM decouple the function of the N-terminal domain of Nrd1 from the pre-mRNA effector function. Posttranslational modifications of Nrd1, such as phos-

phorylation, may similarly decouple these functions, potentially providing a mechanism for regulation of gene expression in response to environmental stimuli.

Given the sequence specificity and the efficiency of Nrd1-dependent control of pre-mRNA accumulation, it seems likely that this pathway is exploited by the cell to control gene expression. However, we cannot rule out the possibility that the response is idiosyncratic to the artificial combination of sequences created by the insertion of the U6R* element into the particular reporter gene system we have used. Further study will be needed to identify natural targets of this potential control mechanism. The precise sequence of the 14-nt region identified by suppressor point mutations occurs only once in the yeast genome, on the nontranscribed strand of the U6 RNA gene, *SNR6*. More than 300 genomic sequences match the eight positions at which spontaneous suppressor mutations have been identified, and these genomic sequences occur in noncoding regions as well as on either strand in coding regions. Furthermore, although the 8 nt are clearly important for Nrd1-dependent pre-mRNA underexpression, they are not sufficient. Our studies demonstrate additional requirements for upstream and downstream sequences whose relevant features remain unclear. Therefore, genomic sequence analysis is of limited value in identifying natural targets of the Nrd1 pathway.

Mammalian CTD-binding proteins related to Nrd1 have been proposed to facilitate pre-mRNA splicing, based on the fact that they contain extensive RE/RS domains similar to those found in many metazoan splicing factors and the observation that the CTD may play an important role in splicing in vitro (68). According to this view, the CTD may act as an assembly platform for splicing components. The similar domain organization shared by Nrd1 and these mammalian proteins suggests that they may play similar roles. However, the *nrd1-1* allele has no detectable effect on splicing of ACT-CUP (Fig. 4) or CYH2 (59a) pre-mRNAs under conditions in which the U6R*-dependent pre-mRNA accumulation defect is fully suppressed. Given that this and other viable mutant alleles of Nrd1 are predicted to lack the RE/RS and RRM domains, any Nrd1-mediated role in yeast pre-mRNA splicing may be nonessential or redundant.

Our results provide a unique glimpse into the critical period in pre-mRNA metabolism occurring after transcript initiation and before splicing and suggest the existence of a novel mechanism by which gene expression can be stringently controlled during this stage. A key player in this pathway, the Nrd1 protein, is related to recently identified mammalian proteins, suggesting that this control mechanism may also be conserved. The genetic approach we have initiated holds promise for further characterization of the Nrd pathway for pre-mRNA control.

ACKNOWLEDGMENTS

We are grateful to Cammie Lesser and Christine Guthrie for supplying pGAC24 and yeast *cup1*Δ strain 46α, Michael Kaiser and Phil James for yeast strains, and the Michael Culbertson lab for yeast strains and Sen1 oligonucleotides. We thank Jeff Corden for communication of results prior to publication. Discussions with and encouragement from members of the D.A.B. and J. E. Dahlberg labs contributed to this work.

This investigation has been aided by a Searle Scholar Award from the Chicago Community Trust and a Shaw Scientist Award from the Milwaukee Foundation (to D.A.B.) and a grant from the Jane Coffin Childs Memorial Fund for Medical Research (to E.J.S.).

REFERENCES

1. **Alani, E., S. Subbiah, and N. Kleckner.** 1989. The yeast RAD50 gene encodes a predicted 153-kD protein containing a purine nucleotide-binding domain and two large heptad-repeat regions. Genetics **122:**47–57.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Anderson, J. T., M. R. Paddy, and M. S. Swanson.** 1993. PUB1 is a major nuclear and cytoplasmic polyadenylated RNA-binding protein in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **13:**6102–6113.
4. **Anderson, J. T., S. M. Wilson, K. V. Datar, and M. S. Swanson.** 1993. NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. Mol. Cell. Biol. **13:**2730–2741.
5. **Bandziulis, R. J., M. S. Swanson, and G. Dreyfuss.** 1989. RNA-binding proteins as developmental regulators. Genes Dev. **3:**431–437.
6. **Birney, E., S. Kumar, and A. R. Krainer.** 1993. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res. **21:**5803–5816.
7. **Brow, D. A., and C. Guthrie.** 1988. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. Nature (London) **334:**213–218.
8. **Burd, C. G., and G. Dreyfuss.** 1994. Conserved structures and diversity of functions of RNA-binding proteins. Science **265:**615–621.
9. **Caceres, J. F., S. Stamm, D. M. Helfman, and A. R. Krainer.** 1994. Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science **265:**1706–1709.
10. **Christianson, T. W., R. S. Sikorski, M. Dante, J. H. Shero, and P. Hieter.** 1992. Multifunctional yeast high-copy-number shuttle vectors. Gene **110:** 119–122.
11. **Czaplinski, K., Y. Weng, K. W. Hagan, and S. W. Peltz.** 1995. Purification and characterization of the Upf1 protein: a factor involved in translation and mRNA degradation. RNA **1:**610–623.
12. **Dahmus, M. E.** 1995. Phosphorylation of the C-terminal domain of RNA polymerase II. Biochim. Biophys. Acta **1261:**171–182.
13. **Darnell, J. E., Jr.** 1982. Variety in the level of gene control in eukaryotic cells. Nature (London) **297:**365–371.
14. **Das, A.** 1993. Control of transcription termination by RNA-binding proteins. Annu. Rev. Biochem. **62:**893–930.
15. **DeMarini, D. J., M. Winey, D. Ursic, F. Webb, and M. R. Culbertson.** 1992. SEN1, a positive effector of tRNA-splicing endonuclease in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **12:**2154–2164.
16. **Dreyfuss, G., M. J. Matunis, S. Pinol-Roma, and C. G. Burd.** 1993. hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. **62:**289–321.
17. **Duan, D. R., A. Pause, W. H. Burgess, T. Aso, D. Y. T. Chen, K. P. Garrett, R. C. Conaway, J. W. Conaway, W. M. Linehan, and R. D. Klausner.** 1995. Inhibition of transcription by the VHL tumor suppressor gene. Science **269:**1402–1406.
18. **Elliott, D. J., F. Stutz, A. Lescure, and M. Rosbash.** 1994. mRNA nuclear export. Curr. Opin. Genet. Dev. **4:**305–309.
19. **Frank, D., and C. Guthrie.** 1992. An essential splicing factor, SLU7, mediates 3′ splice site choice in yeast. Genes Dev. **6:**2112–2124.
20. **Fry, D. C., S. A. Kuby, and A. S. Mildvan.** 1986. ATP-binding site of adenylate kinase: mechanistic implications of its homology with *ras*-encoded p21, F1-ATPase, and other nucleotide-binding proteins. Proc. Natl. Acad. Sci. USA **83:**907–911.
21. **Fu, X. D.** 1995. The superfamily of arginine/serine-rich splicing factors. RNA **1:**663–680.
22. **Good, P. J.** 1995. A conserved family of elav-like genes in vertebrates. Proc. Natl. Acad. Sci. USA **92:**4557–4561.
23. **Gorbalenya, A. E., E. V. Koonin, A. P. Donchenko, and V. M. Blinov.** 1989. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. Nucleic Acids Res. **17:**4713–4730.
24. **Greenblatt, J., J. R. Nodwell, and S. W. Mason.** 1993. Transcriptional antitermination. Nature (London) **364:**401–406.
25. **Guthrie, C., and G. R. Fink (ed.).** 1991. Guide to yeast genetics and molecular biology, vol. 194. Academic Press, San Diego, Calif.
26. **Hamer, D. H., D. J. Thiele, and J. E. Lemontt.** 1985. Function and autoregulation of yeast copperthionein. Science **228:**685–690.
27. **Hodgman, T. C.** 1988. A new superfamily of replicative proteins. Nature (London) **333:**22–23.
28. **Hoffman, C. S., and F. Winston.** 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. Gene **57:**267–272.
29. **Kenan, D. J., C. C. Query, and J. D. Keene.** 1991. RNA recognition: towards identifying determinants of specificity. Trends Biochem. Sci. **16:**214–220.
30. **Kibel, A., O. Iliopoulos, J. A. DeCaprio, and W. G. Kaelin.** 1995. Binding of the von Hippel-Lindau tumor suppressor protein to elongin B and C. Science **269:**1444–1446.
31. **Koonin, E. V.** 1992. A new group of putative RNA helicases. Trends Biochem. Sci. **17:**495–497.
32. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154:** 367–382.
33. **Lee, F. J., and J. Moss.** 1993. An RNA-binding protein gene (RBP1) of *Saccharomyces cerevisiae* encodes a putative glucose-repressible protein containing two RNA recognition motifs. J. Biol. Chem. **268:**15080–15087.

34. **Leeds, P., J. M. Wood, B.-S. Lee, and M. R. Culbertson.** 1992. Gene products that promote mRNA turnover in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **12:**2165–2177.
35. **Legrain, P., and M. Rosbash.** 1989. Some cis- and trans-acting mutants for splicing target pre-mRNA to the cytoplasm. Cell **57:**573–583.
36. **Lesser, C. F., and C. Guthrie.** 1993. Mutational analysis of pre-mRNA splicing in *Saccharomyces cerevisiae* using a sensitive new reporter gene, CUP1. Genetics **133:**851–863.
37. **Lesser, C. F., and C. Guthrie.** 1993. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. Science **262:**1982–1988.
38. **Liu, X., and J. E. Mertz.** 1995. HnRNP L binds a cis-acting RNA sequence element that enables intron-independent gene expression. Genes Dev. **9:**1766–1780.
39. **Matunis, M. J., E. L. Matunis, and G. Dreyfuss.** 1993. PUB1: a major yeast poly(A)$^+$ RNA-binding protein. Mol. Cell. Biol. **13:**6114–6123.
40. **Mayeda, A., and A. R. Krainer.** 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell **68:**365–375.
41. **Michael, W. M., M. Choi, and G. Dreyfuss.** 1995. A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. Cell **83:**415–422.
42. **Nagai, K., K. Oubridge, N. Ito, J. Avis, and P. Evans.** 1995. The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. Trends Biochem. Sci. **20:**235–240.
43. **Negishi, Y., Y. Nishita, Y. Saegusa, I. Kakizaki, I. Galli, F. Kihara, K. Tamai, N. Miyajima, S. M. Iguchi-Ariga, and H. Ariga.** 1994. Identification and cDNA cloning of single-stranded DNA binding proteins that interact with the region upstream of the human c-myc gene. Oncogene **9:**1133–1143.
44. **Nevins, J. R.** 1983. The pathway of eukaryotic mRNA formation. Annu. Rev. Biochem. **52:**441–466.
45. **Page, B. D., and M. Snyder.** 1992. CIK1: a developmentally regulated spindle pole body-associated protein important for microtubule functions in *Saccharomyces cerevisiae*. Genes Dev. **6:**1414–1429.
46. **Parker, R., P. G. Siliciano, and C. Guthrie.** 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. Cell **49:**229–239.
47. **Pikielny, C. W., and M. Rosbash.** 1985. mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. Cell **41:**119–126.
48. **Pinol-Roma, S., and G. Dreyfuss.** 1992. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. Nature (London) **355:**730–732.
49. **Platt, T.** 1994. Rho and RNA: models for recognition and response. Mol. Microbiol. **11:**983–990.
50. **Ripmaster, T. L., and J. L. Woolford, Jr.** 1993. A protein containing conserved RNA-recognition motifs is associated with ribosomal subunits in *Saccharomyces cerevisiae*. Nucleic Acids Res. **21:**3211–3216.
51. **Robinow, S., A. R. Campos, K. M. Yao, and K. White.** 1988. The elav gene product of Drosophila, required in neurons, has three RNP consensus motifs. Science **242:**1570–1572.
52. **Rose, M. D., P. Novick, J. H. Thomas, D. Botstein, and G. R. Fink.** 1987. A *Saccharomyces cerevisiae* genomic plasmid bank based on a centromere-containing shuttle vector. Gene **60:**237–243.
53. **Rothstein, R.** 1991. Targeting, disruption, replacement, and allele rescue: integrative DNA transformation in yeast. Methods Enzymol. **194:**281–301.
54. **Russo, P., and F. Sherman.** 1989. Transcription terminates near the poly(A) site in the CYC1 gene of the yeast *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **86:**8348–8352.
55. **Schiestl, R. H., and R. D. Gietz.** 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Genet. **16:**339–346.
56. **Schneiter, R., T. Kadowaki, and A. M. Tartakoff.** 1995. mRNA transport in yeast: time to reinvestigate the functions of the nucleus. Mol. Biol. Cell **6:**357–370.
57. **Sikorski, R. S., and P. Hieter.** 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. Genetics **122:**19–27.
58. **Soulard, M., V. Della Valle, M. C. Siomi, S. Pinol-Roma, P. Codogno, C. Bauvy, M. Bellini, J. C. Lacroix, G. Monod, G. Dreyfuss, et al.** 1993. hnRNP G: sequence and characterization of a glycosylated RNA-binding protein. Nucleic Acids Res. **21:**4210–4217.
59. **Spencer, C. A., and M. Groudine.** 1990. Transcription elongation and eukaryotic gene regulation. Oncogene **5:**777–785.

59a. **Steinmetz, E. J., and D. A. Brow.** Unpublished observations.

60. **Story, R. M., and T. A. Steitz.** 1992. Structure of the recA protein-ADP complex. Nature (London) **355:**374–376.
61. **Sun, Q., A. Mayeda, R. K. Hampson, A. R. Krainer, and F. M. Rottman.** 1993. General splicing factor SF2/ASF promotes alternative splicing by binding to an exonic splicing enhancer. Genes Dev. **7:**2598–2608.
62. **Tian, Q., M. Streuli, H. Saito, S. F. Schlossman, and P. Anderson.** 1991. A polyadenylate binding protein localized to the granules of cytolytic lymphocytes induces DNA fragmentation in target cells. Cell **67:**629–639.
63. **Treco, D. A.** 1989. *Saccharomyces cerevisiae*, p. 13.12.11–13.12.13. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. Greene Publishing Associates and Wiley Interscience, New York.
64. **Ursic, D., D. J. DeMarini, and M. R. Culbertson.** 1995. Inactivation of the yeast Sen1 protein affects the localization of nucleolar proteins. Mol. Gen. Genet. **249:**571–584.
65. **Vijayraghavan, U., R. Parker, J. Tamm, Y. Iimura, J. Rossi, J. Abelson, and C. Guthrie.** 1986. Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome. EMBO J. **5:**1683–1695.
66. **Wilson, S. M., K. V. Datar, M. R. Paddy, J. R. Swedlow, and M. S. Swanson.** 1994. Characterization of nuclear polyadenylated RNA-binding proteins in *Saccharomyces cerevisiae*. J. Cell Biol. **127:**1173–1184.
67. **Yang, X., M. R. Bani, S. J. Lu, S. Rowan, Y. Ben-David, and B. Chabot.** 1994. The A1 and A1B proteins of heterogeneous nuclear ribonucleoparticles modulate 5′ splice site selection in vivo. Proc. Natl. Acad. Sci. USA **91:**6924–6928.
68. **Yuryev, A., M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gentile, M. Gebara, and J. L. Corden.** 1996. The CTD of RNA polymerase II interacts with a novel set of SR-like proteins. Proc. Natl. Acad. Sci. USA **93:**6975–6980.