MOLECULAR AND CELLULAR BIOLOGY, May 1993, p. 2730–2741
0270-7306/93/052730-12$02.00/0

Vol. 13, No. 5

NAB2: a Yeast Nuclear Polyadenylated RNA-Binding Protein Essential for Cell Viability

JAMES T. ANDERSON, SCOTT M. WILSON, KSHAMA V. DATAR, AND MAURICE S. SWANSON*

Department of Immunology and Medical Microbiology, College of Medicine, University of Florida, Gainesville, Florida 32610-0266

Received 22 December 1992/Returned for modification 1 February 1993/Accepted 5 February 1993

A variety of nuclear ribonucleoproteins are believed to associate directly with nascent RNA polymerase II transcripts and remain associated during subsequent nuclear RNA processing reactions, including pre-mRNA polyadenylation and splicing as well as nucleocytoplasmic mRNA transport. To investigate the functions of these proteins by using a combined biochemical and genetic approach, we have isolated nuclear polyadenylated RNA-binding (NAB) proteins from *Saccharomyces cerevisiae*. Living yeast cells were irradiated with UV light to covalently cross-link proteins intimately associated with RNA in vivo. Polyadenylated RNAs were then selectively purified, and the covalent RNA-protein complexes were used to elicit antibodies in mice. Both monoclonal and polyclonal antibodies which detect a variety of NAB proteins were prepared. Here we characterize one of these proteins, NAB2. NAB2 is one of the major proteins associated with nuclear polyadenylated RNA in vivo, as detected by UV light-induced cross-linking. Cellular immunofluorescence, using both monoclonal and polyclonal antibodies, demonstrates that the NAB2 protein is localized within the nucleus. The deduced primary structure of NAB2 indicates that it is composed of at least two distinct types of RNA-binding motifs: (i) an RGG box recently described in a variety of heterogeneous nuclear RNA-, pre-rRNA-, mRNA-, and small nucleolar RNA-binding proteins and (ii) CCCH motif repeats related to the zinc-binding motifs of the largest subunit of RNA polymerases I, II, and III. In vitro RNA homopolymer/single-stranded DNA binding studies indicate that although both the RGG box and CCCH motifs bind poly(G), poly(U), and single-stranded DNA, the CCCH motifs also bind to poly(A). *NAB2* is located on chromosome VII within a cluster of ribonucleoprotein genes, and its expression is essential for cell growth.

Gene expression is regulated at multiple levels following the initiation of transcription by RNA polymerase II. Nascent pre-mRNA transcripts are extensively modified to generate mRNAs (59). These modifications, all of which occur in the nucleus, may include 5′-end capping (46, 52), 3′-end cleavage and polyadenylation (64), splicing (27, 53), and modification of individual bases (16). All of these activities occur while pre-mRNAs are associated with a set of nuclear factors which may include both heterogeneous nuclear ribonucleoproteins (hnRNPs) and small nuclear ribonucleoproteins (snRNPs) (21, 23, 30, 61). Once formed in the nucleus, mRNAs are transported into the cytoplasm by an unknown mechanism (40). mRNA turnover in both the nucleus and cytoplasm may also dramatically influence the amount of each type of mRNA available for translation (18).

Although considerable information exists concerning the functional role of snRNPs in nuclear RNA processing (27, 30, 53, 61), very little is known about the specific functions of individual hnRNPs. Previous work focused on the idea that hnRNPs are primarily pre-mRNA-packaging proteins functionally analogous to the role that histones play in chromatin organization (17). More recently, it has been proposed that hnRNPs package pre-mRNAs in a transcript-specific manner so that the nuclear machinery required for specific pre-mRNA-processing events, such as snRNPs, recognizes each transcript as a unique entity (9, 10, 63). This hypothesis would help explain why many of these proteins bind RNA in a sequence-specific fashion (62). The use of alternative splicing and polyadenylation sites (59, 64) and the possible nucleocytoplasmic transport of different classes of mRNAs via specific nuclear pore complexes (38, 40) might begin with this transcript-specific packaging by proteins of the hnRNP complex. Experimental support for this idea is beginning to emerge. Analysis of the proteins associated with isolated spliceosomes prepared in vitro indicates that different sets of hnRNPs bind to different types of pre-mRNAs as well as to different regions within these pre-mRNAs which are functionally important for splicing (10). Moreover, in vitro studies on the mechanism of alternative pre-mRNA splicing have shown that the relative concentrations of the hnRNP A1 protein and splicing factor SF2 can influence splice site choice (43). Collectively, these results suggest that individual hnRNPs may bind to discrete domains within a pre-mRNA and compete with other RNA-processing factors to induce or repress utilization of a particular splice site. Finally, recent work demonstrates that the hnRNP A1, but not C, protein shuttles between the nucleus and cytoplasm (47). A subset of hnRNPs may therefore be involved in nucleocytoplasmic mRNA trafficking.

Saccharomyces cerevisiae offers several unique advantages for investigation of the function of proteins which directly associate with nuclear polyadenylated RNAs in vivo. First, only a few protein-encoding genes in this yeast produce pre-mRNAs which are spliced, and there are presently no examples of alternative pre-mRNA splicing (53). In higher eukaryotes, both constitutive and alternative pre-mRNA splicing are extensively used to modify primary RNA polymerase II transcripts (27, 41, 59), and therefore other fundamental pathways of RNA processing, such as pre-mRNA packaging and nucleocytoplasmic mRNA trafficking, may be experimentally masked. Second, once candidate nuclear pre-mRNA/mRNA-binding proteins are isolated,

* Corresponding author.

TABLE 1. Oligonucleotides and plasmids

Designation	Sequence or description
Oligonucleotides	
MSS44	5′-TAATACGACTCACTATAGGGAGA-3′
MSS45	5′-CATACGATTTAGGTGACACTATAG-3′
MSS47	5′-CCCGGATCCCGGTACAGCGGATAGCGC-3′
MSS48	5′-CCCGAATTCGATCATAAGGAAGTGGAAG-3′
MSS49	5′-CCCGAATTCATGTCTCAAGAACAGTAC-3′
MSS51	5′-CCCGGATCCCGCCCCTCTTTCTTGG-3′
MSS52	5′-CCCGAATTCGCAATGCAGACAGATG-3′
MSS53	5′-CCCGAATTCGGAATGGCGGGTGAG-3′
MSS54	5′-CCCGGATCCGGCAATGCAAATGGCC-3′
MSS58	5′-CCCGGTACCAAAATGAAAATCGAAGAAGGTAA-3′
Plasmids[a]	
pRNP327	Genomic 4.7-kb fragment (Fig. 3) subcloned into *Eco*RI-cut pSP72 (Promega)
pNAB2-3B6	*NAB2* coding sequence, generated by PCR with MSS47 and MSS48 primers and inserted into pSP72 cut with *Eco*RI-*Bam*HI (includes nt −29 to 1831); encodes the NAB2* protein
pNAB2.0	pNAB-3B6 cut with *Nar*I (includes nt −29 to 550)
pNAB2.2	NAB2 minus the last 20 amino acids (nt −29 to 1417), produced by PCR with MSS48 and MSS54 and cloned into *Eco*RI-*Bam*HI-cut pSP72
pNAB2.3	Partial *NAB2* coding region (nt −29 to 696), produced by PCR with MSS48 and MSS51 and cloned into *Eco*RI/*Bam*HI-cut pSP72
pNAB2ex1	Entire *NAB2* coding region fused in frame to pMAL-c2 (New England Biolabs), generated by PCR with primers MSS47 and MSS49 (includes nt 1 to 1831)
pNAB2ex2	NAB2 RGG box (nt 423 to 696), produced by PCR with MSS51 and MSS52 and cloned in frame into *Eco*RI-*Bam*HI-cut pMAL-c2
pNAB2ex3	NAB2 C_3H motifs (nt 635 to 1417), produced by PCR with MSS53 and MSS54 and cloned in frame into *Eco*RI-*Bam*HI-cut pMAL-c2
pNAB2ex4	NAB2 RGG box and C_3H motifs (nt 423 to 1417), produced by PCR with MSS52 and MSS54 and cloned into *Eco*RI-*Bam*HI-cut pMAL-c2
pMAL-c2ex	*Kpn*I-*Xmn*I fragment of PCR product of pNAB2ex2 (using MSS51 and MSS58) cloned into *Kpn*I-*Pvu*II-cut pSP72
pNAB2.4	*Kpn*I-*Bam*HI fragment of PCR product of pNAB2ex2 (using MSS51 and MSS58) cloned into *Kpn*I-*Bam*HI-cut pSP72
pNAB2.5	As pNAB2.4 except using pNAB2ex3, MSS54, MSS58
pNAB2.6	As pNAB2.4 except using pNAB2ex4, MSS54, MSS58
pNAB2.8	cDNA clone (1,758 nt) isolated from a λZAP BJ926 expression library in pBS (Stratagene) encoding the NAB2 protein
pNAB2.14	*Eco*RI insert from pRNP327 cut with *Sty*I, blunt ended, and cut with *Nsi*I, with the 2.2-kb fragment cloned into *Eco*RV-*Pst*I-cut pSP72. The 1.2-kb *Eco*RV-*Bal*I fragment within the *NAB2* ORF was replaced with a blunt-ended 2.2-kb *Xho*I-*Sal*I fragment containing the *LEU2* gene
pNAB2.15	*Pvu*I-*Nsi*I fragment (nt −496 to 1961), blunt ended and cloned into the *Sma*I site of pRS316 (58)

[a] Nucleotide numbers refer to Fig. 5. Genomic DNA (BJ926) was used for all PCR amplifications unless otherwise stated.

this yeast provides an excellent genetic system with which to investigate the functions of individual proteins in vivo. This last attribute is key since the pathway of nuclear RNA processing may be extremely dynamic, involving the constant association and dissociation of multiple proteins and RNP complexes with pre-mRNAs or mRNAs, making it difficult to dissect the functions of individual proteins biochemically in vitro.

We have therefore investigated whether hnRNPs exist in *S. cerevisiae*. Our approach to the isolation of this group of RNPs from yeast cells is dependent on the observation that many hnRNPs from higher eukaryotes are readily and selectively cross-linked in vivo by UV light to nuclear polyadenylated RNAs. Here we report the isolation of a novel group of yeast proteins which are intimately associated with nuclear polyadenylated RNAs, the nuclear polyadenylated RNA-binding (NAB) proteins. In addition, we characterize one of these proteins, NAB2, whose expression is essential for cell growth.

MATERIALS AND METHODS

Yeast strains and genetic manipulations. Cell labeling, immunofluorescence, and isolation of genomic DNA and poly(A)$^+$ RNA were done by using BJ926 (*MAT***a**/α *prb1-1122/prb1-1122 prc1-407/prc1-407 pep4-3/pep4-3 can1/can1 gal2/gal2 his1/HIS1 TRP1/trp1*) (Yeast Genetic Stock Center). The yeast strain used for cross-linking was YJA504 (*MAT***a**/α *leu2Δ2/leu2Δ2 ura3-52/ura3-52 PUB1::LEU2/PUB1::URA3*) (6). The *NAB2* gene transplacement was performed by using YJA501 (*MAT***a**/α *leu2Δ2/leu2Δ2 ura3-52/ura3-52*) (6) to generate YJA511 (*MAT***a**/α *NAB2/NAB2::LEU2 leu2Δ2/leu2Δ2 ura3-52/ura3-52*). The *NAB2* deletion was complemented by transforming YJA511 with pNAB2.15 to generate YJA512 (*MAT***a**/α *NAB2/NAB2::LEU2 leu2Δ2/leu2Δ2 ura3-52/ura3-52* p*NAB2URA3CEN*).

Yeast media, including YPD, synthetic minimal SD, and amino acid supplements, were prepared according to Rose et al. (51). All genetic manipulations and dissection of asci for segregation analysis were performed by using standard procedures (29).

Oligonucleotides and plasmids. Table 1 lists all of the oligonucleotides and plasmids used in this study.

Preparation of UV cross-linked yeast RNPs. A modification of a previously described procedure (1) was followed. For each preparation, 2 liters of YJA504 cells was grown to an optical density at 600 nm (OD_{600}) of 1.0 to 2.0, pelleted by

centrifugation at 4,500 × *g* for 10 min in a JA-10 rotor (Beckman) at 4°C, resuspended in phosphate-buffered saline (PBS), and repelleted as described above. Yeast cells were resuspended in PBS to a cell density of 1×10^9 to 5×10^9 cells per ml (approximately 200 ml); 50-ml aliquots were placed in a 15-cm petri dish floating in an ice-water bath and irradiated with UV light at a distance of 10 cm from the light source (UV Stratalinker 1800; Stratagene) twice for 2.5 min each time, with gentle agitation between each 2.5-min irradiation period. Cross-linked cells were then pelleted by centrifugation at 4,500 × *g* for 10 min in a swinging bucket rotor (Jouan) at 4°C, resuspended in lysis buffer (20 mM Tris-Cl [pH 7.4]–1 mM EDTA–50 mM LiCl–1% sodium dodecyl sulfate [SDS]–1% β-mercaptoethanol–1 mg of heparin per ml–10 mM vanadyl adenosine–0.5 mM phenylmethylsulfonyl fluoride–5 μg of pepstatin A per ml–1 μg of chymostatin per ml–1 mM *p*-aminobenzamidine–1 mM ε-aminocaproic acid–1 μg of leupeptin per ml–2 μg of aprotinin per ml) at room temperature, and passed once through a French press (20,000 lb/in^2). Insoluble material was pelleted for 10 min at 4,500 × *g* (Jouan), and the supernatant was batch bound to oligo(dT)-Sepharose (GIBCO/BRL) for 30 min at room temperature. Poly(A)$^+$ cross-linked RNPs were reselected on oligo(dT), concentrated by using *sec*-butanol, and precipitated with 3 volumes of ethanol in the presence of 200 mM LiCl. Each preparation (2 liters of starting culture) yielded 6 to 8 OD_{260} units of cross-linked RNPs.

Polyclonal antisera, monoclonal antibodies, and isolation of genomic and cDNA clones. Polyclonal antisera against yeast polyadenylated RNPs were raised by immunizing two BALB/c mice with three injections (2 weeks between injections) of approximately 20 OD_{260} units of cross-linked RNPs per injection, using MPL/TDM adjuvant (RIBI Immunochemical Research). Monoclonal antibodies were prepared by using SP2/O myeloma cells, and hybridomas were screened by immunoblotting and cellular immunofluorescence (22). Polyclonal antisera were also prepared against a fusion protein containing the maltose-binding protein (MBP) and the entire *NAB2* open reading frame (ORF). The MBP-NAB2 fusion protein was produced in *Escherichia coli* TB1 cells from an expression plasmid, pNAB2ex1, constructed by using pMAL-c2 (New England Biolabs) (see plasmid constructs discussed above). Mice were injected with the fusion protein three times, using 5 μg per injection.

The 1G1 monoclonal antibody reactive against the polyadenylate-binding protein (PAB1) was prepared by using an MBP-PAB1 fusion protein containing the entire ORF of *PAB1*. The MBP-PAB1 fusion protein was synthesized by using a vector constructed by cloning the *PAB1* gene (1) in frame into pMAL-c2. Polyclonal antisera and monoclonal antibodies were prepared as described above.

The *NAB2* gene was isolated from a λgt11 genomic expression library as previously described (1, 60), using a 1:300 dilution of the polyclonal antisera elicited against the cross-linked RNPs described above. *NAB2* cDNA clones were isolated from a λZAP cDNA expression library, which was prepared according to the instructions of the manufacturer (Stratagene), using 5 μg of poly(A)$^+$ RNA isolated from BJ926 cells. The cDNA library was screened by using a *Pst*I-*Sca*I fragment (859 nucleotides [nt]; see Fig. 3) encompassing ~58% of *NAB2*.

Cell labeling, subcellular fractionation, and immunopurification. BJ926 yeast cells were labeled with TRAN^{35}S (ICN) for 2 h as previously described (1) and disrupted by vortexing with glass beads (29). Highly purified nuclei (7, 8) were isolated as described previously. Immunopurification of NAB2 was performed by using labeled yeast extract preclarified by centrifugation at maximum speed (14,300 × *g*) in a microcentrifuge. Briefly, 3 μl of anti-NAB2 fusion protein antiserum (prepared by using pNAB2ex1; Table 1) was bound to protein A-Sepharose (Pharmacia) in PBS containing 1 mM EDTA, 1% Triton X-100, 0.5% deoxycholic acid, 0.1% SDS, 0.5% aprotinin, and 0.02% sodium azide for 60 min at 4°C. The antibody-coupled Sepharose beads were then incubated with labeled cell extract for 15 min at 4°C, and bound proteins were purified by being washed five times with the buffer described above and then briefly centrifuged.

Gel electrophoresis and immunoblotting. Proteins were resolved by SDS-polyacrylamide gel electrophoresis (PAGE), using 12.5% (final acrylamide concentration) separation gels (22). Coomassie blue-stained gels were impregnated with 2,5-diphenyloxazole for visualization of ^{35}S-labeled proteins. Following electrophoresis, proteins were transferred to nitrocellulose by using a semidry electroblotter (Hoefer) for 1 h at 100 mA. For immunoblotting, all antibody preparations were diluted in 5% milk–PBS (anti-cross-linked RNP antiserum, 1:1,000; anti-NAB2 antiserum, 1:1,000; D77 monoclonal antibody to NOP1, 1:500; 1G1 monoclonal antibody to PAB1, 1:1,000; goat anti-mouse horseradish peroxidase-conjugated secondary antibody, 1:5,000), and reactive antigens were visualized by enhanced chemoluminescence (Amersham), using film exposures of 5 to 30 s.

DNA and RNA blot analysis. High-molecular-weight yeast genomic DNAs were isolated and fractionated in a 0.8% agarose gel, and Southern blot transfer and hybridization were performed as described previously (14, 51). Polyadenylated RNAs were prepared from BJ926, fractionated on a 1.4% formaldehyde–agarose gel, and transferred to nitrocellulose. Hybridization was performed under high-stringency conditions as described previously (1, 14).

In vitro transcription/translation and RNA/ssDNA binding studies. In vitro transcriptions/translations of pSP72-based expression plasmids were performed by using T7 RNA polymerase and rabbit reticulocyte lysates as instructed by the manufacturer (Promega). DNA sequence analysis of the expression plasmids confirmed the correct sequences between *NAB2* and pMAL-c2. In vitro binding reactions of full-length protein and various subdomains of the NAB2 protein to both RNA homopolymers and single-stranded DNAs (ssDNAs) were performed essentially as described previously (62) except that 2 μl of each in vitro translation mixture was used per binding reaction and the heparin wash was omitted.

Cellular immunofluorescence. Immunofluorescence of yeast cells was performed by using a modification of previously published procedures (8, 20). BJ926 cells were grown to an OD_{600} of 0.1 to 0.2 in YPD, and 10% formaldehyde (freshly prepared from paraformaldehyde in water) was directly added to a final concentration of 4%. Following incubation at room temperature for 2 h with gentle agitation on a Vari-Mix (Thermolyne), fixed cells were pelleted at 1,500 × *g* (Jouan) for 5 min, washed twice with 25 ml of 100 mM KH_2PO_4 (pH 7.5) and once with 25 ml of 100 mM KH_2PO_4 (pH 7.5)–1.2 M sorbitol, and finally resuspended in 1 ml of the latter buffer including 40 μl of yeast lytic enzyme (10 mg/ml; 100,000 U/g; ICN). Cell wall digestion was carefully monitored by phase-contrast microscopy (magnification of 400×) and required 10 to 30 min for >90% spheroplast formation. Spheroplasts were centrifuged at 2,940 × *g* (6,000 rpm) in an Eppendorf Microfuge, washed

once with 1 ml of 100 mM KH_2PO_4–1 M sorbitol, and resuspended in 500 μl of the same buffer; 10 μl per well was allowed to attach to polylysine-coated 10-well HTC Blue slides (Cell-Line Associates) for 15 min at 4°C. Slides were washed with ice-cold PBS and sequentially incubated at −20°C in 100% methanol for 5 min and 100% acetone for 30 s, washed three times with PBS, incubated with 0.1% Triton X-100–PBS, washed three times in PBS, and finally stored at 4°C in PBS–0.05% sodium azide. Slides were generally used immediately but could be stored for at least 1 week without significant loss of cellular morphology. Polyclonal antisera (1:1,000), 1G1 (1:5,000), A66 (1:1,000), and fluorescein-isothiocyanate-conjugated goat anti-mouse secondary antibodies (1:2,000; Cappel) were diluted as indicated in 3% bovine serum albumin (BSA)–PBS. Cells in each slide well were sequentially incubated with 10 μl of 3% BSA–PBS (preblock) for 30 min, primary antibody for 60 min, and secondary antibodies for 30 min at room temperature. Secondary antibodies were pretreated with heat-killed BJ926 cells for 60 min at 4°C and were cleared by centrifugation at 14,300 × *g* for 5 min in a Microfuge prior to use. Slides were washed three times with PBS, treated with 0.5 μg of 4′,6-diamidino-2-phenylindole (Sigma) per ml, washed three more times with PBS, and mounted. A Nikon Optiphot-2 microscope equipped with an ×100 fluorescence/differential interference contrast objective was used for all cellular immunofluorescence studies.

DNA sequencing and analysis. DNA sequences were determined for both strands (57) by using restriction fragments cloned into either M13mp18 and M13mp19 (44), Bluescript SK− (Stratagene) (for the cDNA clones), or pSP72 and SP6/T7 sequencing primers. The University of Wisconsin Genetics Computer Group programs were used to analyze all sequence information. Data base searches were performed by using the BLAST network service at the National Center for Biotechnology Information (3).

Nucleotide sequence accession numbers. The nucleotide sequences of both *NAB2* genes have been listed in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under accession numbers L08079 (encoding the NAB2* protein) and L10288 (encoding the NAB2 protein).

RESULTS

Isolation and characterization of yeast NAB proteins. A previous study described several proteins which are readily cross-linked by UV light to polyadenylated RNA within yeast cells (1). One of these proteins is the polyadenylate tail-binding protein PAB1, which has been extensively characterized by both biochemical (1, 13, 54) and genetic (55) techniques. We have also recently isolated the other major protein cross-linked by UV light to polyadenylated RNA in vivo, PUB1 (6). Both PAB1 and PUB1 are predominantly, but not exclusively, associated with cytoplasmic mRNA (1, 6). While *PAB1* is an essential gene in yeast cells, YJA504 bearing a homozygous deletion of *PUB1* grows normally (1, 6, 54). Cytoplasmic mRNA accounts for >90% of the polyadenylated RNA in yeast cells (28), and thus it is not surprising that mice injected with cross-linked poly(A)$^+$ RNA-RNP complexes produce antibodies which primarily recognize PAB1 and PUB1 (1, 6). To avoid eliciting antibodies to PUB1, YJA504 was selected for the preparation of cross-linked poly(A)$^+$ RNA-RNP complexes. These complexes were then injected into mice for the preparation of polyclonal antibodies.

Total yeast cell protein was immunoblotted with the antiserum from one of the mice injected with cross-linked poly(A)$^+$ RNA-RNP complexes (Fig. 1A); the antiserum from the other mouse yielded nearly identical results. Several polypeptides ranging from 45 to 120 kDa were recognized by the antiserum. One of the proteins migrating at ~72 kDa was identified as PAB1 by using a monoclonal antibody, 1G1 (see Materials and Methods). Surprisingly, when the distribution of these cross-linked antigens within the cell was examined by using this mouse antiserum, a predominant nuclear fluorescence was obtained (Fig. 2C).

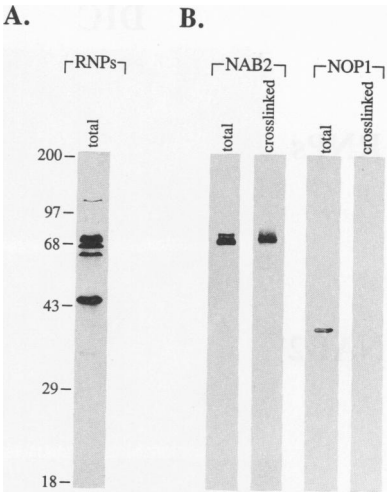

FIG. 1. Yeast cell proteins cross-linked by UV light to polyadenylated RNA in vivo. (A) Immunoblot of total cell extract from BJ926 (lane total) probed with polyclonal antibodies elicited against purified poly(A)$^+$ RNA-RNP cross-linked complexes. Sizes are indicated in kilodaltons. (B) Immunoblot analysis of either total BJ926 extract (lanes total) or purified poly(A)$^+$ RNA-RNP cross-linked complexes (lanes crosslinked) probed with polyclonal antisera raised against the purified MBP-NAB2 fusion protein (panel NAB2) or a monoclonal antibody against the yeast pre-rRNA-binding protein fibrillarin (panel NOP1).

As described below, the mouse antiserum was used to screen a yeast λgt11 genomic expression library to isolate and identify the other cross-linked proteins. Once individual genes had been identified by DNA sequencing and RNA blot analysis, they were subcloned into an MBP-expression vector. Fusion proteins produced in *E. coli* were used to elicit polyclonal antibodies. Two criteria were used to identify nuclear polyadenylated RNA-binding proteins. First, antibodies to individual proteins had to recognize an antigen which was efficiently cross-linked to poly(A)$^+$ RNA. Second, these antibodies had to localize their respective antigens to the nucleus by cellular immunofluorescence. Antibodies to one of these antigens, which was termed NAB2, recognized a protein of ~72 kDa which was efficiently cross-linked to poly(A)$^+$ RNA, whereas a nucleolar pre-rRNA-binding protein, NOP1, was not (Fig. 1B). Cross-linking of NAB2 to poly(A)$^+$ RNA resulted in a slight increase in its apparent size as a result of cross-linked ribonucleotides (Fig. 1B; compare total versus cross-linked). This antiserum also recognized a minor polypeptide which migrated just above the major NAB2 protein (Fig. 1B, panel NAB2, lane total). As shown below, NAB2 often migrated as a doublet even when the protein was produced by in vitro transcription/translation of the cloned *NAB2* gene. NAB2 was localized to the nucleus, and anti-NAB2 antibodies gave

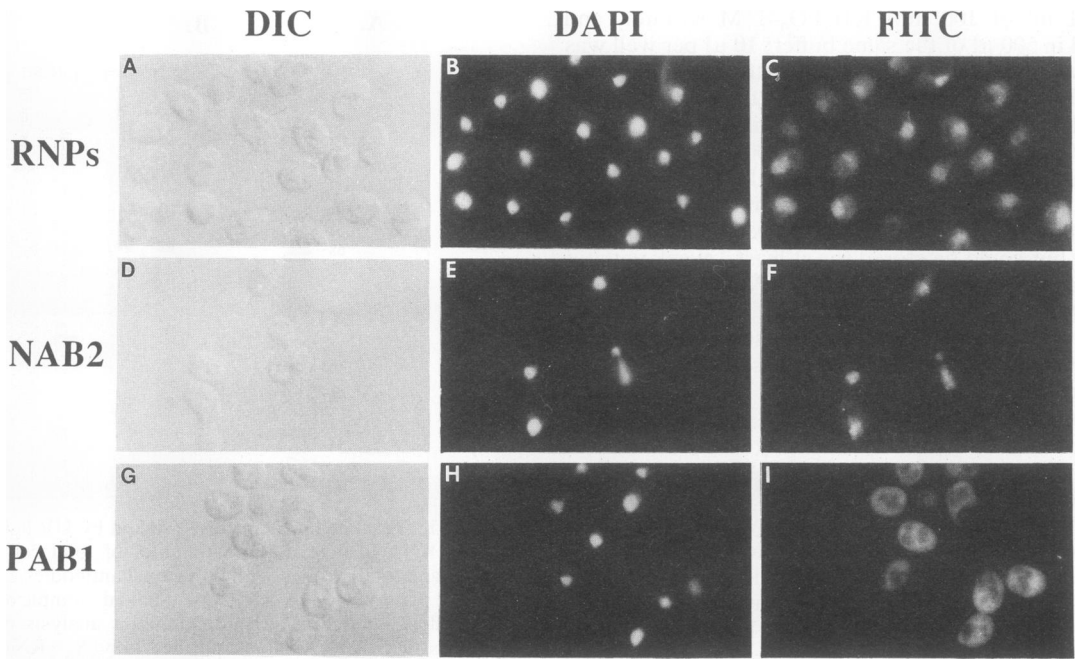

FIG. 2. Localization of the NAB2 protein to the nucleus. Indirect immunofluorescence was used to study the subcellular distribution of either cross-linked poly(A)$^+$ RNA-RNP complexes (A to C), the NAB2 protein (D to F), or the PAB1 protein (G to I) in BJ926 cells. Cells were visualized by differential interference contrast microscopy (DIC), DNA distribution was visualized by 4′,6-diamidino-2-phenylindole (DAPI) staining, and the various antigens were localized by using a fluorescein isothiocyanate (FITC)-conjugated secondary antibody. The immunofluorescence patterns seen with BJ926 are indistinguishable from those obtained with the PUB1-deleted strain YJA504 (6).

a strong overall nuclear staining pattern very similar to the diffuse nucleoplasmic staining seen in human cells with use of anti-hnRNP antibodies (15, 21) (Fig. 2D to F). This result is in contrast to the immunofluorescence patterns seen for various nucleolar antigens, including NOP1 (8) and NSR1 (39), which are predominantly subnuclear and crescent shaped, as well as PAB1 (Fig. 2G to I), since the 1G1 monoclonal antibody gives an intense cytoplasmic, and a significantly weaker nuclear, immunofluorescence signal. Biochemical fractionation and immunoblot studies of highly purified nuclei (7, 8) also confirmed that the majority of NAB2 is detectable in the nucleoplasm fraction.

***NAB2* resides within a cluster of three RNP genes on chromosome VII.** As described above, the *NAB2* gene was isolated from a λgt11 genomic expression library by using the antiserum from a mouse injected with cross-linked poly(A)$^+$ RNA-RNP complexes. Full-length genomic clones, identified by hybridization, were subsequently isolated. Figure 3 depicts a 4.7-kb genomic clone which contains the complete *NAB2* gene. This particular genomic clone was isolated separately as an expression clone of *CAB1*, which encodes another RNP that was cross-linked to polyadenylated RNA in vivo (19). The *NAB2* gene is situated on chromosome VII and resides ~1 kb from *CAB1* and ~0.6 kb from *SUP44*, which encodes the S4 protein of the 40S ribosomal subunit (2). Several other, smaller ORFs, which begin with methionine codons situated within an appropriate initiation sequence context, exist on this genomic fragment. However, the deduced amino acid sequences of these ORFs do not possess any homology to known proteins or structural motifs, and we presently do not have any evidence that they are transcribed. Genomic blot analysis indicated the *NAB2* gene is present in single copy in the yeast genome (Fig. 4A).

A major poly(A)$^+$ RNA of 1.65 kb was detected by RNA blot analysis using a *NAB2* hybridization probe (Fig. 4B). Although longer exposures revealed two low-abundance, or weakly hybridizing, RNAs of 1.9 and 2.5 kb, no consensus signals for splicing (27, 30, 53), such as the TACTAAC branchpoint sequence, exist within 1 kb of the *NAB2* gene. It is therefore highly unlikely that the *NAB2* gene is spliced (30, 53). The infrequent use of alternative polyadenylation sites, or cross-hybridizing RNAs produced from related genes, could account for the observed larger poly(A)$^+$ RNAs.

We have also isolated two anti-NAB2 monoclonal anti-

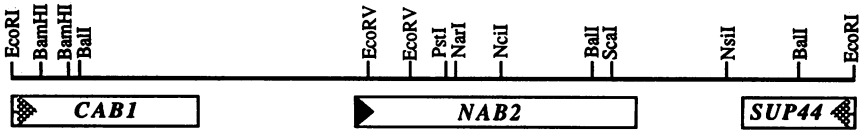

FIG. 3. Localization of the *NAB2* gene on chromosome VII between two other RNP genes. Shown is a restriction map of a 4.7-kb genomic fragment (pRNP327) which contains *NAB2*, *CAB1*, and *SUP44* (ribosomal small-subunit S4 gene). The boxes representing the three genes are drawn to scale, and the arrows indicate the direction of transcription. Stippled arrows indicate that only part of the ORF is encoded on this genomic fragment.

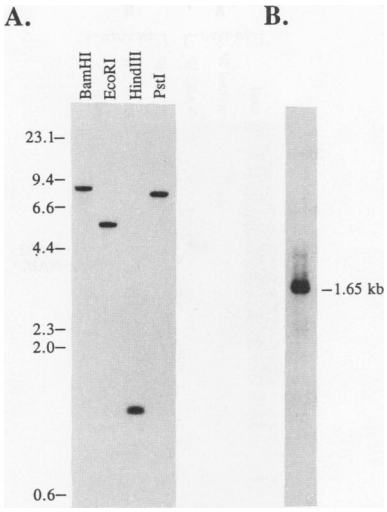

FIG. 4. Genomic DNA and RNA blot analyses of NAB2. (A) Genomic DNA (BJ926) was isolated, digested with the indicated restriction enzymes, and probed with a 1,470-bp *Pst*I-*Nsi*I restriction fragment from pRNP327 (see Fig. 3). Sizes are indicated in kilobases. (B) Polyadenylated RNA was isolated, fractionated on a 1.4% formaldehyde–agarose gel, and probed with a 780-bp *Pst*I-*Bal*I restriction fragment from pRNP327.

bodies, 1F2 and 3F2, which were prepared by using the MBP-NAB2 fusion protein described above. In agreement with the results obtained with polyclonal antisera, these monoclonal antibodies localize NAB2 to the nucleus and demonstrate that this protein is strongly cross-linked to polyadenylated RNA in vivo.

NAB2 is structurally related to other known RNA-binding proteins. The amino acid sequence deduced from the *NAB2* gene (Fig. 5A) revealed a basic protein (pI = 8.01) of 497 amino acids with a calculated molecular weight of 54,950. The apparent molecular size of NAB2 by SDS-PAGE was 71 to 72 kDa when the protein was synthesized in vivo (Fig. 1B). Since a variety of human hnRNPs, including the C1/C2 (13), K/J (42), L (48), and U (37) proteins, display anomalous electrophoretic migration, aberrant SDS-PAGE mobility could have accounted for all of the difference between the calculated and apparent sizes for the NAB2 protein. However, as described below, the *NAB2* gene sequence varies between yeast strains, and this variation is primarily responsible for the observed differences in molecular weights.

As illustrated in Fig. 5A and 6A, the NAB2 protein consists of at least three interesting protein structural motifs. The first is a 44-amino-acid glutamine-rich region (101 to 144) related to glutamine-rich regions found in a wide array of eukaryotic transcriptional factors as well as many RNA-binding proteins, including PAN1 (56), a PAB1-dependent nuclease of *S. cerevisiae*, the *Drosophila elav* protein involved in neurogenesis (50), and the human hnRNP U protein (37). The second NAB2 motif is an arginine/glycine-rich region of 20 amino acids (181 to 200) that is related to the RGG box, which is conserved among a number of RNA-binding proteins and has been shown to be an RNA-binding motif (37). Figure 6B compares the RGG boxes of the known yeast proteins which possess this motif. Unlike NAB2, they are all thought to be localized to the nucleolus and involved in pre-rRNA processing. In this respect, it is interesting to note that the sequence of NAB2 varies somewhat from the

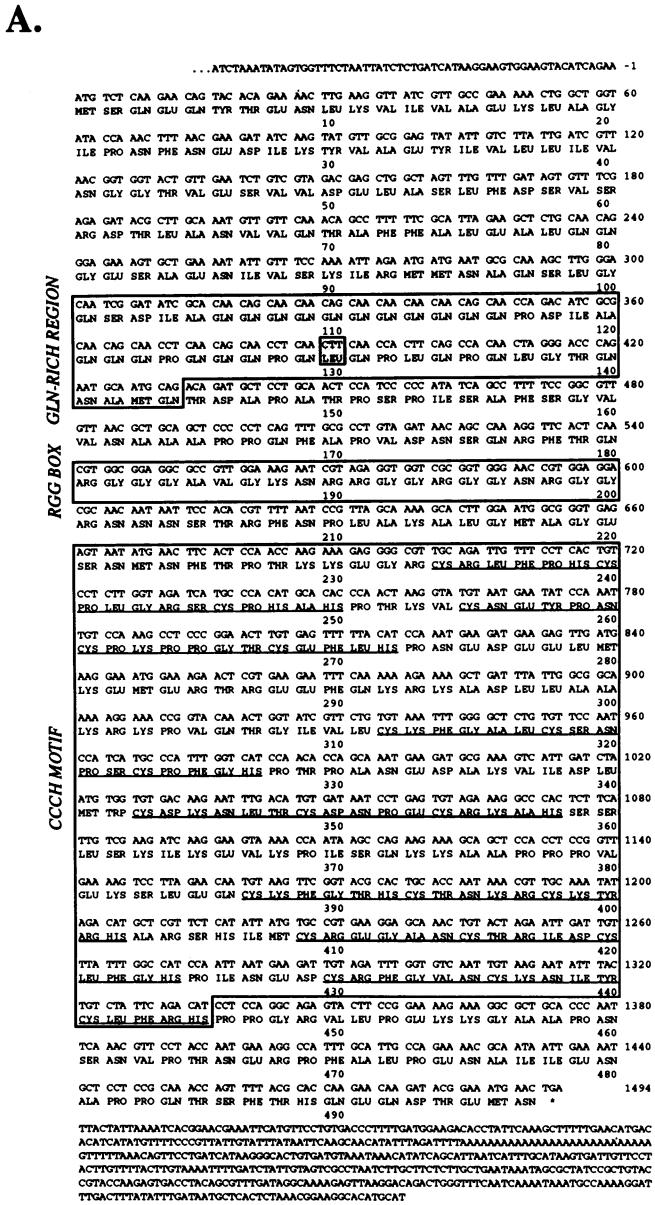

FIG. 5. Nucleotide and deduced amino acid sequences of *NAB2*. (A) Genomic DNA sequence from pNAB327, including nt −59 to 1961 (*Nsi*I site; see Fig. 3). Protein structural features are highlighted by open boxes, and the C_3H motif repeats are underlined. The shaded box indicates the leucine residue which is interrupted by the 84-nt insert after the first nucleotide of the codon (C ↓ TT). (B) Sequence of the 84-nt insert in BJ926 cDNA and genomic clones.

consensus, with the second RGG replaced with KNR and variations in the spacing between the RGG tripeptides. Also shown are the RGG boxes of the human hnRNP A1 and U proteins.

Finally, nearly half of the NAB2 primary structure is

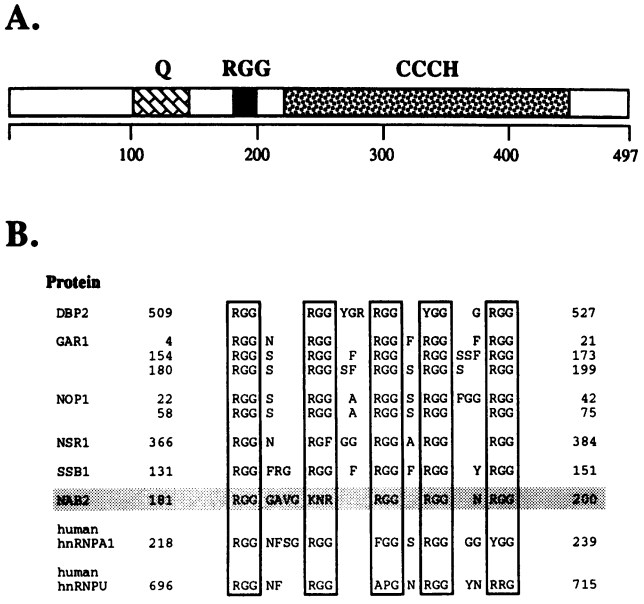

FIG. 6. Relationship of protein structural motifs in NAB2 with other RNA/ssDNA-binding proteins. (A) Illustration of distinctive structural motifs in NAB2, including the glutamine-rich (Q), RGG box (RGG), and C_3H repeat (CCCH) regions. (B) Relatedness of RGG boxes in various rRNA- and pre-mRNA-binding proteins. The RGG box of NAB2 (stippled) is aligned with those of the yeast nucleolar DBP2 (33), GAR1 (26), NOP1 (31), NSR1 (39), and SSB1 (34) proteins, as well as the human hnRNP complex A1 and U proteins (37). (C) The seven C_3H repeats in NAB2 (CCCH motif) aligned with the C_3H motifs present in RNA polymerases I, II, and III (PolI, PolII, and PolIII) from *S. cerevisiae* (65). Also shown are the HIV NC protein (11), which contains zinc knuckle RNA/ssDNA-binding motifs (CCHC motif), and the seven repeats in the sterol-mediated transcriptional repressor protein CNBP (49). Amino acid numbers for the HIV NC protein refer to the Gag polyprotein.

composed of a novel type of repeat motif related to the CCCH putative zinc-binding domain within the largest subunit of RNA polymerases I, II, and III (65) (Fig. 6C). In NAB2, seven repeats comprise a CCCH (C_3H) motif with a consensus $CX_5CX_{4-6}CX_3H$, whereas in the three RNA polymerases a single C_3H motif exists in each large subunit, with a consensus $CX_2CX_{6-12}CX_2H$ (65). This putative zinc-binding domain in RNA polymerases has been suggested to be responsible for the RNA and ssDNA-binding properties of these enzyme subunits as well as provide an interaction site for the zinc-binding domain of the second-largest subunit of eukaryotic RNA polymerases (65, 66). The C_3H motif is also related to the zinc knuckle family of RNA/ssDNA-binding proteins exemplified by two CCHC repeats found in the human immunodeficiency virus (HIV) nucleocapsid (NC) protein (Fig. 6C). Moreover, a sequence-specific ssDNA-binding protein, or cellular nucleic acid-binding protein (CNBP), is highly related to HIV NC and possesses seven CCHC motifs (49).

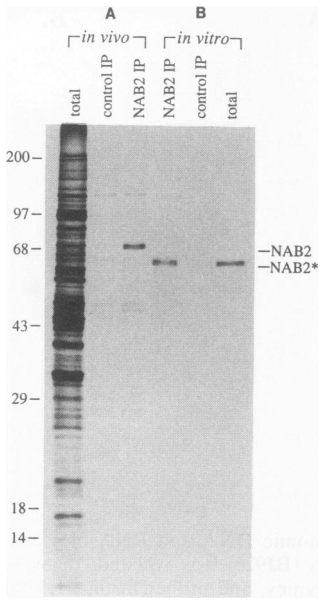

FIG. 7. Immunopurification of NAB2 produced in vivo and in vitro. Yeast cells (BJ926) were labeled in the presence of TRAN^{35}S (A), and either total protein (lanes total) or proteins immunopurified by using the MBP-NAB2 antiserum (lanes NAB2 IP) or preimmune antiserum (lanes control IP) were fractionated by SDS-PAGE. The in vitro-synthesized protein (B) is indicated by NAB2*. NAB2 migrates as a doublet either when isolated from yeast cells or when produced in vitro.

Sequence divergence of *NAB2* between yeast strains results in variations in the size of the glutamine-rich region. To determine whether the *NAB2* gene depicted in Fig. 5A encoded the mature NAB2 protein produced in the yeast cell, polymerase chain reaction (PCR) primers immediately upstream of the initiation codon and downstream of the termination codon were used to produce an in vitro transcription subclone of *NAB2*, pNAB2-3B6. NAB2 synthesized in vitro migrated as a doublet of 64 to 65 kDa (Fig. 7, NAB2*), while the protein isolated from TRAN^{35}S-labeled BJ926 yeast cells ran as a 71- to 72-kDa doublet (Fig. 7, NAB2). A larger genomic fragment from pRNP327, extending from −273 to +1961 (*Sty*I-*Nsi*I; Fig. 3 and 5A) and cloned into pSP72, also produced proteins of 64 to 65 kDa by in vitro transcription/translation (6). Therefore, the different migration of the in vitro-synthesized polypeptide is not due to a sequence artifact induced by PCR or to a lack of sequence information upstream of the initiation codon or downstream of the termination codon.

To resolve the discrepancy between the protein synthesized in BJ926 cells and the protein synthesized in vitro from

the pNAB327 genomic clone, we constructed a λZAP BJ926 cDNA library and isolated 20 *NAB2* cDNA clones. The 5′ end of the longest cDNA began 20 nt upstream of the initiation codon of the original λgt11 genomic clone, and the 3′ ends of all of the *NAB2* cDNAs were positioned within an A_{28} tract 144 nt downstream of the termination codon (Fig. 5A). In vitro translation of this cDNA produced NAB2 proteins which comigrated by SDS-PAGE with the in vivo-synthesized 71- to 72-kDa proteins (6) (see Fig. 10). DNA sequencing of both the BJ926 cDNA and a BJ926 genomic clone revealed that the BJ926 *NAB2* sequence was identical to the pNAB327 sequence except for an in-frame insertion of 84 nt after the first nucleotide for a codon encoding a leucine residue at position 110 (Fig. 5A, shaded). This insert results in the addition of a 28-amino-acid tetrapeptide repeat (QQQP) to NAB2 within the glutamine-rich region (Fig. 5B). Therefore, NAB2 from BJ926 is composed of 525 amino acids and has a calculated molecular weight of 58,321, or 3,371 larger than the protein produced from the gene shown in Fig. 5A. The addition of seven copies of the tetrapeptide repeat occurs immediately following two identical tetrapeptide repeats which possess the same nucleotide sequence. The most likely mechanism for generating this type of insert is allelic recombination during DNA replication. Presumably this recombinational event is relatively recent, since codon usage within the repeat has not yet diverged.

Both the RGG box and C_3H motifs bind RNA and ssDNA in vitro. To test whether the RGG box and C_3H motif regions were functional RNA-binding domains, we used an in vitro RNA/ssDNA binding assay. Full-length NAB2 protein, produced by in vitro transcription/translation, was tested for its ability to bind to the four RNA homopolymers as well as ssDNA at various salt concentrations. This binding assay has been successful at distinguishing the RNA binding specificities of a variety of hnRNA- and mRNA-binding proteins and has been used to study how the four individual domains of PAB1 bind not only to poly(A) but also to poly(G) and poly(U) (14, 62). For these RNA binding studies, the BJ926 *NAB2* gene (or cDNA), which encodes the 71- to 72-kD protein, was used.

Preliminary results indicated that NAB2 bound to poly(G), poly(A), and poly(U) up to 0.5 M NaCl but dissociated at between 0.5 and 1.0 M NaCl and did not bind at all to poly(C). NAB2 was also able to bind to ssDNA at 0.1 M NaCl but dissociated at between 0.1 and 0.5 M NaCl. In this assay, the binding of various subregions of the NAB2 protein was studied by using PCR to produce in vitro transcription constructs which were progressively truncated from the carboxyl terminus (Fig. 8A) or by cloning subregions in frame downstream of the MBP of *E. coli* (Fig. 8B). Figure 8C is an illustration of the *NAB2* constructs used for this study. Removal of 20 amino acid residues from the carboxyl terminus of NAB2 (Fig. 8A, NAB2.2) decreases the apparent molecular size of the protein by ~5 kDa, as determined by SDS-PAGE. This truncation has little effect on RNA/ssDNA binding, although binding to poly(U) at 0.5 M NaCl is diminished. When most of the C_3H motif region is removed (Fig. 8C, NAB2.3), the binding to poly(A) is eliminated. In addition, binding at 0.5 M NaCl is either eliminated for poly(U) or reduced for poly(G), and binding to ssDNA is also severely reduced. The region of NAB2 amino terminal to the RGG box (Fig. 8C, NAB2.0) does not bind to either RNA or ssDNA in this assay.

As another approach to study which domains of NAB2 were able to bind RNA and ssDNA in vitro, we tested the ability of various regions of NAB2 to convert a non-RNA-binding protein, the MBP, into an RNA-binding protein. We constructed an in vitro transcription/translation vector which allowed production of the MBP either alone or as an MBP-NAB2 fusion protein. These vectors were generated by positioning the MBP-coding region, with or without NAB2 subregions, downstream of a consensus eukaryotic translational initiation sequence (AAAATGA; see pMAL-c2ex in Table 1). Figure 8B demonstrates that although the MBP (392 amino acids) cannot bind to RNA or ssDNA, addition of the RGG box (93 amino acids) allows the fusion protein to bind to poly(G) and poly(U) and slightly to ssDNA (Fig. 8B, NAB2.4). It is also important to note that the truncation construct NAB2.3 and NAB2.4 gave similar results. In agreement with the results obtained with NAB2.3, addition of the C_3H motif results in binding to poly(A),

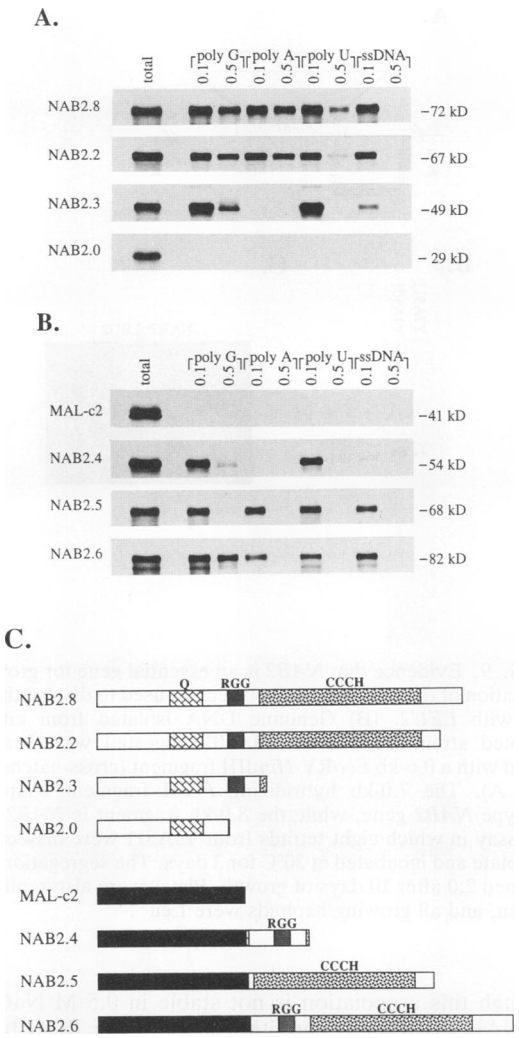

FIG. 8. Binding of RNA and ssDNA by the RGG box and C_3H repeat motifs. Labeled proteins, produced by in vitro transcription of various plasmids and translation of the resulting RNA in the presence of [^{35}S]methionine, were tested for the ability to bind to the RNA homopolymers poly(G), poly(A), and poly(U) as well as ssDNA. (A) Either intact NAB2 protein (NAB2.8 from BJ926) or various deletions from the carboxyl terminus. (B) Assay in which the RGG box and the C_3H repeat motifs of the NAB2 protein from BJ926 were tested for the ability to convert the MBP (MAL-c2) into an RNA/ssDNA-binding protein. (C) Schematic representation of the various plasmids used for in vitro transcription.

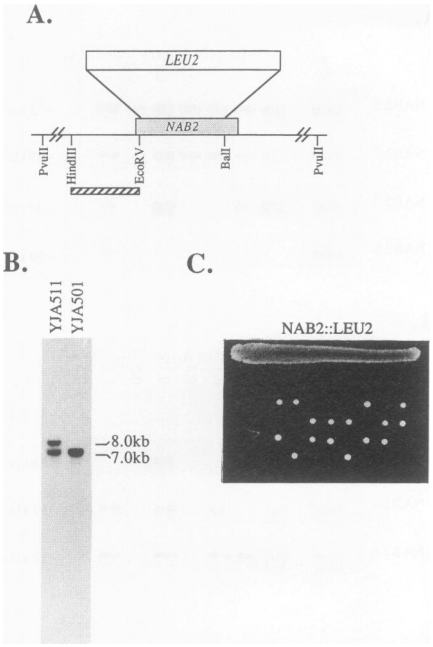

FIG. 9. Evidence that *NAB2* is an essential gene for growth. (A) Illustration of the one-step transplacement used to disrupt the *NAB2* gene with *LEU2*. (B) Genomic DNA isolated from either the disrupted strain YJA511 or YJA501, digested with *Pvu*II, and probed with a 0.6-kb *Eco*RV-*Hin*dIII fragment (cross-hatched bar in panel A). The 7.0-kb hybridizing *Pvu*II fragment contains the wild-type *NAB2* gene, while the 8.0-kb fragment is *NAB2*::*LEU2*. (C) Assay in which eight tetrads from YJA511 were dissected on a YPD plate and incubated at 30°C for 3 days. The segregation pattern remained 2:0 after 10 days of growth. Plates were also replicated to SC-Leu, and all growing haploids were Leu^-.

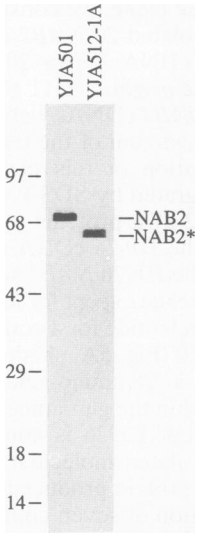

FIG. 10. Evidence that the QQQP tetrapeptide repeat insert is not essential for growth. Shown is an immunoblot of total cell protein from YJA501, containing the NAB2 protein with the 28-amino-acid tetrapeptide repeat, and YJA512-1A, which carries the complementing *Pvu*I-*Nsi*I fragment from pNAB327 on a plasmid encoding the NAB2* protein. Sizes are indicated in kilodaltons.

although this association is not stable in 0.5 M NaCl (Fig. 8B, NAB2.5). Finally, addition of both the RGG box and C_3H motif region to the MBP yields a fusion protein whose binding to RNA and ssDNA is similar to that of full-length NAB2, although binding to poly(A) and poly(U) at 0.5 M NaCl is not stable (Fig. 8A, NAB2.8, versus Fig. 8B, NAB2.6).

These studies indicate that the RGG box and C_3H motif region are responsible for the binding of NAB2 to both RNA homopolymers and ssDNA in vitro. Furthermore, in vitro binding of NAB2 to poly(A) appears to depend on the C_3H motif region.

NAB2 is essential for cell viability. To test whether expression of the NAB2 protein is essential for cell growth, a null allele was created by replacing 80% of the *NAB2* gene with *LEU2* (Fig. 9A). This construct was used to transplace the wild-type *NAB2* gene by transformation of the diploid YJA501 and selection of leucine prototrophs. Genomic DNA blot analysis of one of these Leu^+ diploids, YJA511, confirmed that one copy of the *NAB2* gene had been disrupted (Fig. 9B). YJA511 was sporulated, and 27 of 30 tetrads yielded two viable and two nonviable spores (Fig. 9C). All viable spores were leucine auxotrophs. Furthermore, this growth defect could be complemented by transformation of YJA511 with a CEN plasmid, pNAB2.15, harboring the *NAB2* gene (from pRNP327) and flanking sequences (−496 to +1956) and the selectable marker *URA3*. In 8 of 15 tetrads analyzed, 4 viable spores were recovered, and 16 of the 32 spores were able to grow on media lacking leucine and uracil but were unable to grow in the presence of 5-fluoro-orotic acid, which selectively kills cells expressing *URA3*. From these results, we conclude that NAB2 is essential for cell viability.

Preliminary work demonstrated that YJA501 produced the NAB2 protein which includes the tetrapeptide repeat within the glutamine-rich region. Since the null allele in YJA511 was complemented with the *NAB2* gene encoding the NAB2* protein (Fig. 5A), the 64- to 65-kDa NAB2* protein should have been produced in the Ura^+/Leu^+ haploids from the YJA512 transformation. Comparison of the apparent molecular weights of the NAB2 proteins produced in YJA501 with those of the proteins produced in YJA512-1A, a haploid containing pNAB2.15, demonstrated that the NAB2* protein is indeed synthesized in YJA512 (Fig. 10) and that the tetrapeptide repeat insert is not essential for cell growth.

DISCUSSION

One of the distinguishing features of higher eukaryotic hnRNP complex proteins is that many of them are specifically cross-linked by UV light to nuclear polyadenylated RNAs in vivo. Using the cross-linking technique, we have identified a set of yeast proteins ranging in size from 45 to 120 kDa which are directly associated with polyadenylated RNA in vivo. We have isolated and characterized one of these proteins, NAB2, a nuclear polyadenylated RNA-binding protein which appears to be a good candidate for the first hnRNP complex protein isolated from *S. cerevisiae*.

Relationship between nuclear polyadenylated RNA-binding proteins from *S. cerevisiae* and higher eukaryotes. NAB2 possesses structural features reminiscent of metazoan hnRNPs. A 26-amino-acid peptide, containing an RGG box, has been shown to be necessary for RNA-binding activity of the human hnRNP U proteins in vitro (37). The RNA-binding motif of the human K/J proteins, which bind in a salt-resistant manner to poly(C) in vitro, has not been

experimentally delineated, although the K/J proteins also possess a highly degenerate RGG box (37, 42).

Several hnRNP complex proteins from a variety of organisms contain more than one type of RNA-binding motif within the same protein. The human hnRNP A1 and *Drosophila* hrp40.1/2 proteins possess not only two RNP consensus sequence RNA-binding domains (9) but also an RGG box, and both motifs appear to bind RNA and ssDNA (9, 37). Functional analysis of the RNA-binding regions of the NAB2 protein reported here indicate that the NAB2 protein is also composed of more than one type of RNA-binding motif. Moreover, both the RGG box and C_3H motif repeats of NAB2 bind to RNA homopolymers and ssDNA in vitro, similar to the human hnRNP A1 proteins. However, the C_3H motif repeats are a novel type of RNA/ssDNA-binding motif which appear to be most closely related to the C_3H motifs of the largest subunit of eukaryotic RNA polymerases I, II, and III. This putative zinc-binding domain was originally suggested to conform to the zinc finger CCHH motif. Yano et al. (65, 66) have recently suggested, on the basis of the sequences of both RNA polymerases I and II from *Trypanosoma brucei* in which the second histidine is replaced with an aromatic amino acid, that C_3H is a more reasonable consensus. Since we have demonstrated that the seven C_3H motif repeats of NAB2 constitute an RNA/ssDNA-binding domain, it is likely that the single C_3H motifs of RNA polymerases I, II, and III possess similar binding activities. Of course, there are examples of proteins that do not bind nucleic acids which possess zinc finger-like motifs (11). These metal-binding domains have been suggested to promote protein-protein interactions, and it is thus possible that although the C_3H motifs bind polynucleotides in vitro, they may also facilitate protein-protein interactions in vivo.

The C_3H motif also appears to be related to the zinc knuckle repeats present in a variety of RNA/ssDNA-binding proteins, including retroviral NC proteins (11), the *Drosophila* copia element (45), and a human protein involved in sterol-mediated repression of gene expression, CNBP (49). This latter protein is particularly interesting since it also contains seven CCHC motif repeats, it does not bind to double-stranded DNA, and it recognizes only one of the DNA strands at the binding site. CNBP is believed to be a transcriptional repressor of sterol-mediated gene expression and consists almost entirely of the CCHC motif region. Since NAB2 can bind both RNA and ssDNA in vitro, it is possible that this protein is involved not only in RNA processing but also in transcriptional regulation. In this respect, it is interesting to note that transcription termination and 3′-end maturation appear to be tightly coupled processes in *S. cerevisiae* (64).

Pathway of nuclear mRNA biogenesis and possible functions of NABs. Nuclear mRNA biogenesis in higher eukaryotic cells is believed to consist of at least two steps. The first step includes the initiation of transcription by RNA polymerase II and production of nascent transcripts which appear to be immediately assembled into an RNP complex (5, 17). This complex may consist of hnRNAs and pre-mRNAs, direct hnRNA-binding proteins, and snRNPs as well as other alternative pre-mRNA-splicing and polyadenylation factors (5, 12, 59, 64). The step at which polyadenylation occurs, and whether it follows, precedes, or occurs simultaneously with splicing, is still disputed (27, 64). However, there is now considerable evidence that factors required for 3′-end maturation and splicing, such as snRNPs, associate with pre-mRNAs while they are still being transcribed (5, 12, 48). Since most *S. cerevisiae* genes do not possess introns, the first step in the principal RNA-processing pathway probably consists of cotranscriptional, or immediately posttranscriptional, 5′-end capping and polyadenylation. It has also been suggested that in both yeast cells (24) and higher eukaryotic cells (15), splicing factors may be concentrated in discrete subnuclear domains, and that pre-mRNAs may have to be transported to these regions from their site of transcription.

Once the primary transcript is correctly processed to generate mature mRNA, the second step is believed to be nucleocytoplasmic transport. Much less is known about how transport occurs, although it is thought that mRNA is recognized by the nuclear pore complex, possibly by virtue of its poly(A) tail, and exported into the cytoplasm by an ATP-dependent process (40). The entire export pathway from the site of RNA synthesis to extrusion from the nuclear pore has been visualized as occurring on tracks and has been hypothesized to be dependent on a nucleoskeletal network of fibers (38, 40). Mutations in two yeast genes, *RNA1* and *PRP20*, disrupt nucleocytoplasmic mRNA transport. However, a variety of nuclear RNA-processing reactions, including 3′-end formation and pre-mRNA splicing, as well as nuclear morphology (25, 32), are also severely affected. Recently, mutations in two additional genes which inhibit nucleocytoplasmic trafficking of mRNA have been isolated. When yeast cells with a temperature-sensitive mutation in either the *MTR1* or *RAT1* gene are grown at the nonpermissive temperature, they accumulate poly(A)$^+$ RNA in the nucleus, with only minor effects on pre-rRNA processing and no observable effects on pre-mRNA splicing in the case of *RAT1* (4, 36). The MTR1 protein is now known to be identical to PRP20 (35), while RAT1 is related to proteins which possess 5′-3′ exonuclease activity (4). It is important to note that these trafficking genes have been isolated by using genetic screens for nuclear poly(A)$^+$ RNA accumulation.

How would the NAB proteins be involved in these various processes? Since the pathway of mRNA biogenesis is so multifaceted, it is possible to envision a variety of protein factors which could be involved. These include such proteins as RNA helicases to eliminate unwanted RNA secondary structures, RNA-binding proteins which might protect the RNA from nucleases or package it in a manner which is compatible with export via a specific route, and RNA/ssDNA-binding factors which would not only bind to nascent transcripts but also interact with the transcriptional machinery so that transcriptional initiation and termination could possibly be comodulated. We therefore anticipate that NAB proteins will be found to be both functionally and structurally diverse.

What is the function of NAB2? The essential function of NAB2 in cell growth argues that this protein must play a critical role in this pathway, since it is strongly and specifically associated with nuclear poly(A)$^+$ RNA in vivo. Although this would imply that it is not exported with mRNA from the nucleus, a recent report demonstrating that the nuclear-localized human hnRNP A1 protein is shuttled into the cytoplasm along with mRNA would argue that NAB2 is a candidate for an mRNA transport protein (47). The role that NAB2 plays in such nuclear RNA-processing events as pre-mRNA polyadenylation, splicing, and nucleocytoplasmic mRNA trafficking will be the subject of future investigation, and such studies should be assisted by the recent isolation of a *NAB2* conditional lethal allele.

ACKNOWLEDGMENTS

We thank D. Duke, L. Green, and S. Henry of the Hybridoma Core Laboratory of the Interdisciplinary Center for Biotechnology Research for assistance with the preparation of polyclonal antisera and monoclonal antibodies, J. Aris for the gifts of monoclonal antibodies A66 and D77 against NOP1, and J. Aris, H. Baker, R. Condit, B. Flanegan, and A. Lewin for comments on the manuscript.

This work was supported by grant GM46272 from the National Institutes of Health and by a Junior Faculty Research Award to M.S.S. from the American Cancer Society.

REFERENCES

1. **Adam, S. A., T. Nakagawa, M. S. Swanson, T. K. Woodruff, and G. Dreyfuss.** 1986. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. Mol. Cell. Biol. **6**:2932–2943.
2. **All-Robyn, J. A., N. Brown, E. Otaka, and S. W. Liebman.** 1990. Sequence and functional similarity between a yeast ribosomal protein and the *Escherichia coli* S5 *ram* protein. Mol. Cell. Biol. **10**:6544–6553.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic alignment search tool. J. Mol. Biol. **215**: 403–410.
4. **Amberg, D. C., A. L. Goldstein, and C. N. Cole.** 1992. Isolation and characterization of *RAT1*: an essential gene of *Saccharomyces cerevisiae* required for the efficient nucleocytoplasmic trafficking of mRNA. Genes Dev. **6**:1173–1189.
5. **Amero, S. A., R. Raychaudhuri, C. L. Cass, W. J. van Venrooij, W. J. Habets, A. R. Krainer, and A. L. Beyer.** 1992. Independent deposition of heterogeneous nuclear ribonucleoproteins and small nuclear ribonucleoprotein particles at sites of transcription. Proc. Natl. Acad. Sci. USA **89**:8409–8413.
6. **Anderson, J. T., and M. S. Swanson.** Unpublished data.
7. **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.
8. **Aris, J. P., and G. Blobel.** 1991. Isolation of yeast nuclei. Methods Enzymol. **194**:735–749.
9. **Bandziulis, R. J., M. S. Swanson, and G. Dreyfuss.** 1989. RNA-binding proteins as developmental regulators. Genes Dev. **3**:431–437.
10. **Bennett, M., S. Piñol-Roma, D. Staknis, G. Dreyfuss, and R. Reed.** 1992. Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly in vitro. Mol. Cell. Biol. **12**:3165–3175.
11. **Berg, J. M.** 1990. Zinc fingers and other metal-binding domains. J. Biol. Chem. **265**:6513–6516.
12. **Beyer, A. L., and Y. N. Osheim.** 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. Genes Dev. **2**:754–765.
13. **Burd, C. G., E. L. Matunis, and G. Dreyfuss.** 1991. The multiple RNA-binding domains of the mRNA poly(A)-binding protein have different RNA-binding activities. Mol. Cell. Biol. **11**:3419–3424.
14. **Burd, C. G., M. S. Swanson, M. Gorlach, and G. Dreyfuss.** 1989. Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: a diversity of RNA binding proteins is generated by small peptide inserts. Proc. Natl. Acad. Sci. USA **86**:9788–9792.
15. **Carmo-Fonseca, M., R. Pepperkok, B. S. Sproat, W. Ansorge, M. S. Swanson, and A. I. Lamond.** 1992. *In vivo* detection of snRNP-rich organelles in the nuclei of mammalian cells. EMBO J. **10**:1863–1873.
16. **Cattaneo, R.** 1991. Different types of messenger RNA editing. Annu. Rev. Genet. **25**:71–88.
17. **Chung, S. Y., and J. Wooley.** 1986. Set of novel, conserved proteins fold pre-messenger RNA into ribonucleosomes. Proteins **1**:195–210.
18. **Cleveland, D. W., and T. J. Yen.** 1992. Multiple determinants of eukaryotic mRNA stability. New Biol. **1**:121–126.
19. **Datar, K. V., S. M. Wilson, and M. S. Swanson.** Unpublished data.
20. **Davis, L. I., and G. R. Fink.** 1990. The *NUP1* gene encodes an essential component of the yeast nuclear pore complex. Cell **61**:965–978.
21. **Dreyfuss, G.** 1986. Structure and function of nuclear and cytoplasmic ribonucleoprotein particles. Annu. Rev. Cell Biol. **2**:459–498.
22. **Dreyfuss, G., Y. D. Choi, and S. A. Adam.** 1984. Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. Mol. Cell. Biol. **4**:1104–1114.
23. **Dreyfuss, G., M. S. Swanson, and S. Piñol-Roma.** 1988. Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. Trends Biochem. Sci. **13**:86–91.
24. **Elliott, D. J., D. S. Bowman, N. Abovich, F. S. Fay, and M. Rosbash.** 1992. A yeast splicing factor is localized in discrete subnuclear domains. EMBO J. **10**:3731–3736.
25. **Forrester, W., F. Stutz, M. Rosbash, and M. Wickens.** 1992. Defects in mRNA 3′-end formation, transcription initiation, and mRNA transport associated with the yeast mutation *prp20*: possible coupling of mRNA processing and chromatin structure. Genes Dev. **6**:1914–1926.
26. **Girard, J.-P., H. Lehtonen, M. Caizergues-Ferrer, F. Amalric, D. Tollervey, and B. Lapeyre.** 1992. GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. EMBO J. **11**:673–682.
27. **Green, M. R.** 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. **7**:559–599.
28. **Groner, B., and S. L. Phillips.** 1975. Polyadenylate metabolism in the nuclei and cytoplasm of *Saccharomyces cerevisiae*. J. Biol. Chem. **250**:5640–5646.
29. **Guthrie, C., and G. R. Fink (ed.).** 1991. Guide to yeast genetics and molecular biology. Academic Press, Inc., San Diego, Calif.
30. **Guthrie, C., and B. Patterson.** 1988. Spliceosomal snRNAs. Annu. Rev. Genet. **22**:387–419.
31. **Henriquez, R., G. Blobel, and J. P. Aris.** 1990. Isolation and sequencing of *NOP1*: a yeast gene encoding a nucleolar protein homologous to a human autoimmune antigen. J. Biol. Chem. **265**:2209–2215.
32. **Hopper, A. K., H. M. Traglia, and R. W. Dunst.** 1990. The yeast *RNA1* gene product necessary for RNA processing is located in the cytosol and apparently excluded from the nucleus. J. Cell Biol. **111**:309–321.
33. **Iggo, R. D., D. J. Jamieson, S. A. MacNeill, J. Southgate, J. McPheat, and D. P. Lane.** 1991. p68 RNA helicase: identification of a nucleolar form and cloning of related genes containing a conserved intron in yeasts. Mol. Cell. Biol. **11**:1326–1333.
34. **Jong, A. Y.-S., M. W. Clark, M. Gilbert, A. Oehm, and J. L. Campbell.** 1987. *Saccharomyces cerevisiae* SSB1 protein and its relationship to nucleolar RNA-binding proteins. Mol. Cell. Biol. **7**:2947–2955.
35. **Kadowaki, T., and A. Tartakoff.** 1992. A yeast gene implicated in mRNA export from nucleus to cytoplasm. Mol. Biol. Cell **3**:134a.
36. **Kadowaki, T., Y. Zhag, and A. M. Tartakoff.** 1992. A conditional yeast mutant deficient in mRNA transport from nucleus to cytoplasm. Proc. Natl. Acad. Sci. USA **89**:2312–2316.
37. **Kiledjian, M., and G. Dreyfuss.** 1992. Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. EMBO J. **11**:2655–2664.
38. **Lawrence, J. B., R. H. Singer, and L. M. Marselle.** 1989. Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. Cell **57**:493–502.
39. **Lee, W.-C., Z. Xue, and T. Melese.** 1991. The *NSR1* gene encodes a protein that specifically binds nuclear localization sequences and has two RNA recognition motifs. J. Cell Biol. **113**:1–12.
40. **Maquat, L. E.** 1991. Nuclear mRNA export. Curr. Opin. Cell Biol. **3**:1004–1012.
41. **Mattox, W., L. Ryner, and B. S. Baker.** 1992. Autoregulation and multifunctionality among trans-acting factors that regulate alternative pre-mRNA processing. J. Biol. Chem. **267**:19023–19026.
42. **Matunis, M. J., W. M. Michael, and G. Dreyfuss.** 1992. Characterization and primary structure of the poly(C)-binding heter-

ogeneous nuclear ribonucleoprotein complex K protein. Mol. Cell. Biol. **12**:164–171.
43. **Mayeda, A., and A. R. Krainer.** 1992. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. Cell **68**:365–375.
44. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
45. **Mount, S. M., and G. M. Rubin.** 1985. Complete nucleotide sequence of the *Drosophila* transposable element copia: homology between copia and retroviral proteins. Mol. Cell. Biol. **5**:1630–1638.
46. **Ohno, M., N. Kataoka, and Y. Shimura.** 1990. A nuclear cap binding protein from HeLa cells. Nucleic Acids Res. **18**:6989–6995.
47. **Piñol-Roma, S., and G. Dreyfuss.** 1992. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. Nature (London) **355**:730–732.
48. **Piñol-Roma, S., M. S. Swanson, J. G. Gall, and G. Dreyfuss.** 1989. A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts. J. Cell Biol. **109**:2575–2587.
49. **Rajavashisth, T. B., A. K. Taylor, A. Andalibi, K. L. Svenson, and A. J. Lusis.** 1989. Identification of a zinc finger protein that binds to the sterol regulatory element. Science **245**:640–643.
50. **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**:1540–1572.
51. **Rose, M. D., F. Winston, and P. Hieter.** 1990. Methods in yeast genetics. A laboratory course manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
52. **Rozen, F., and N. Sonenberg.** 1987. Identification of nuclear cap specific proteins in HeLa cells. Nucleic Acids Res. **15**:6489–6500.
53. **Ruby, S. W., and J. Abelson.** 1991. Pre-mRNA splicing in yeast. Trends Genet. **7**:79–85.
54. **Sachs, A. B., M. W. Bond, and R. W. Davis.** 1986. A single gene from yeast for both the nuclear and cytoplasmic polyadenylate-binding proteins: domain structure and expression. Cell **45**:827–835.
55. **Sachs, A. B., and R. W. Davis.** 1989. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. Cell **58**:857–867.
56. **Sachs, A. B., and J. A. Deardorff.** 1992. Translation initiation requires the PAB-dependent poly(A) ribonuclease in yeast. Cell **70**:961–973.
57. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
58. **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.
59. **Smith, C. W. J., J. G. Patton, and B. Nadal-Ginard.** 1989. Alternative splicing in the control of gene expression. Annu. Rev. Genet. **23**:527–577.
60. **Snyder, M., S. Elledge, D. Sweeter, R. A. Young, and R. W. Davis.** 1987. λgt11: gene isolation with antibody probes and applications. Methods Enzymol. **154**:107–128.
61. **Steitz, J. A., D. L. Black, V. Gerke, K. A. Parker, D. Frendewey, and W. Keller.** 1988. Functions of the abundant U-snRNPs, p. 115–154. *In* M. Birnstiel (ed.), Structure and function of major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Heidelberg.
62. **Swanson, M. S., and G. Dreyfuss.** 1988. Classification and purification of proteins of heterogeneous nuclear ribonucleoprotein particles by RNA-binding specificities. Mol. Cell. Biol. **8**:2237–2241.
63. **Swanson, M. S., and G. Dreyfuss.** 1988. RNA-binding specificity of hnRNP proteins: a subset bind to the 3′ end of introns. EMBO J. **7**:3519–3529.
64. **Wahle, E., and W. Keller.** 1992. The biochemistry of 3′-end cleavage and polyadenylation of messenger RNA precursors. Annu. Rev. Biochem. **61**:419–440.
65. **Yano, R., and M. Nomura.** 1991. Suppressor analysis of temperature-sensitive mutations of the largest subunit of RNA polymerase I in *Saccharomyces cerevisiae*: a suppressor gene encodes the second-largest subunit of RNA polymerase I. Mol. Cell. Biol. **11**:754–764.
66. **Yano, R., M. Oakes, M. Yamaghishi, J. A. Dodd, and M. Nomura.** 1992. Cloning and characterization of *SRP1*, a suppressor of temperature-sensitive RNA polymerase I mutations, in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **12**:5640–5651.