## Expression of RNase P RNA in *Saccharomyces cerevisiae* is controlled by an unusual RNA polymerase III promoter

(RPR1 gene/transcription factor TFIIIC/internal promoters)

JAE-YONG LEE, CLAIRE F. EVANS<sup>†</sup>, AND DAVID R. ENGELKE<sup>‡</sup>

Department of Biological Chemistry, The University of Michigan, Ann Arbor, MI 48109-0606

Communicated by Minor J. Coon, May 20, 1991 (received for review March 20, 1991)

ABSTRACT The RNA subunit of Saccharomyces cerevisiae nuclear RNase P is encoded by a single-copy, essential gene, RPR1. The 369-nucleotide mature form of the RNA has an apparent precursor with an 84-nucleotide 5' leader and  $\approx$ 33 nucleotides of additional 3' sequence. Analysis of RPR1 transcription in a strain with a temperature-sensitive lesion in RNA polymerase III shows that the gene is transcribed in vivo by RNA polymerase III. Examination of potential promoter regions using both progressive upstream deletions and point mutations indicates that at least two sequences contained within the 5' leader region are essential for expression in vivo, while sequences farther upstream influence efficiency. The required leader elements resemble tRNA gene-like A-box and B-box internal promoters in sequence and spacing. As in the tRNA genes, transcription factor TFIIIC binds to this region in vitro and binding is severely reduced by either A-box or B-box point mutations that impair expression in vivo. It thus appears that the yeast RNase P RNA gene has adopted a promoter strategy that places an RNA polymerase III "internal" promoter upstream of the mature structural domain to help drive transcription.

A variety of eukaryotic genes encoding small RNAs are transcribed by RNA polymerase (pol) III. Characterization of the promoter elements of these genes has identified two divergent strategies for placement of transcriptional regulatory elements in the DNA. Early work on genes for 5S rRNA and tRNAs demonstrated that their expression was controlled largely by DNA elements within the RNA coding regions, although in these cases levels of activity can also be modulated by upstream flanking regions (reviewed in refs. 1) and 2). 7SL RNA genes, several small viral RNA genes, and pol III transcription units within several highly repetitive DNA elements also fall into this promoter class (1, 2). All members of this class except 5S have related bipartite internal promoters consisting of an "A box" and "B box" separated by 30-90 base pairs (bp), with the A box 8-25 bp downstream of the transcription start site. The A and B boxes are initially bound by a complex transcription factor, TFIIIC, and a full transcription complex is subsequently assembled by the addition of at least one other factor (TFIIIB) and pol III (3-5). More recently it has been established that another class of genes, including 7SK (6, 7) and U6 from plants (8) and vertebrates (9-12), is regulated primarily by upstream sequences similar to those found in pol II promoters. In at least one reported case there appears to be a combination of these two strategies. The yeast U6 gene has upstream promoter elements similar to those found in the vertebrate U6 gene but also has a normally positioned A box within the U6 RNA coding sequence and a functional B box downstream of the coding region (13).

In this report a pol III promoter strategy is proposed that circumvents the need for promoters located within the RNA coding region. The RPR1 gene from Saccharomyces cerevisiae encodes the highly structured 369-nucleotide RNA subunit of the nuclear RNase P (14). This RNA appears to be made as a precursor with an 84-nucleotide 5' leader and  $\approx 30$ extra 3' nucleotides, ending in a  $U_{4-6}$  sequence common to pol III terminators. It was previously demonstrated that the human RNase P RNA (H1 RNA) gene could be transcribed by pol III in vitro (15). Here a temperature-sensitive lesion in pol III is used to show that RPRI RNA is synthesized by pol III in vivo. Further, a combination of upstream deletions and point mutations suggests that while DNA upstream of the 5' leader influences the efficiency of expression, the essential portions of the promoter are an A box and a nonconsensus B box located within the leader region. Thus an "internal" promoter has been placed upstream of a structural RNA to drive its expression.

## **MATERIALS AND METHODS**

Strains and Genetic Methods. S. cerevisiae strain W3031A (MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1-100) and its isogenic strain W3031B (MATa ade2-1 his3-11,15 leu2-3,112 trp1-1 ura3-1 can1-100) were obtained from R. Rothstein (Columbia University). A temperature-sensitive pol III mutant, D27-7C (MATa pep4::HIS<sup>+</sup>MAT ura3-52 his3-200 trp1-901 lys2-801 ade2-101 rpc160-41) was obtained from Pierre Thuriaux and Andre Sentenac (16) along with its parental strain, YNN281. Yeast transformation was performed by the lithium acetate procedure (17). Sporulation and tetrad analysis were performed using standard genetic methods (18). Cells were grown either in selection medium (0.67% yeast nitrogen base without amino acids, 2% dextrose, and appropriate amino acid supplements) or in YPD medium (2% dextrose, 2% peptone, and 1% yeast extract).

**Preparation of RPRI Gene Variants.** An RPR1 gene subclone containing the 1.3-kilobase (kb) Bgl II-HindIII fragment in the BamHI-HindIII polylinker sites of pUC19 (14) was used for making a series of deletions from the 5' flanking region of the RPR1 gene. The plasmid was first cut with one of the following restriction endonucleases: EcoRI at -407, Hpa I at -182, Hae I at -130, Stu I at -69, or Rsa I at -51 (see Fig. 1). EcoRI octamer linkers were added to the resulting ends and the gene fragment from the new EcoRI site to the downstream HindIII site was cloned between the EcoRI and HindIII sites of YCp50. To delete the entire upstream region of RPR1, the region downstream of +1 was amplified by polymerase chain reaction using an oligonucleotide that introduced an EcoRI site at +1 (14) and an

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: pol, RNA polymerase.

<sup>&</sup>lt;sup>†</sup>Present address: Department of Neuropharmacology, Scripps Clinic and Research Foundation, 10666 North Torrey Pines Road, La Jolla, CA 92037.

<sup>&</sup>lt;sup>‡</sup>To whom reprint requests should be addressed.

oligonucleotide from a site downstream of the HindIII site. Point mutations were created by site-directed mutagenesis (19) using a 1.3-kb Bgl II-HindIII fragment in phage vector M13mp18. Mutations introduced into A- and B-box promoter candidates are shown in Fig. 1. To specifically detect RPR1 RNA expression from plasmid-borne RPR1 genes, a 7-bp segment at positions 69-75 in a loop region of the RPRI RNA (ref. 20 and unpublished observations) was replaced in some constructs (see Fig. 3) after it was confirmed that the gene with the 7-bp replacement could complement the RPR1 disruption. The 7-bp substitution was combined with upstream deletions and candidate internal promoter mutations by ligating the appropriate EcoRI-BamHI-HindIII fragments or EcoRI-Sty I-HindIII fragments of M13mp18 replicative form into the EcoRI-HindIII sites of YCp50. Identities of mutated and recombined genes were confirmed by DNA sequencing.

**Complementation by RPR1 Gene Variants.** Diploid strain W303 with one copy of RPR1 disrupted by HIS3 (14) was transformed with single-copy yeast vectors containing RPR1 gene variants and a URA3 gene as a selection marker and was grown on selection plates lacking uracil. Transformants were sporulated, tetrads were dissected, and spores were grown on YPD plates to test viability.

**RNA Blot Analysis.** For analysis of RNAs from the pol III temperature-sensitive mutant and the parental wild-type strain, inocula from fresh cultures were grown at 25°C or 37°C for  $\approx 20$  hr in YPD medium until reaching an OD<sub>650</sub> of 1.0. Total RNA was isolated as described (21). The amount of RNA giving the same intensity of rRNA bands (about 10  $\mu$ g) was determined on an ethidium bromide-stained agarose gel and electrophoresed in 2 M formaldehyde/0.8% agarose gels

-	agatctgco	aattgaacat	aacatggtag	ttacatatac	tagtaatatg	gttcggcaca
-	348 cattaaaagt	: ataaaaacta	tctgaattac	gaattacata	tattggtcat	aaaaatcaat
-	288 caatcatcgt	: gtgttttata	tgtctcttat	ctaagtataa	gaatatccat	agttaatatt
-	228 cacttacgct	accttttaac	ctgtaatcat	tgtcaacagg	-182 tatgttaacg	acccacattg
-	168 ataaacgcta	gtatttcttt	tteetettet	-130	ctgtctctat	acteccetat
-:	108	******		-69	A Bo	-51
	48		B* Box		1	CC
1	13	CGCGGCTG <u>GG</u>	G C	TGGGAGCTGC	GATTGGCA <u>GT</u>	GGAACAGTGG
_	TAATTCCTAC	GATTAAGAAA	CCTGTTTACA	GAAGGATCCC	CACCTATGGG	CGGGTTATCA GTCC
	GATATTATCA TGC	GGTGGGAAAT	TCGGTGGAAC	ACAGTGGAGC	CTIGICCICC	<u>GGGTTAATGT</u>
1	.33 CGCTTTTGGC	ATTGGCCCCT	GCTCCTGAGA	GAAGAAATAT	ACTOCCOARC	CAGTOTTAC
					ACIGOGGAAC	CAULCILIAC
1	.93 CGACCGTTGT	TATCAGAAAT	TCACGGAGTT	EGGCCTAGGT	1 Box CGGACTCCGA	TGGGAACGGC
1	93 CGACCGTTGT 53 AACGGTTGTT	TATCAGAAAT CCGTTTGACT	TCACGGAGTT	B CGGCCTAGGT C TACGGCGTGA	1 Box CGGACTCCGA T SCGTCAAGGT	TGGGAACGGC CTGTTGAGTG
1 2 3	93 CGACCGTTGT 53 AACGGTTGTT 13 CAATCGTAGG	TATCAGAAAT CCGTTTGACT ACGTCATTAG	TCACGGAGTT TGTCGCCCGC	B CGGCCTAGGT C TACGGCGTGA	1 Box CGGACTCCGA T SCGTCAAGGT ACTGCTGCTG	TGGGAACGGC CTGTTGAGTG Box 369 TTCCAGCCCA
1 2 3	93 CGACCGTTGT 53 AACGGTTGTT 13 CAATCGTAGG 73 TATCCAACTT	TATCAGAAAT CCGTTTGACT ACGTCATTAG CCAATTTAAT	TCACGGAGTT TGTCGCCCGC TGGCGAACCC	B CGGCCTAGGT C TACGGCGTGA GATACCGATT	1 Box CGGACTCCGA T GCGTCAAGGT ACTGCTGCTG :tatttgcga 1	TGGGAACGGC CTGTTGAGTG Box 369 TTCCAGCCCA T G Lacagaaaga

FIG. 1. Features of the *RPR1* gene sequence and positions of mutations. Positions are numbered relative to the first nucleotide in the mature RNase P RNA (14, 20). The 5' termini for upstream deletions are indicated with bold numbers (-407, -182, -130, -69, -51, and 1). The internal promoter homologies are contained in boxes, and nucleotides shown below the line indicate the replacement nucleotides in point mutations used to test candidate internal promoters. A neutral 7-bp substitution at positions 69–75, used to generate a unique oligonucleotide hybridization site, is also shown. The underlined sequences represent the mature RNA coding region, and the uppercase letters show the extent of the larger, putative precursor RNA.

(19). Membranes from parallel blot transfers were hybridized with antisense RPRI or tRNA<sub>3</sub><sup>Leu</sup> radioactive probes, washed, and subjected to autoradiography (14, 18). As a control for RNA synthesis by pol II, duplicate blots were probed with a 5'- $^{32}$ P oligonucleotide that specifically detected LEU2 mRNA, which remained stable or increased slightly at the nonpermissive temperature in the pol III mutant (data not shown). To uniquely detect levels of RPRI RNA from plasmid genes, a neutral 7-bp substitution was made in the coding region (positions 69-75) in addition to the upstream and internal promoter candidate mutations indicated in Fig. 3. Total RNA was prepared as above, except that growth was at 30°C in synthetic medium lacking uracil. Ten micrograms of each RNA was electroblotted from a denaturing 8% polyacrylamide gel and probed with the  $5^{\prime}$ -<sup>32</sup>P antisense 28-mer oligonucleotide that was used to create the 7-bp substitution. Radioactivity in the combined RPR1 RNA bands was determined using a Betascope 603 blot analyzer (Betagen, Waltham, MA).

DNA Binding Assays with TFIIIC. A SUP53 tRNA gene fragment was prepared by end-filling the vector *Eco*RI site of pUC53 (70 bp upstream of the tRNA gene; ref. 22) with  $[\alpha^{-32}P]$ dATP and  $[\alpha^{-32}P]$ dTTP (3000 cpm/mmol), followed by cleavage downstream of the gene with HindIII to release a 652-bp fragment. End-labeled fragments of the RPR1 gene and mutated variants used in mobility shifts and footprinting of the sense strand were prepared by cleavage at the Eae I site at position -136, labeling with polynucleotide kinase (23), and cleavage at a HindIII site created at position +472 to give a 608-bp fragment. For footprinting the antisense strand, an *RPR1* fragment extending from the Taq I site at position -90to the HindIII site was inserted into the Acc I-HindIII sites of plasmid pGEM-Blue (Promega). End labeling of the antisense strand at the vector EcoRI site was done as above for SUP53, and the EcoRI-HindIII fragment was isolated. Purification of TFIIIC from extracts of yeast strain PP1002 through chromatography on Bio-Rex 70 (Bio-Rad), DEAE-Sephadex A-25 (Pharmacia), and B-box DNA affinity columns has been described (22, 24). The amount of factor required to give the mobility shifts in Fig. 4 was determined empirically and was approximately the same for wild-type tRNA and RPR1 genes. In mobility-shift experiments endlabeled DNA fragments were incubated with or without 1.5  $\mu$ l of affinity-purified TFIIIC at 25°C for 12 min in a final volume of 10 µl containing 20 mM Hepes (pH 7.9), 70-85 mM ammonium sulfate, 4 mM MgCl<sub>2</sub>, 5.5% (vol/vol) glycerol, 0.1% (wt/vol) bovine serum albumin, 100 ng of competitor pUC9 DNA (cleaved with Rsa I), and 200 ng of additional plasmid DNA competitor where indicated. Samples were run at 300-400 V in nondenaturing 4% polyacrylamide gels in 0.25× TBE buffer (1× TBE: 89 mM Tris borate/2 mM EDTA, pH 8.3) plus 5% glycerol at 4°C. For DNase I footprinting, 2 ng of end-labeled DNA was preincubated with 4  $\mu$ l of TFIIIC ( $\approx$ 4-fold excess) under the same conditions as for shifts, except glycerol concentrations were held to 0-3%. DNase I digestions and analysis of cleavage sites were performed as described (25). Products of dideoxynucleotide sequencing reactions (data not shown) were used as size markers for the positions of DNase I cleavages.

## RESULTS

The predominant RNA associated with purified nuclear RNase P from S. cerevisiae is 369 nucleotides long, and a defective form of the gene encoding this RNA, RPR1, is defective in 5' end processing of tRNA precursors (14, 20). A longer, less abundant form of the RNA is also present in extracts and highly purified enzyme and has an 84-nucleotide 5' leader and 27-33 nucleotides of extra 3' sequence. Although a precursor-product processing relationship has not



FIG. 2. *RPR1* RNA and tRNA<sup>1</sup><sub>2</sub>eu levels in a pol III mutant at nonpermissive temperature. Total RNA isolated from wild-type cells or a derivative pol III temperature-sensitive (ts) mutant grown at  $25^{\circ}$ C or  $37^{\circ}$ C was subjected to Northern blot analysis using antisense *RPR1* RNA (*Upper*) and tRNA<sup>1</sup><sub>2</sub>eu (*Lower*) probes. The identities of the various RNAs given at right have been confirmed (14, 26).

been rigorously established between these two forms, the coordinate response of the two RNAs to pol III and promoter mutations (see below) is consistent with this assignment. The extent of the two RNAs within the *RPR1* gene region is shown in Fig. 1, with positions numbered according to the smaller, mature form of the RNA. The positions of deletions, pol III internal promoter homologies, and nucleotides substituted in directed mutations used in this work are also indicated.

To establish that *RPR1* is transcribed by pol III, the level of *RPR1* was determined in an *S. cerevisiae* mutant harboring a temperature-sensitive lesion in the large subunit of pol III

	<b>•</b>	Comple- mentation	% Steady state level of RNA
5'∆- <b>40</b> 7	-407 -84 A B* B1 B2	+	100
5' <b>Δ-182</b>	· <u>182</u> <b>CCC</b>	+	41
5'∆-1 <b>30</b>	-1 <u>30</u>	+	26
5' <b>Δ-6</b> 9	- <sup>69</sup> 00- <b></b> 0-0	+†	11
5'∆- <b>5</b> 1	- <u>51</u> -C3- <b>C3-C3-C3-C3-</b>	_	0
5'∆ <b>1</b>	<sup>1</sup>	_	0
<i>rpr1</i> no A	-407		4
<i>rpr1</i> no B*		—	0
RPR1 no B1	00	+	94
RPRI no B2	00	+	88

FIG. 3. Complementation and *RPR1* RNA expression using variant *RPR1* genes. Expression of the indicated *RPR1* variants from single-copy plasmids was tested in a diploid strain carrying one normal and one disrupted chromosomal copy of the essential *RPR1* gene (14) by both complementation and RNA blot analysis. The effects of progressive 5' deletions and point mutations in suspected A-box and B-box promoter elements on complementation and levels of *RPR1* RNA are shown. The extent of the larger *RPR1* RNA is indicated by an arrow at the top. The coding region for the smaller RNA, the major form associated with RNase P activity (20), is shown as a dark bar, and the A-box and B-box promoter homologies (see Figs. 1 and 6) are shown as open boxes. "X" in boxes indicates point mutations; † indicates a slightly reduced growth rate. Complementation analyses and quantitation of RNA from the plasmid-borne *RPR1* variants by RNA blots are described in *Materials and Methods*.

(Fig. 2). The level of both forms of *RPRI* RNA was severely reduced in the mutant grown at nonpermissive temperature (37°C) relative either to growth at the permissive temperature or to the parental strain grown at either temperature. The degree of reduced expression was comparable to that of



FIG. 4. Binding of TFIIIC to the SUP53 tRNA gene and RPR1 gene variants. Electrophoretic mobility shift of the SUP53 tRNA gene or RPR1 gene variant DNA fragments was tested with binding of affinity-purified transcription factor TFIIIC. (Left) Radiolabeled DNA fragment probes contained either the wild-type RPR1 gene (RPR1 lanes) or variants with mutated A or B\* internal promoters [rpr1 (no A box) and rpr1 (no B\* box) lanes]. Binding reactions with additional unlabeled competitor plasmid ( $0.2 \mu g$ ) without (vector lane) or with (SUP53 lane) the SUP53 tRNA gene insert are indicated. (Right) The radiolabeled DNA probe in all lanes contained the SUP53 tRNA gene. The indicated DNA competitors were contained in unlabeled plasmid DNA ( $0.2 \mu g$ ). SUP53 and RPR1 denote wild-type genes; sup53 and rpr1 denote gene variants with mutated internal promoters. The upper shifted bands correspond to full TFIIIC complexes. Less intense bands migrating between the unbound probe and the full complex correspond to shifts observed with partially proteolyzed TFIIIC (unpublished observations).



FIG. 5. Binding of TFIIIC to the *RPR1* leader region. Excess affinity-purified *S. cerevisiae* TFIIIC was bound to *RPR1* gene fragments labeled upstream of the coding region on either the sense or the antisense strand. DNase I footprinting was performed as described in the text. A DNA fragment digested with DNase I in the absence of TFIIIC was run for comparison, and DNA sequencing lanes (not shown) served as position markers. Indicated to the left of each autoradiograph are positions corresponding to the *RPR1* mature domain (thick solid bar, +1 and above) and leader region (thin solid bar, -84 to -1). The positions of the A box and the B\* box are also indicated. Cleavages intensified in the presence of TFIIIC are denoted by arrows to the right of each autoradiograph, with the minimum extent of the footprint as defined by protected cleavage sites denoted by open bars.

tRNA, which is a pol III transcript known to be diminished in this mutant (16).

Identification of DNA elements required for pol III expression in vivo was accomplished by constructing two types of **RPR1** mutations. Progressive upstream deletions were made because sequence comparisons with the yeast U6 gene suggested numerous homologies with RPR1 (see Discussion). In addition, computer analysis revealed a tRNA-like A-box promoter properly positioned within the 5' leader and two consensus B boxes (B1 and B2) in the correct orientations within the mature RNA coding domain (see Fig. 1). Although the B1 and B2 elements are much farther downstream from the RNA 5' ends than is optimum for tRNA gene promoters, a B box similarly far downstream in the yeast U6 gene has been shown to regulate transcription (13). Another B box (B\*) with a relatively poor match to the consensus was identified in the leader region only when TFIIIC was found to bind exclusively to the leader region (see Fig. 5). Double point mutations were made at invariant positions in each

а		
A box cons A at	ensus : <u>T</u> RGCN -64 : TGGCG	N <u>a</u> gyn <u>gg</u> Cacatgg
B box cons B* at -	ensus : G <u>GTTC</u> 30 : GGAAC	G <u>a</u> nt <u>c</u> c Gaaact
b		
	Sequence	Position
U6: <i>RPR1</i> :	TTTTTTTCG TTTCTTTCG	-18 -102 (-18)
U6 :	CTATA	-31
KFKI:	CTATA CTATA	-112 (-28) -122 (-38)
U6 :	TCCACTAT	-48
<i>RPR1</i> :	TCCCCTAT	-116 (-32)
U6 :	AGTATTTCGT	-57 •
RPRI:	AGTATTTCTT	-159 (-75)
U6 :	TAAAAGTAT	-61
RPR1:	TAAAAGTAT	-345 (-261)
U6 :	TCAACA	-87
RPRI:	TCAACA	-196 (-112)

FIG. 6. Sequence comparisons of internal promoters and RPRI/U6 upstream regions. (a) Invariant positions in the tRNA internal promoter consensus (ref. 1) are underlined. (b) Sequence similarities are shown between the upstream region of RPRI and the S. cerevisiae U6 gene between the transcription start (+1) and the beginning of the solo delta repetitive element (-90). Position numbers indicate the first nucleotide of each sequence. In the case of RPRI sequences, the first position number corresponds to the numbering scheme relative to the first nucleotide of the smaller, mature RNA as in Fig. 1. The numbers in parentheses refer to the position relative to the 5' end of the leader RNA.

candidate internal promoter element to ensure inactivation. The variant genes were tested for complementation of an *RPR1* gene disruption and for levels of RNA production *in vivo* (Fig. 3).

Deletions of the upstream region as far as -130 had no effect on cell growth when the variant RPR1 genes were used to complement an RPR1 disruption in a haploid strain. A deletion to -69, to within the leader but not removing the A box, gave a slightly slow growth phenotype. However, a further deletion (to -51) that removes the A box was lethal. Double point mutations in the A box or B\* box were also lethal, whereas mutations in B1 and B2 gave normal growth phenotypes. Analysis of RNA synthesized from these RPRI variants gave parallel results, although the upstream region was found to substantially influence the efficiency of expression. Successive deletions of the 5' flanking region up to the putative A box reduced RNA expression by as much as 10-fold. However, further deletions that removed the A box, or point mutations in the B\* box, eliminated detectable expression, and a double point mutation in the A box reduced expression more than 20-fold. In contrast, mutations in either the B1 or the B2 box within the coding region had little effect on RPR1 RNA levels, and these small reductions might have been due to partial destabilization of the mature RNase P holoenzyme. Northern blot analysis of RPR1 RNA expressed from the variants (data not shown) indicated that the larger and smaller (mature form) RNAs were reduced or eliminated coordinately by mutations, consistent with their being derived from a single transcript.

These observations indicated that the A and B\* boxes in the 5' leader region were essential elements of the *RPR1* promoter. To further investigate whether these sequences behaved like the bipartite internal promoters of tRNA genes, binding of the transcription complex assembly factor, TFIIIC,

to RPR1 genes containing normal or mutated A and B\* boxes was examined. DNA mobility-shift experiments (Fig. 4) showed that TFIIIC bound to both the RPR1 gene and the SUP53 tRNA gene and that binding to either gene could be competitively inhibited by the other. As for the tRNA gene (1, 27), mutations in either the A or B\* elements severely reduced or eliminated TFIIIC binding. Mutations in the RPRI B1 or B2 elements had no effect on binding (data not shown).

The region of RPR1 to which TFIIIC bound was determined by DNase footprinting of both the sense and the antisense strand (Fig. 5). Stable TFIIIC footprints were not observed in the upstream or mature RNA coding regions (data not shown), but binding of TFIIIC protected most of the leader region, from at least -74 to -4 (9 bp upstream of the A box to 15 bp downstream of the B\* box). Both the extent of protection (open bars in Fig. 5) and the existence of intensified cleavages near the ends of the protected region (denoted by arrows) were comparable to the patterns seen with tRNA gene internal promoters.

## DISCUSSION

Efficiency of *RPR1* transcription is strongly influenced by upstream DNA sequences, but the essential pol III promoter elements correspond to an A box/B box combination in the 5' leader region that binds transcription assembly factor TFIIIC. This arrangement has the advantage that it can use an efficient pol III internal promoter without constraining the functional RPR1 RNA sequence by requiring that the promoter be retained in the mature form. The RPR1 5' leader region does not appear to be a bona fide tRNA gene, however, since its sequence cannot be folded into a cloverleaf structure and does not closely correspond to the sequence of any known yeast tRNA (unpublished observations). It is not known whether the leader RNA either has any distinct function in the cell or is related to any other uncharacterized RNAs.

Comparison of the RPRI A and B\* boxes to consensus Aand B-box sequences (1) shows that the A box is a good match but B\* has several differences, notably an adenine at an invariant thymine position (Fig. 6a). The B box is the primary binding site for TFIIIC and can normally bind the factor in the absence of the A box, although at reduced efficiency (ref. 28; see also Fig. 4). The departure of  $B^*$  from the consensus might make binding of TFIIIC more dependent on the A box, which would explain why the A-box point mutations in RPR1 confer a severe defect in TFIIIC binding (Fig. 4).

It is not clear what general features or precise sequences in the RPR1 upstream region influence expression, although the progressive decrease in RNA production with the 5' deletions argues that there might be several positive control signals. A comparison was made between this region and the upstream region of the S. cerevisiae U6 gene to -90 (Fig. 6b), where a delta repetitive element begins (13, 29). U6 transcription appears to be controlled primarily by A-box and B-box internal promoters in vivo, but it is not clear what quantitative effects might be conferred by the immediate upstream region. This region of the U6 gene is of interest because a "TATA" element at -30 and a proximal sequence element at -59resemble putative control regions of mammalian U6 genes and because in vitro transcription of the yeast gene is strongly influenced by the upstream region (13, 27). A number of striking sequence homologies were found, often at nearly corresponding distances upstream from the internal A box. U6 has a TATAATA box at -30, consistent with the observation that factor TFIID can stimulate U6 transcription in vitro (30). Although no classical TATA element was found in RPRI, the -31 CTATA sequence from U6 is present as direct repeat at -28 and -38 relative to the start of the *RPR1* leader. The -59 AAAGTTATTCG proximal sequence element common to mammalian and U6 genes also overlaps two strong U6/RPRI similarities at -75 and -261 relative to the RPRI leader. The thymine-rich element at -18 has been suggested as a potential U6 regulatory site (27), but the sequence similarities at -48 and -87 (U6 positions) have not previously been noted through comparison with the mammalian genes or investigated for functional significance. Further investigation will be required to elucidate how control mechanisms that might operate through these or other upstream sequences cooperate with the required internal promoters to increase expression of *RPR1*.

We thank J. Warner, P. Thuriaux, A. Sentenac, and E. Esposito for cloned DNA and yeast strains. We thank D. Kindelberger, C. Rohlman, M. Hull, and G. Butler for critical reading of the manuscript. This work was supported by National Institutes of Health Grant GM34869 and National Science Foundation Grant DMB-8901559 to D.R.E.

- 1. Geiduschek, E. P. & Tocchini-Valentini, G. P. (1988) Annu. Rev. Biochem. 57, 873-914.
- Murphy, S., Moorefield, B. & Pieler, T. (1989) Trends Genet. 5, 2. 122-126.
- 3. Huibregtse, J. M. & Engelke, D. R. (1989) Mol. Cell. Biol. 9, 3244-3252.
- Kassavetis, G. A., Riggs, D. L., Negri, R., Nguyen, L. H. & 4. Geiduschek, E. P. (1989) Mol. Cell. Biol. 9, 2551-2566.
- Kassavetis, G. A., Braun, B. R., Nguyen, L. H. & Geiduschek, E. P. (1990) Cell 60, 235-245.
- Murphy, S., DiLiegro, C. & Melli, M. (1987) Cell 51, 81-87. 6.
- Murphy, S., Tripodi, M. & Melli, M. (1986) Nucleic Acids Res. 14, 7. 9243-9261.
- Waibel, F. & Filipowicz, W. (1990) Nature (London) 346, 199-202. Krol, A., Carbon, P., Ebel, J.-P. & Appel, B. (1987) Nucleic Acids Res. 15, 2463-2478.
- 10. Kunkel, G. R., Maser, R. L., Calvet, J. P. & Pederson, T. (1986) Proc. Natl. Acad. Sci. USA 83, 8575-8579
- 11. Reddy, R., Henning, D. & Das, G. (1987) J. Biol. Chem. 262, 75-81.
- Bark, C., Weller, P., Zabielski, J., Janson, L. & Pettersson, U. 12. (1987) Nature (London) 328, 356-359.
- 13.
- Brow, D. A. & Guthie, C. (1990) Genes Dev. 4, 1345-1356. Lee, J.-Y., Rohlman, C. E., Molony, L. E. & Engelke, D. R. (1991) 14. Mol. Cell. Biol. 11, 721-730.
- 15. Baer, M., Nilsen, T. Y., Costigan, C. & Altman, S. (1990) Nucleic Acids Res. 18, 97-101.
- Gudenus, R., Mariotte, S., Moenne, A., Ruet, A., Memet, S., 16. Buhler, J.-M., Sentenac, A. & Thuriaux, P. (1988) Genetics 119, 517-526.
- 17. Ito, H., Fukuda, Y., Murata, K. & Kimura, A. (1983) J. Bacteriol. 153, 163-168.
- Sherman, F., Fink, G. R. & Hicks, J. B. (1986) Methods in Yeast 18. Genetics (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- 19. Ausubel, F. M., Brent, R., Kington, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1987) Current Protocols in Molecular Biology (Wiley, New York). Lee, J.-Y. & Engelke, D. R. (1989) Mol. Cell. Biol. 9, 1058-1064.
- 20
- Hanning, E. M., Thiele, D. J. & Leibowitz, M. J. (1984) Mol. Cell. 21. Biol. 4, 101-109.
- Huibregtse, J. M. & Engelke, D. R. (1986) Gene 44, 151-158.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular 23. Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- Evans, C. F. (1990) Ph.D. thesis (The University of Michigan, Ann 24.
- Arbor). Newman, A. J., Ogden, R. C. & Abelson, J. (1983) Cell 35, 117-125. 25.
- Strobel, M. C. & Abelson, J. (1986) Mol. Cell. Biol. 6, 2674-2684.
- 27. Moenne, A., Camier, C., Anderson, G., Margottin, F., Beggs, J. & Sentenac, A. (1990) EMBO J. 9, 271-277
- Baker, R. E., Gabrelsen, O. & Hall, B. D. (1986) J. Biol. Chem. 261, 28. 5275-5282.
- Brow, D. A. & Guthrie, C. (1988) Nature (London) 334, 213-218. 29.
- 30. Margottin, F., Dujardin, G., Gerard, M., Egly, J.-M., Huet, J. & Sentenac, A. (1991) Science 251, 424-426.