EUKARYOTIC CELL, Jan. 2009, p. 69–76 Vol. 8, No. 1
1535-9778/09/$08.00+0 doi:10.1128/EC.00254-08

Promoter Analysis of Palindromic Transcription Units in the Ribosomal DNA Circle of *Entamoeba histolytica*▽

Sunil K. Panigrahi,[1] Gagan Deep Jhingan,[1] Indrani Som,[2] Alok Bhattacharya,[3] William A. Petri, Jr.,[4] and Sudha Bhattacharya[1]*

School of Environmental Sciences, Jawaharlal Nehru University, New Mehrauli Road, New Delhi 110 067, India[1]; LCDB/NIDDK, National Institutes of Health, Bethesda, Maryland 20892[2]; School of Life Sciences, Jawaharlal Nehru University, New Mehrauli Road, New Delhi 110 067, India[3]; and Division of Infectious Diseases and International Health, University of Virginia, Charlottesville, Virginia 22908-1340[4]

Received 29 July 2008/Accepted 23 October 2008

rRNA genes of *Entamoeba histolytica* are organized as palindromic ribosomal DNA (rDNA) units (I and II) in a 24.5-kb circle. Although the two rDNAs are identical in sequence, their upstream spacers are completely different. Since the intergenic sequences (IGS) of all rDNA copies in other organisms are conserved and contain transcription regulatory sequences, the lack of sequence conservation in the IGS prompted the question of whether both rDNAs are indeed transcriptionally active. We mapped the transcriptional start points (tsp's) and promoters of the two rDNAs. A 51-bp sequence immediately upstream of the tsp's was highly conserved in both units. In addition, both units had an A+T-rich stretch upstream of the 51-bp core. Analysis of reporter gene transcription showed promoter activity to reside in the regions from positions −86 to +123 (rDNA I) and positions −101 to +140 (rDNA II). The promoter-containing fragments from both units could bind and compete with each other for protein(s) from nuclear extracts. Protein binding was especially dependent on the A+T-rich region upstream of the 51-bp core (positions −53 to −68). The requirement of >80 bp downstream of the tsp was striking. Although this sequence was not conserved in the two units, it could potentially fold into very long stem-loops. Both rDNAs transcribed with comparable efficiency, as measured by nuclear runon. Thus, both rDNAs share very similar organization of promoter sequences, and in exponential culture both rDNAs are transcribed. It remains to be seen whether the different IGS affect the regulation of the two units under adverse conditions.

Cell division rate is intimately linked with cellular ability to synthesize ribosomes. Many oncoproteins and tumor suppressor proteins are known to modulate ribosome biogenesis (11). This control is primarily exercised at the level of rRNA gene transcription (9, 20). Studies on this important regulatory mechanism are currently limited to a few model systems. Since ribosomes perform a highly conserved function in all cell types, it will be interesting to understand the diversity of mechanisms that nature may have evolved to regulate ribosome biogenesis in organisms adapted to a variety of niches. The unicellular parasitic protist *Entamoeba histolytica* resides in the human gut. Its rRNA genes are organized in an unusual manner on extrachromosomal circular molecules (2). Some strains of *E. histolytica* (e.g., HK-9) contain only one transcription unit per circle, while in other strains (e.g., HM-1:IMSS) each 24.5-kb circle (EhR1) contains two palindromic ribosomal DNA (rDNA) units (rDNA I and II) separated by upstream and downstream intergenic spacers (IGS) (Fig. 1). While the coding sequence of the two rDNAs of EhR1 is identical, the sequence of the upstream IGS and external transcribed spacers (ETS) of the two units is completely different. Strains with only one rDNA unit typically contain the unit termed "rDNA I" in EhR1. No strains with rDNA II alone have yet been found (2). Hence, sequences corresponding to the PvuI and ScaI repeats of the upstream IGS of rDNA II and the transcript (Tr) of the rDNA II ETS (Fig. 1) are missing in strains such as HK-9. IGS contain important sequences required for transcription regulation (core promoters, spacer promoters, enhancers, and terminators) (18, 19). The lack of sequence conservation in the rDNA I and II IGS in EhR1 is therefore significant. The transcription start point (tsp) of the single rDNA unit of strain HK-9 has been mapped (10). The tsp's of rDNA I and II in EhR1 have not been mapped, and it is not known whether both units are transcriptionally active, especially since they are so different in their IGS and ETS sequences.

Here we show that both rDNA transcription units of EhR1 are active and share a 51-bp conserved core promoter sequence. We have further identified the important promoter elements by mutation analysis and find that, in addition to sequences upstream of the tsp, both units require downstream sequences also for promoter activity.

MATERIALS AND METHODS

***E. histolytica* culture.** Trophozoites of strains HM-1:IMSS and Rahman were axenically cultured in TYI-S-33 medium at 35.5°C (4).

Primer extension. Assays were carried out by using mMuLV reverse transcriptase (Pharmacia) and an oligonucleotide (5′-CAATACTACGTACTACCT ACTTAT-3′) complementary to the nucleotides at positions +115 to +139 with respect to the tsp of rDNA II. A portion (40 μg) of total *E. histolytica* RNA was incubated with γ-^{32}P-labeled oligonucleotide (6×10^5 cpm). Annealing was carried out at 65°C for 2 min, followed by 37°C for 30 min and extension at 37°C for 90 min with 200 U of reverse transcriptase. The products were separated on

* Corresponding author. Mailing address: School of Environmental Sciences, Jawaharlal Nehru University, New Mehrauli Road, New Delhi 110 067, India. Phone: 91-11-26704308. Fax: 91-11-26172438. E-mail: sb@mail.jnu.ac.in.

▽ Published ahead of print on 31 October 2008.

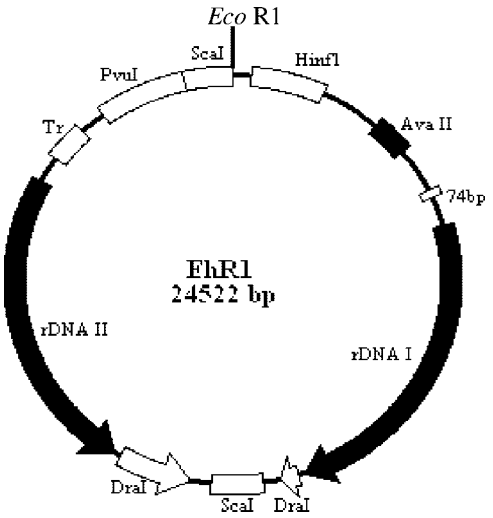

FIG. 1. Sequence organization of the *E. histolytica* HM-1:IMSS rDNA plasmid, EhR1 (17). The two inverted rDNA repeats (5.9 kb each) are indicated by filled arrows. The EcoRI site at the top of the figure corresponds to nucleotide 1. Various families of short tandem repeats in the regions upstream and downstream of the rDNAs are marked as PvuI, ScaI, HinfI, AvaII, 74bp, and DraI. The Tr region upstream of the rDNA II codes for a 0.7-kb transcript.

denaturing 6% urea-polyacrylamide gels, together with the sequencing reaction using the same oligonucleotide (15).

RNase protection assay. Fragments encompassing the putative tsp's were cloned in pGEM-T vector for in vitro transcription. The PCR primer pairs used to generate the desired fragments were as follows: for rDNA I, 5′-TTGCCGAGAAAAGATGACGT-3′ and 5′-TTGGTACCCCTCATTCTATTCA-3′ were used for the fragment spanning positions −278 to +115 and 5′-TTGCCGAGAAAAGATGACGT-3′ and 5′-ATTTTGAAAAATCCACATCAAAC-3′ were used for the negative control (positions −278 to −55); and for rDNA II, the fragments spanning positions −182 to +141 (5′-TGATACATGTGAATTGTTGT-3′ and 5′-CAATACTACGTACTACCTACTTAT-3′) were used and, for a negative control, the fragments spanning positions −182 to −28 (5′-TGATACATGTGAATTGTTGT-3′ and 5′-ATTTGACTCTTTGTAGCTAACGTTTT-3′) were used. Cloned fragments were transcribed in antisense orientation (with respect to rRNA) in the presence of 20 μCi of [^{32}P]UTP using the T3/T7 in vitro transcription system (Roche). The RNA probes (393 nucleotides for rDNA I and 323 nucleotides for rDNA II) were purified on a 5% denaturing polyacrylamide gel. Total *E. histolytica* RNA (10 μg) was hybridized at 42°C for 16 h overnight with 90,000 cpm of the probes and then digested with an RNase A and RNase T_1 mix at 37°C for 1 h as described in the Ambion RPAIII kit (Ambion, Austin, TX). Transcripts were also obtained from control fragments corresponding to regions upstream of the tsp's and similarly processed. The protected fragments, along with the sequencing reaction (rDNA I), were separated on a 5% denaturing polyacrylamide gel and visualized by autoradiography.

Methyl guanosine capping assay. Total RNA (40 to 50 μg) was added to 50 μl of capping buffer (50 mM Tris-HCl [pH 7.9], 1.25 mM $MgCl_2$, 60 mM KCl, 0.1 mM *S*-adenosylmethionine, 2.5 mM dithiothreitol, 1 mg of acetylated bovine serum albumin/ml, RNase inhibitor [Amersham] at 500 U/ml), 100 μCi of [^{32}P]GTP, and 8 U of guanylyltransferase (7). After incubation for 1 h at 37°C, the mixture was phenol extracted, and RNA was ethanol precipitated, centrifuged, and washed with 80% ethanol. The pellet was dissolved in hybridization buffer (80% deionized formamide, 1 mM EDTA, 40 mM PIPES [pH 7.4], and 0.2 M sodium acetate [pH 5.2]). Five different fragments from rDNA (as indicated in Fig. 2) were cloned in pBS vector and in vitro transcribed by the T3/T7 in vitro transcription system (Roche). Transcribed products were phenol extracted and ethanol precipitated. The pellet was dissolved in hybridization buffer. Both in vitro-capped and in vitro-synthesized RNAs were pooled, denatured at 85°C for 5 min, and hybridized at 37°C for 16 h, followed by RNase One digestion for 1 h. The reaction was stopped (by adding 20 μl of 10% sodium dodecyl sulfate) and ethanol precipitated. The pellet was dissolved in formamide gel loading buffer and loaded on 5% urea-polyacrylamide gels.

Luciferase reporter constructs. To assay for rDNA I and II promoter activity, deletion constructs were cloned in the promoterless vector pBS-Luc-Dra, which contained 1.7 kb of firefly luciferase sequence cloned upstream of a 0.6-kb fragment containing 40 nucleotides of the 3′ end of *E. histolytica* large subunit rRNA, followed by the downstream sequence which includes DraI repeats (17). All of the deletion fragments were PCR amplified from genomic DNA and inserted at EcoRI and KpnI sites upstream of luciferase. The end points of all of the deletion fragments with respect to the tsp are indicated in Fig. 3. The orientation and sequence of each construct was confirmed by restriction digestion and DNA sequence analysis.

Transient-transfection and reporter assay. Cells were transfected by electroporation (12) and transferred to plastic flasks (Techno Plastic Products, Switzerland) containing 45 ml of TYI-S-33 medium, followed by incubation at 36.5°C for 30 h. Total RNA was isolated from the transfected trophozoites using Tripure reagent (Roche) and treated with DNase I. Reverse transcription-PCR (RT-PCR) was carried out with luciferase-specific primers (5′-CGCCCTGGTTCCTGGAAC-3′ and 5′-GAGAATCTGACGCAGGCAGTTC-3′) spaced 290 nucleotides apart.

Gel shift assays. End-labeled DNA fragments (0.5 to 1 ng, up to 20,000 cpm) were incubated with 15 μg of nuclear extract, 1 μg of poly(dI-dC), and 10% glycerol in the presence of DNA-protein binding buffer (12 mM HEPES [pH 7.9],150 mM KCl, 1 mM dithiothreitol, 1 mM EDTA, 4 mM Tris-HCl [pH 7.9], 1 mM spermidine, and 1.5 mM $MgCl_2$) for 15 min at room temperature (13). The bound and unbound complexes were separated on 7% nondenaturing polyacrylamide gels in 0.5× TBE (44.5 mM Tris-HCl [pH 7.9], 44.5 mM boric acid, and 1 mM EDTA) at room temperature (~25°C) and 200 V for 4 h. After electrophoresis, the gel was dried and visualized by autoradiography. Competition assays were performed using a 20- or a 100-fold excess of unlabeled competing fragments. Sheared calf thymus DNA was used as the nonspecific competitor (350-fold excess).

Preparation of nuclei and elongation of nascent RNA by nuclear runon. Nuclei were prepared essentially as described previously (3). About 2×10^7 trophozoites were harvested, washed twice with cold phosphate-buffered saline, and resuspended in 40 volumes of lysis buffer (10 mM Tris-HCl [pH 7.4], 3 mM $CaCl_2$, and 2 mM $MgCl_2$). After incubation for 10 min on ice, cells were harvested by centrifugation with 500 × *g* for 5 min at 4°C and resuspended in 5 volumes of lysis buffer supplemented with 1% NP-40. Lysis of cells and the integrity of nuclei were monitored by phase-contrast microscopy. The nuclear pellet was washed several times in lysis buffer without NP-40 (10). Purified nuclei were incubated at 37°C in a buffer consisting of 80 mM Tris-HCl (pH 7.5), 50 mM KCl, 50 mM NaCl, 8 mM $MgCl_2$, 0.5 mM spermidine, 0.15 mM spermine, 5 mM dithiothreitol, 50% glycerol, 8 mM ATP, 4 mM CTP, 4 mM GTP, 20 μM UTP, 300 U of RNA Guard (Promega)/ml, and 250 μCi of [^{32}P]UTP (3,300 Ci/mmol) for 7 min. The radioactively labeled RNA was isolated by using Tripure (Roche). The total *E. histolytica* genomic DNA and various cloned DNAs were purified, denatured, and applied to nylon membrane by using a dot blot apparatus. Hybridization with nascent RNA was performed as described previously (14). Posthybridization washing of the membrane was performed sequentially for 20 min each time with 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at room temperature, 2× SSC supplemented with 0.1% SDS at 65°C, and 0.1× SSC at room temperature.

RESULTS

Mapping the tsp of rDNA I and II. Primer extension analysis was performed to map the tsp's. The tsp of rDNA I is already mapped in strain HK-9 to lie 2.67 kb upstream of the mature 18S rRNA (10). We confirmed that the rDNA I tsp in strain HM-1:IMSS also mapped to the same position (data not shown). We further mapped the tsp of rDNA II to a position 1.224 kb upstream of the mature 5′ end of 18S rRNA by primer extension (Fig. 2A). The nucleotide sequence of the region in which each of the tsp's mapped is highly conserved in the *E. histolytica* strains sequenced in our lab, although polymorphism exists in the numbers (but not sequence) of tandem repeats that are found further upstream and downstream of the tsp's (2, 17). The tsp assignment of both rDNA units was further confirmed by RNase protection (Fig. 2B) and by methyl guanosine capping of 5′-NTP (Fig. 2C) using guanylyl

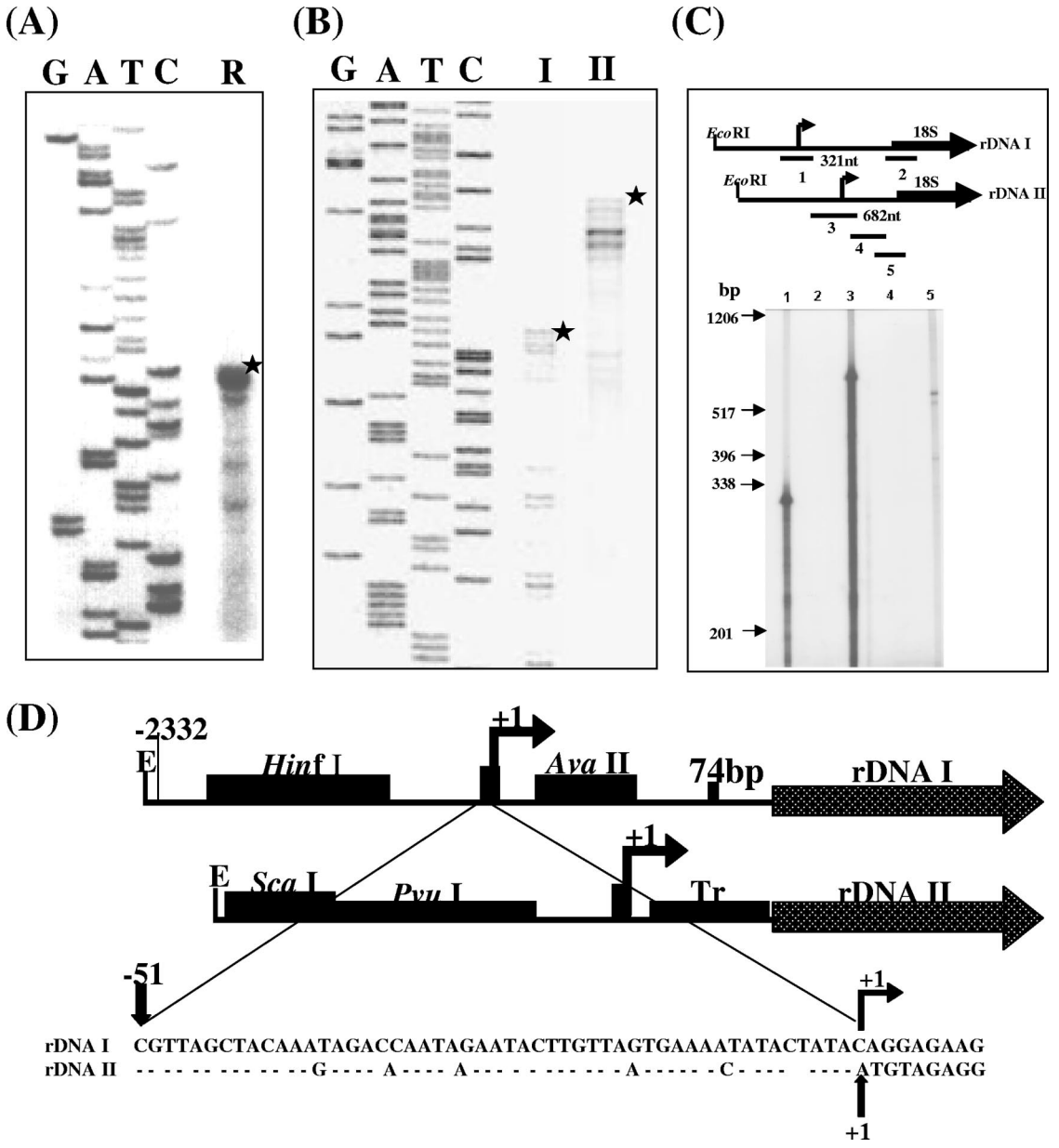

FIG. 2. Identification of tsp's and putative core promoters of rDNA I and II. (A) Primer extension products obtained with a primer located upstream of the expected tsp of rDNA II were analyzed by electrophoresis alongside a sequencing ladder extended with the same primer. Lane R shows the primer-extended products. The size of the largest extended product was 139 nucleotides. (B) Tsp assignment was done by RNase protection assay. The sizes of protected fragments were determined by running a sequencing reaction (using primer for rDNA I) alongside. The size of the largest protected fragment (indicated by an asterisk) for rDNA I (lane I) was 115 nucleotides, while that for rDNA II (lane II) was 141 nucleotides. Fragments upstream of the tsp's were also used for RNase protection as detailed in Materials and Methods. These fragments did not give any protection (not shown). The tsp for each rDNA, as inferred from these data, is indicated by an arrow in panel D. (C) Methyl guanosine capping confirmed the tsp assignment. Total RNA was labeled with [α-^{32}P]GTP and guanylyl transferase and hybridized with in vitro-transcribed RNA corresponding to regions 1 to 5 in rDNAs I and II, as indicated. The single-stranded RNA remaining after hybridization was removed by treatment with RNase One, and the sizes of the protected fragments were determined by denaturing polyacrylamide gel electrophoresis. Lanes 1 to 5 correspond to fragments protected by regions 1 to 5, respectively. The expected size of the fragment protected in lane 1 is 321 nucleotides, and that in lane 3 is 682 nucleotides. (D) Sequence homology in a 51-nucleotide stretch (putative core promoter) upstream of the tsp's (arrow at position +1) in rDNA I and II.

transferase. The tsp assignment of rDNA I was identical by both primer extension and RNase protection. For rDNA II, the tsp assignment by RNase protection compared to primer extension was shifted by two nucleotides, that is, 1.226 kb upstream of the mature 5′ end of 18S rRNA. Upon manual inspection of the sequence immediately upstream of the tsp's in rDNA I and II (assigned by RNase protection) it was found that a 51-nucleotide sequence was highly conserved (86% identity; Fig. 2D). This is significant since the rest of the sequence upstream and downstream of the tsp's in the two rDNAs is very

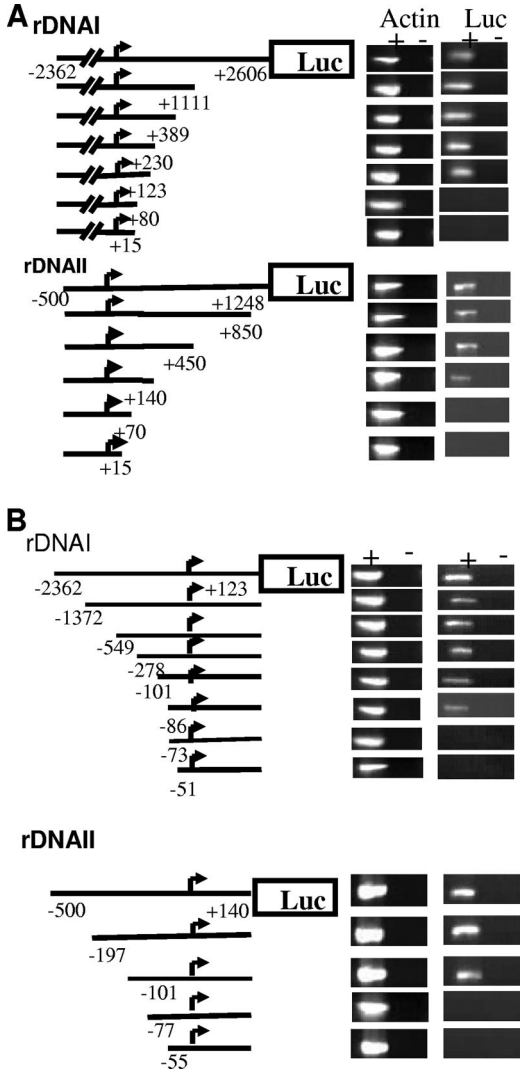

FIG. 3. Mapping the promoter of rDNA I and rDNA II by measuring transcription of reporter constructs in transiently transfected cells. Deletions of sequences adjoining the tsp (indicated by a bent arrow) were cloned upstream of luciferase (Luc) gene. Deletions downstream of the tsp are indicated in panel A, and those upstream are indicated in panel B. The endpoints of each fragment are indicated with respect to the tsp. Plasmid constructs containing each deletion fragment were introduced into cells by electroporation, and transcription was measured by RT-PCR with luciferase-specific primers (as described in Materials and Methods). PCR was performed in the presence (+) or absence (−) of reverse transcriptase. RT-PCR was also performed with actin primers as an endogenous control for RNA.

divergent (17). This 51-nucleotide sequence could correspond with the core promoter, by analogy with rRNA genes in other systems.

Deletion analysis to define the promoter fragments of both rDNAs that support reporter gene transcription. To map the functional promoter, deletion constructs of DNA sequences adjoining the putative core promoter were made. Since proper transcription termination may be required for efficient initiation, we decided to introduce the sequence found downstream of the 28S rRNA (including the DraI repeats, Fig. 1) into our reporter construct, downstream of the luciferase gene. For this purpose, the luciferase gene and DraI repeat sequence were introduced at KpnI/BamHI and BamHI/PstI sites, respectively, in the MCS of $pBS^{+/-}$ phagemid vector. Different deletion fragments to be tested for promoter activity were cloned upstream of luciferase reporter sequence using EcoRI/KpnI sites, and the plasmids transiently transfected into *E. histolytica*. Total RNA was isolated from transfectants, treated with DNase I and amplified by RT-PCR with gene-specific primers for luciferase (reporter gene) and actin (endogenous control for RNA). (No luciferase enzyme activity was detectable in the transfected cells, possibly because the polymerase I (Pol I) transcript could not be suitably modified for translation. The same luciferase gene sequence when cloned downstream of *E. histolytica* lectin promoter gave 2,000 U of luciferase activity per mg of cell lysate in transiently transfected cells.)

Results of RT-PCR showed that a region of similar length, spanning nucleotides −86 to +123 (in rDNA I) and nucleotides −101 to +140 (in rDNA II) was required to support luciferase transcription (Fig. 3). Genomic PCR was done with DNA from all samples to confirm that transfection had indeed occurred efficiently in samples scoring negative in RT-PCR. In addition, at least three independent transfections were done for each deletion and similar results were obtained. These data show that, apart from the conserved 51-nucleotide putative core promoter, the sequence between positions −51 to −86 (for rDNA I) and positions −51 to −101 (for rDNA II) was also required for transcription. Although this sequence is not conserved in the two units, it is highly A+T-rich in both rDNAs (83% A+T in rDNA I and 92% A+T in rDNA II). Sequences downstream of tsp were also essential for promoter activity. Transcription was not found in constructs that included only +80 of the downstream region.

The rDNA I and II promoters transcribe with comparable efficiencies in their native location. In the transient-transfection analysis with the luciferase reporter, the rDNA promoters were removed from their native location in the rDNA circle. To compare transcription efficiency of the two rDNA promoters in their native location, nuclear runon analysis was performed. Radioactive RNA, labeled in the nuclear runon reaction, was hybridized with ETS probes from the two rDNAs (Fig. 4). Probes from both units hybridized with comparable intensity. The hybridization intensity obtained was proportional to the probe size. For example, probe 1B (389 bp) gave less signal than 2B (454 bp), while probe 1C (1111 bp) was brighter than probe 2C (855 bp). An average of five independent measurements showed similar hybridization intensity of the two rDNAs. Thus, no significant difference was found in the transcription efficiency of the two rDNA promoters.

Both promoters bind and compete for similar nuclear proteins in a gel shift assay. The assembly of the preinitiation complex of RNA Pol I takes place at the rRNA gene promoter, in conjunction with various transcription factors. Therefore, it is expected that promoter-containing DNA fragments should bind to specific proteins in nuclear extracts. To demonstrate this, radiolabeled double-stranded DNA probes from promoters of both rDNAs (positions −101 to +123 for rDNA I and positions −101 to +140 for rDNA II) were incubated with total nuclear extract, and a gel shift assay was performed. A prominent shift was observed with the rDNA I promoter fragment, which was efficiently competed with excess cold DNA of

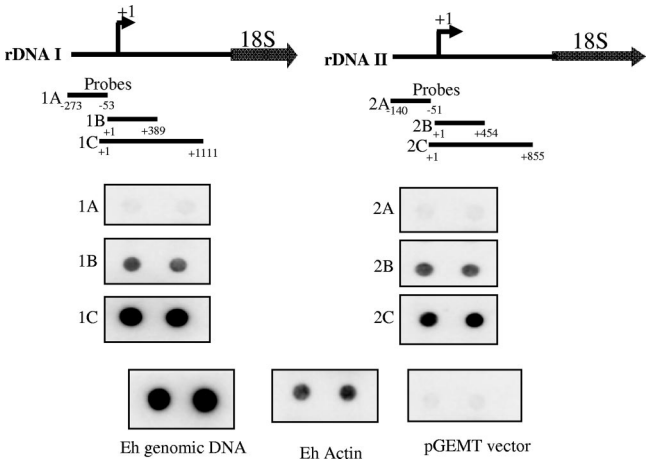

FIG. 4. Measurement of rDNA I and II transcription by nuclear runon. RNA was transcribed in nuclear extracts in the presence of 250 μCi of [α-^{32}P]UTP. RNA was hybridized with DNA probes cloned in pGEM-T vector (spotted in duplicate), corresponding to regions upstream and downstream of tsp, as indicated (probes 1A to C and probes 2A to C). *E. histolytica* genomic DNA and actin DNA served as positive controls, and pGEM-T vector DNA alone was the negative control.

rDNA I promoter itself and rDNA II promoter (data not shown). The competition was specific since DNA from a downstream region of rDNA I (positions +101 to +230) and sheared calf thymus DNA had no effect. A reciprocal experiment was performed with labeled rDNA II promoter as a probe. This promoter also showed a prominent shift when incubated with nuclear extract. The shift was competed with self-DNA and rDNA I promoter region, while there was no competition with sheared calf thymus DNA. This shows that both promoters bind to nuclear protein(s) and can compete with each other in this binding.

Thus, promoters of both rDNA I and II can support transcription and compete for binding to similar nuclear proteins.

Mapping the protein binding region in the rDNA I promoter. It is obvious from the gel shift analysis that one or more nuclear proteins bound specifically to the rDNA promoter fragments under our assay conditions. To determine the minimum region involved in DNA-protein interaction, we performed gel shift assays using various deletions of the rDNA I promoter (Fig. 5A). Deletion 1 (positions −101 to +15; Fig. 5A, lanes 5 and 6) competed with the total promoter (positions −101 to +123), whereas deletion 2 (positions +1 to +123, lanes 7 and 8) did not. Further deletions of region from −101 to +15 were done from both the 5′ and the 3′ directions. In Fig. 5B, deletions from −73 to +15 (lanes 5 to 8) could compete, while those from −60 to +15 (lanes 9 and 10) could not. From the 3′ end (Fig. 5C), deletion of −73 to −32 (lanes 7 and 8) could compete, but further deletions (−73 to −42; lanes 9 and 10) could not compete. Hence, a minimum region of rDNA I promoter essential for protein binding in our gel shift assay condition was located between positions −73 and −32 with respect to the tsp. The same region from rDNA II (positions −77 to −32) could compete for protein binding with rDNA I (lanes 15 and 16; Fig. 5C). Mutation analysis of this region of rDNA I was done to determine the nucleotides

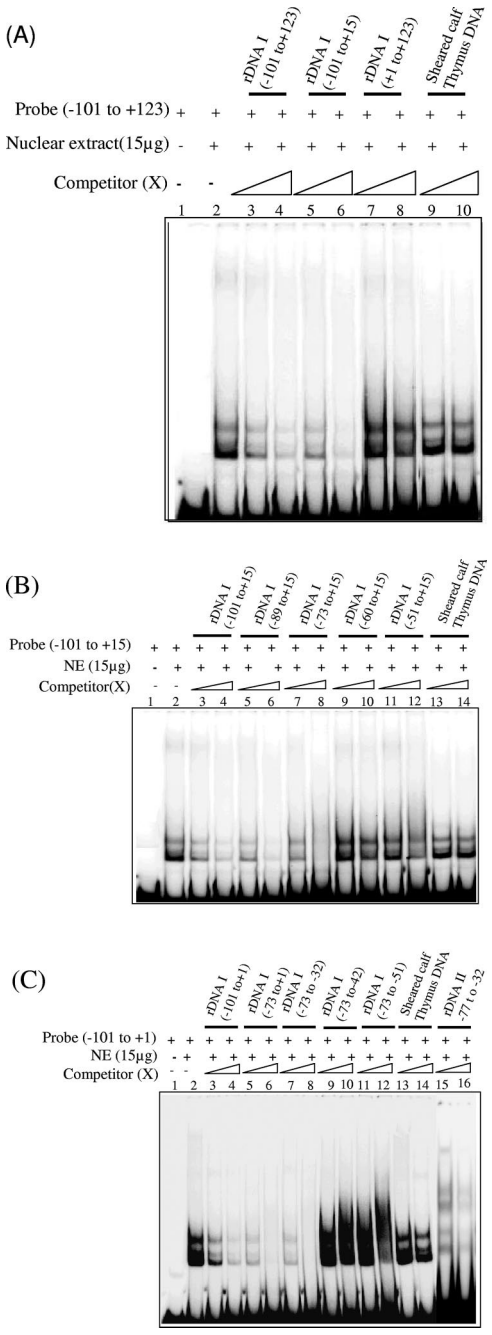

FIG. 5. Gel shift competition assay to map the minimum region of rDNA I promoter needed for protein binding. DNA probes from the promoter of rDNA I (positions −101 to +123 [A], positions −101 to +15 [B], and positions −101 to +1 [C]) were incubated with nuclear extract (NE). A gel shift assay was performed in the presence of 20- and 100-fold excesses of the indicated unlabeled competitors. In panel C, rDNA II fragment (positions −77 to −32) was used as a competitor (lanes 15 and 16).

essential for DNA-protein interaction. Site-directed mutants were generated by introducing base changes (A→C, T→G, and vice versa) in contiguous stretches of five nucleotides. A gel shift assay showed that protein binding was actually enhanced when the sequence between positions −73 to −69 was

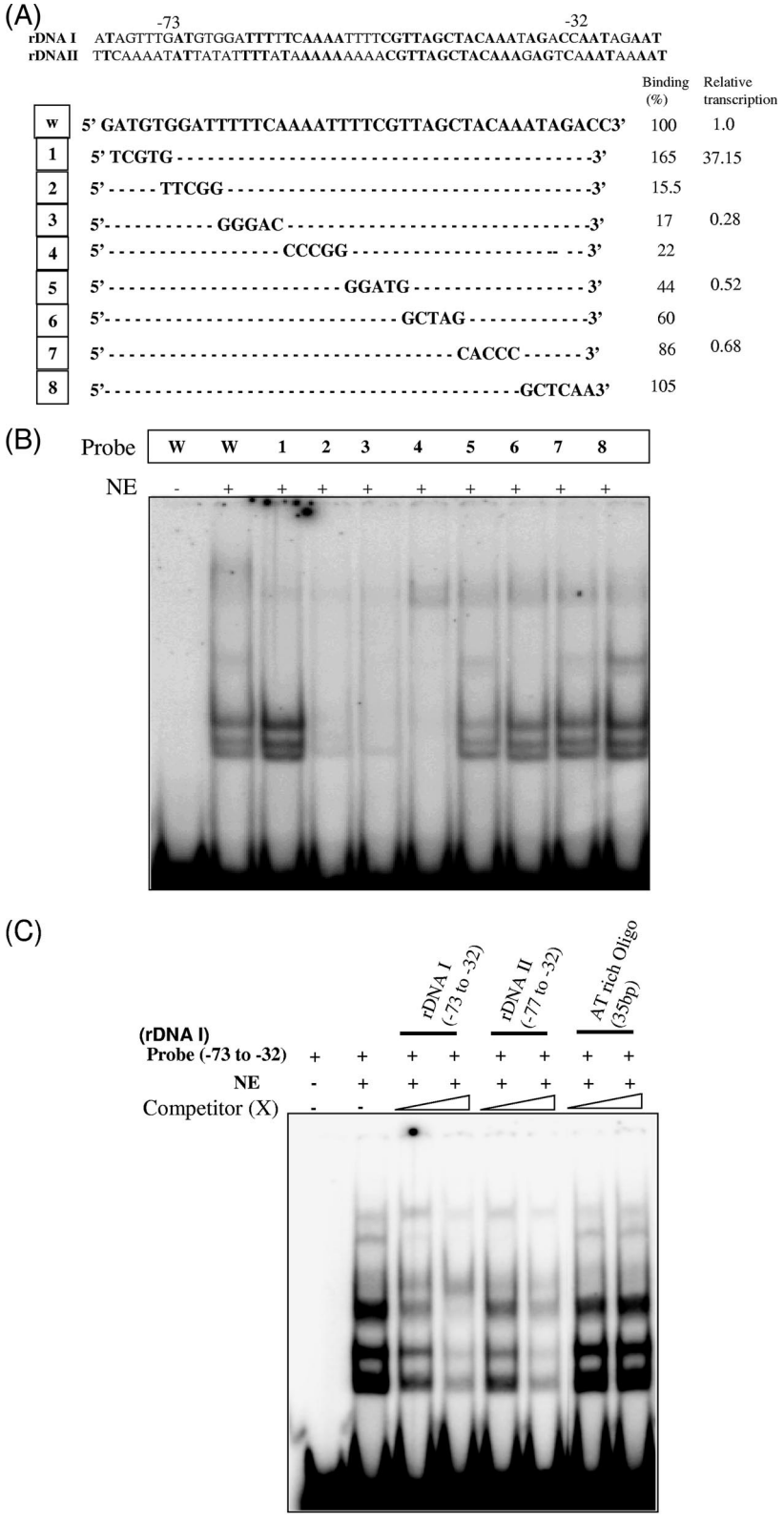

FIG. 6. Gel shift assay with promoter mutants and (A+T)-rich oligonucleotide. (A) The sequences of the minimum fragment required for protein binding (positions −73 to −32) in rDNA I (and its counterpart in rDNA II) are shown at the top. Conserved nucleotides are in boldface. The sequence from positions −73 to −32 of rDNA I (marked W) was mutated (A→C, T→G, and vice versa) a stretch of five nucleotides at a time, as shown. (B) A gel shift assay was done with the wild-type probe and all of the mutants. W, wild-type probe; lanes 1 to 8, mutations 1 to 8, respectively. The binding percentage was calculated with reference to the W probe and is shown in panel A. The wild-type and mutant oligonucleotides were cloned upstream of luciferase reporter gene and introduced into *E. histolytica* by transient transfection (as in Fig. 3). The levels of luciferase RNA were measured by quantitative PCR. The relative transcription efficiency of mutants with respect to W is given in panel A. (C) A gel shift assay was performed with the W probe and competed with 20- and 100-fold excesses of rDNA II (positions −77 to −32) and a 35-bp oligonucleotide composed entirely of A's and T's. The sequence of the oligonucleotide was 5′-TAATAATAATAATA ATAATTAAATAATAATAATAA-3′.

mutated, while mutating the sequence between positions −68 to −53 reduced the protein binding very significantly (Fig. 6). The percent binding values were calculated from an average of three independent gel shift experiments. The transcription efficiency of these mutants was determined by cloning the mutant promoters upstream of luciferase reporter gene as in Fig. 3. Transcription was measured by quantitative PCR of luciferase RNA from transiently transfected cells. Again, three independent measurements were made. A direct correlation was observed between protein binding and relative transcription efficiency (Fig. 6A). This experiment showed that nucleotides between positions −68 and −53 (5′-GGATTTTTCAAAATT-3′) are crucial in the DNA-protein interaction. Since this sequence is highly (A+T)-rich, it is possible that the nuclear proteins binding to this region are those that generally recognize such sequences. To check this, a 35-bp oligonucleotide composed exclusively of A's and T's was used in the gel shift competition assay. However, a 100-fold excess of this oligonucleotide failed to show any competition (Fig. 6C).

DISCUSSION

E. histolytica rRNA genes show two novel features not commonly found in other systems. First, these genes are exclusively located on extrachromosomal circular molecules, and in many strains each circle has two copies arranged as palindromes. Second, the IGS and ETS sequences, which in other systems contain important regulatory elements, are completely different in the two palindromic copies. However, the hundreds of rDNA copies in the cell are all identical, and there is no sequence divergence in the IGS and ETS sequences from various *E. histolytica* strains (2). Therefore, the sequence differences seen in the two palindromic copies are stably inherited and fixed in the population. Thus, the regulation of rRNA gene transcription in this parasite is likely to have some unique features. To explore this, it is necessary to first define the core promoter sequences required for transcription. Here we have identified the core promoter of both rDNA units and show that under exponential conditions of growth both units are transcriptionally active. The general organization of regulatory sequences of rRNA genes follows a common theme in most eukaryotes. *E. histolytica* also displays some of these conserved features. For example, (i) the tsp maps 1 to 2 kb upstream of the mature 18S rRNA (2.67 kb in rDNA I and 1.224 kb in rDNA II); (ii) the IGS of both rDNAs contain tandem repetitive elements (Fig. 1); and (iii) the promoter is located within 100 nucleotides upstream of the tsp. From deletion analysis we could assign the upstream limit of the rDNA I promoter at position −86 and rDNA II promoter at −101. Of this, the region from position −1 to position −51 is highly conserved in the two rDNAs. Sequences further upstream (from positions −68 to −53) are crucial for binding of nuclear protein/s. In rDNA I the sequence between −52 and −66 is entirely composed of tracts of A's and T's, with only one C residue. In rDNA II the sequence between positions −52 and −74 is entirely composed of tracts of A's and T's. Sequence changes in this region abolish DNA-protein interaction. However, this is probably not because nuclear proteins directly bind to these (A+T)-rich sequences, since a DNA fragment composed entirely of A's and T's did not compete in the gel shift (Fig. 6C). It is possible that this region may adopt a favorable conformation required for DNA-protein interaction further downstream. A- and T-rich tracts are associated with bent DNA. rRNA promoter upstream sequences in a variety of organisms, including *Physarum polycephalum* (16), *Arabidopsis thaliana* (8), *E. coli* (23), and humans (21), have been shown to harbor regions of bent DNA.

Most RNA Pol I promoters contain two distinct sequence elements: a core promoter (spanning approximately positions −50 to +10) and an upstream element (till about position −150), both of which are required in vivo (11). However, this bipartite arrangement is not seen in *Acanthamoeba castellani* (1) and *A. thaliana* (5). In these organisms the upstream element is absent. The promoter defined by us in *E. histolytica* is equivalent to the core promoter in other systems. No upstream sequences were found necessary for promoter activity. rRNA promoters in plants have a very highly conserved sequence between nucleotide positions −6 to +6 immediately surrounding the tsp (5). This sequence, which is A+T-rich from positions −6 to −1 and G (or A+G)-rich from positions +2 to +6 is also quite conserved in other organisms. It is proposed that this sequence may be equivalent to the RNA Pol II INR (initiator), which can assemble the Pol II transcription complex and determine transcription start site in the absence of other promoter domains (5). In *E. histolytica* this sequence in rDNA I is TACTATACAGGA and in rDNA II is ATAATGTAGAGG (with the underlined positions being +1), showing a similar sequence arrangement as reported in plants. An unusual feature observed by us in the *E. histolytica* promoter is a requirement for a longer stretch of downstream sequences (>80 nucleotides) compared to other systems where sequences till + 20 are generally sufficient. In mouse rDNA the sequence from positions +10 to +18 is involved in stem-loop formation with a further downstream sequence (positions +31 to +39). Removal of sequences between positions +23 and +53 reduces transcription efficiency due to disruption of the stem-loop (22). We checked for propensity to form stem-loop structures in the downstream region of *E. histolytica* rRNA genes, using mfold (http://mfold.bioinfo.rpi.edu). The region from +1 to +80 of rDNA I folded into a 25-bp long stem (from positions +13 to +76), while in rDNA II a 16-bp stem formed between nucleotides +27 and +69, in addition to shorter stem-loops. It remains to be seen whether these long stems are inhibitory to the progression of RNA Pol I and whether further downstream sequences till position +123 relieve this inhibition. Although the requirement for downstream sequences in rRNA promoter activity is not clearly understood in any system, it has been reported that the U3 proteins that are part of the SSU processome associate with this region in *Saccharomyces cerevisiae* and stimulate transcription, possibly by influencing chromatin structure (6).

Our data show that in exponentially growing cells of *E. histolytica* both rDNA I and II are transcribed with comparable efficiency (as measured by nuclear runon). Preliminary work from our lab has shown that the two units are differentially expressed during growth stress (unpublished data). Further work is required to understand transcription regulation during stress and to determine whether the two units are differentially expressed during conversion of actively dividing trophozoite to

the resting cyst stage. Our work has laid the basis for these future investigations.

ACKNOWLEDGMENTS

This research was supported by a FIRCA grant (TW 5655-03) from the NIH. S.K.P. and G.D.J. acknowledge a senior fellowship from the CSIR (India).

REFERENCES

1. **Bateman, E., C. T. Iida, P. Kownin, and M. R. Paule.** 1985. Footprinting of rRNA genes by transcription initiation factor and RNA polymerase I. Proc. Natl. Acad. Sci. USA **82:**8004–8008.
2. **Bhattacharya, S., I. Som, and A. Bhattacharya.** 1998. The ribosomal DNA plasmids of entamoeba. Parasitol. Today **14:**181–185.
3. **Bruchhaus, I., M. Leippe, C. Lioutas, and E. Tannich.** 1993. Unusual gene organization in the protozoan parasite *Entamoeba histolytica*. DNA Cell Biol. **12:**925–933.
4. **Diamond, L. S., D. R. Harlow, and C. C. Cunnick.** 1978. A new medium for the axenic cultivation of *Entamoeba histolytica* and other entamoebas. Trans. R. Soc. Trop. Med. Hyg. **72:**431–432.
5. **Doelling, J. H., and C. S. Pikaard.** 1995. The minimal rRNA gene promoter of *Arabidopsis thaliana* includes a critical element at the transcription initiation site. Plant J. **8:**683–692.
6. **Gallagher, J. E., D. A. Dunbar, S. Granneman, B. M. Mitchell, Y. Osheim, A. L. Beyer, and S. J. Baserga.** 2004. RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. Genes Dev. **18:**2506–2517.
7. **Graves, T., M. Dante, L. Eisenhour, and T. W. Christianson.** 1998. Precise mapping and characterization of the RNA primers of DNA replication for a yeast hypersuppressive petite by in vitro capping with guanylyltransferase. Nucleic Acids Res. **26:**1309–1316.
8. **Kneidl, C., E. Dinkl, and F. Grummt.** 1995. An intrinsically bent region upstream of the transcription start site of the rRNA genes of *Arabidopsis thaliana* interacts with an HMG-related protein. Plant Mol. Biol. **27:**705–713.
9. **Mayer, C., H. Bierhoff, and I. Grummt.** 2005. The nucleolus as a stress sensor: JNK2 inactivates the transcription factor TIF-1A and down-regulates rRNA synthesis. Genes Dev. **19:**933–941.
10. **Michel, B., P. M. Lizardi, A. Alagon, and M. Zurita.** 1995. Identification and analysis of the start site of rRNA transcription of *Entamoeba histolytica*. Mol. Biochem. Parasitol. **73:**19–30.
11. **Moss, T., F. Langlois, T. Gagnon-Kugler, and V. Stefanovsky.** 2007. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. Cell Mol. Life Sci. **64:**29–49.
12. **Nickel, R., and E. Tannich.** 1994. Transfection and transient expression of chloramphenicol acetyltransferase gene in the protozoan parasite *Entamoeba histolytica*. Proc. Natl. Acad. Sci. USA **91:**7095–7098.
13. **Perez, D. G., C. Gomez, E. Lopez-Bayghen, E. Tannich, and E. Orozco.** 1998. Transcriptional analysis of the EhPgp5 promoter of *Entamoeba histolytica* multidrug-resistant mutant. J. Biol. Chem. **273:**7285–7292.
14. **Rudenko, G., and L. H. Van der Ploeg.** 1989. Transcription of telomere repeats in protozoa. EMBO J. **8:**2633–2638.
15. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
16. **Schroth, G. P., J. S. Siino, C. A. Cooney, J. P. Th'ng, P. S. Ho, and E. M. Bradbury.** 1992. Intrinsically bent DNA flanks both sides of an RNA polymerase I transcription start site: both regions display novel electrophoretic mobility. J. Biol. Chem. **267:**9958–9964.
17. **Sehgal, D., V. Mittal, S. Ramachandran, S. K. Dhar, A. Bhattacharya, and S. Bhattacharya.** 1994. Nucleotide sequence organisation and analysis of the nuclear ribosomal DNA circle of the protozoan parasite *Entamoeba histolytica*. Mol. Biochem. Parasitol. **67:**205–214.
18. **Sollner-Webb, B., and E. B. Mougey.** 1991. News from the nucleolus: rRNA gene expression. Trends Biochem. Sci. **16:**58–62.
19. **Sollner-Webb, B., and J. Tower.** 1986. Transcription of cloned eukaryotic rRNA genes. Annu. Rev. Biochem. **55:**801–830.
20. **Stefanovsky, V., F. Langlois, T. Gagnon-Kugler, L. I. Rothblum, and T. Moss.** 2006. Growth factor signaling regulates elongation of RNA polymerase I transcription in mammals via UBF phosphorylation and r-chromatin remodeling. Mol. Cell **21:**629–639.
21. **Sylvester, J. E., R. Petersen, and R. D. Schmickel.** 1989. Human ribosomal DNA: novel sequence organization in a 4.5-kb region upstream from the promoter. Gene **84:**193–196.
22. **Yamamoto, O., N. Takakusa, Y. Mishima, R. Kominami, and M. Muramatsu.** 1984. Determination of the promoter region of mouse rRNA gene by an in vitro transcription system. Proc. Natl. Acad. Sci. USA **81:**299–303.
23. **Zacharias, M., G. Theissen, C. Bradaczek, and R. Wagner.** 1991. Analysis of sequence elements important for the synthesis and control of rRNA in *Escherichia coli*. Biochimie **73:**699–712.