Volume 8 Number 21 1980 Nucleic Acids Research

The structure of the yeast ribosomal RNA genes. 2. The nucleotide sequence of the initiation site for ribosomal RNA transcription

A.A.Bayev, O.I.Georgiev*, A.A.Hadjiolov*, M.B.Kermekchiev*, N.Nikolaev*, K.G.Skryabin and V.M.Zakharyev

Institute of Molecular Biology, Academy of Sciences of the USSR, Vavilov str. 32, Moscow B-312, USSR, and *Institute of Molecular Biology, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

Received 12 August 1980

ABSTRACT

The 5'-terminal coding sequence for the 37 S precursor to rRNA of *Saccharomyces cerevisiae* is identified by reverse transcriptase extension and protection mapping with nuclease S1. The sequence of a 419 bp rDNA fragment containing the transcription initiation site and its adjacent region is determined.

INTRODUCTION

Determination of the primary structure of transcription units and their flanking regions is important for the elucidation of regulatory mechanisms involved in gene expression. Although the nucleotide sequence of promotor regions in prokaryotes is studied to considerable details (1,2), little is known about putative regulatory sequences for transcription initiation in eukaryotes. Recently, the nucleotide sequences adjacent to 5 S rRNA (3-5) and some pre-mRNA (6) genes were determined. Analysis of initiation sites for rRNA genes in eukaryotes has been hampered by the instability of primary pre-rRNA and uncertainty about its identity with the primary gene transcript (7,8). Identification of 5'-triphosphates in eukaryotic primary pre-rRNA provided evidence that it is likely to be identical with the primary transcript (9-14).

In this work we describe the localization of initiation sites for the transcription of 37 S pre-rRNA in *Saccharomyces cerevisiae* and the nucleotide sequence of the adjacent rDNA region. Advantage was taken of the existence of extensive sequence information on the structure of rDNA in the region between 5 S rRNA and 18 S rRNA genes (15,16). The use of an osmotic-dependent yeast mutant (17) yielding low ribonuclease lysates (18) permitted the isolation of bulk amounts of intact 37 S pre-rRNA.

MATERIALS AND METHODS

The recombinant plasmid pY1rB3, containing yeast rDNA was used (19).

Restriction enzymes and polynucleotide kinase were prepared in our laboratory by standard procedures. Reverse transcriptase was a gift by Dr.J.Beard. The [α-^{32}P]-dATP (300 Ci/mmol) and ^{32}P-orthophosphate were obtained from the Radiochemical Centre, Amersham, U.K. The [γ-^{32}P]-ATP (about 1000 Ci/mmol) was prepared by the method of Glynn and Chappell (20). Plasmid DNA was isolated according to Tanaka and Weissblum (21). The fragments EcoRI-G and EcoRI-B (19) were purified by sucrose density gradient centrifugation or by electrophoresis in preparative 8 % acrylamide gels (22). Sequencing of DNA was done according to Maxam and Gilbert (23) using a revised protocol with several modifications (24).

Isolation and purification of 37 S pre-rRNA

The 37 S pre-rRNA was isolated from the osmotic-dependent *S.cerevisiae* mutant VY 1160 as described earlier (10). Yeast cells were grown at 30° C in YM 5 medium supplemented with 10 % sorbitol. At a culture density of 5-6x10^6 cells/ml, cycloheximide (100 µg/ml) and [^{3}H]-uracil were added. After 20 min. the cells were harvested on frozen 10 % sorbitol, washed, lysed and total RNA isolated (10). The 37 S pre-rRNA was purified from total RNA by several successive sedimentation runs in 10 to 30 % linear sucrose density gradients, containing 0.1 M Tris-acetate (pH 7.4), 5 mM EDTA and 0.5 % Na dodecyl sulfate, using the zonal (TZ 28) or vertical (TV 850) rotors of a Sorvall OTD-50 refrigerated ultracentrifuge. Before each sedimentation run the RNA sample was dissolved in 80 % dimethylsulfoxide, 10 % dimethylformamide, 10 % distilled water, denatured for 1 min. at 60° C, diluted 5-fold with the gradient buffer and applied directly onto the gradients. The position of 37 S pre-rRNA was followed by its [^{3}H]-label and A_{260}. Isolated samples were analyzed by agar/urea gel electrophoresis (25). Batches of 400-500 mg of total RNA yielded about 0.5 mg of purified 37 S pre-rRNA.

Hybridization of 37 S pre-rRNA with rDNA fragments and nuclease S1 mapping

Nuclease S1 mapping was carried out essentially according to Berk and Sharp (26). Five pmoles of the DNA fragment, dissolved in 9 µl 90 % formamide were denatured for 3 min. at 90° C and rapidly chilled on ice. Denatured DNA was added to a 37 S pre-rRNA pellet (about 0.5 pmoles) and after addition of 1 µl 20 x SSC the mixture was heated for 10 min. at 60° C and incubated for 120 min. at 50° C. After incubation, 2 µl tRNA (1 mg/ml) were added as a carrier and the hybrids were precipitated with 3 vol. of ethanol, containing 0.3 M Na acetate (pH 6.0). The obtained RNA-DNA hybrids were dissolved in 0.2 M NaCl, 50 mM Na acetate (pH 4.5), 1 mM $ZnCl_2$ and treated with 5 to 30 units of nuclease S1 for 15 min. at 37° C. The hybrids were extracted with

phenol and ethanol-precipitated. The products resistant to nuclease S1 were run onto a sequencing gel in parallel with a DNA nucleotide sequence ladder produced by the method of Maxam and Gilbert (23).

Sequence analysis by reverse transcription

This was carried out essentially as described by Levy et al.(27). A hybrid of 37 S pre-rRNA with the appropriate primer DNA fragment was obtained as described above. The hybrid was dissolved in 8 µl buffer H (250 mM NaCl, 170 mM Tris-HCl, pH 8.3, 30 mM $MgCl_2$, 25 mM dithiotreitol) and supplemented with 2 µl of actinomycin D (final conc.0.12 µg/µl). The solution was divided into 4 samples and to each sample 1 µl of a dNTP mixture (0.06 mM [α-^{32}P]dATP plus 0.5 mM each of dCTP, dTTP and dGTP) was added. To separate samples a ddNTP terminator was added as follows: for "A" termination 1 µl 0.2 mM ddATP, for "T", "G" and "C" termination 1 µl 0.5 mM ddTTP, ddGTP and ddCTP, respectively. To each sample 2-3 units of reverse transcriptase were added and the mixture incubated for 15 min. at 42° C. The reaction was stopped by addition of 1 µl N NaOH and alkaline digestion carried out for 3 min. at 90°C. After neutralization, the reaction products were ethanol-precipitated and analyzed further on 8 % acrylamide sequencing gels.

RESULTS AND DISCUSSION

Nucleotide Sequence of the 37 S pre-rRNA 5'-Terminus

The inclusion of several denaturation steps in the purification procedure removed substantial amounts of A_{260}-absorbing material co-sedimenting with 37 S pre-rRNA. The last sucrose density gradient step yielded a homogeneous 37 S pre-rRNA peak and this material gave a single band upon denaturation treatment and further agar/urea gel electrophoresis (Figure 1). These results show that the purified 37 S pre-rRNA is homogeneous, intact, without aggregated shorter RNA chains and can be used for the analysis of its 5'-terminus.

Due to the high molecular weight of 37 S pre-rRNA direct sequencing of its 5'-terminus is difficult. Therefore, we analyzed the reverse transcriptase extension product using 37 S pre-rRNA as template and a known rDNA fragment as primer. Analysis of the 5'-terminus of 37 S pre-rRNA requires the selection of a rDNA fragment mapped in this region. It is known that the 5'-end of the 18 S rRNA gene is located in the Eco RI-C fragment at 376 nucleotides downstream from the Eco RI site (24). Therefore, the restriction subfragments of the Eco RI fragments B, G and C had to be considered (Figure 2). The Eco RI-G fragment of 277 nucleotide pairs was ^{32}P-labelled at its 5'-termini and used as a reverse transcriptase primer with 37 S pre-rRNA as template. Estimation of the size of the product by acrylamide gel electrophoresis showed

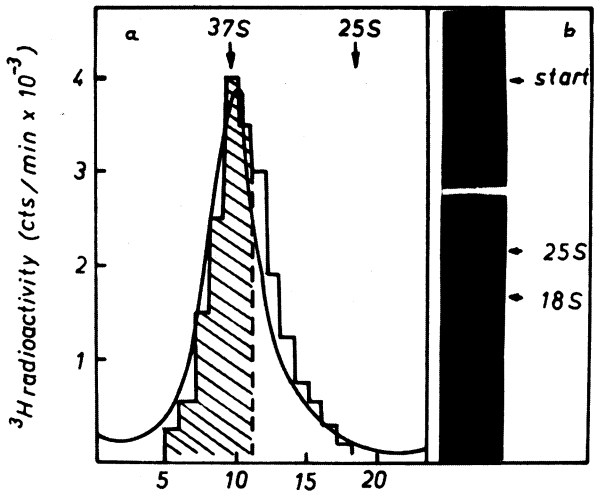

Figure 1. Last step in the purification of 37 S pre-rRNA. The 37 S pre-rRNA is purified from total RNA by several successive sucrose density gradient runs. (a) Last centrifugation in a linear 10-30 % sucrose density gradient. The continuous line represents $A_{254\,nm}$; the bars give the radioactivity of 50 µl samples from each 1 ml fraction. The material from the hatched zone of the gradient is collected and used for further analyses. (b) Autoradiogram of 37 S pre-rRNA after denaturation treatment and agar/urea gel electrophoresis. Labelling with [^{3}H]-uracil in (a) and ^{32}P-phosphate in (b).

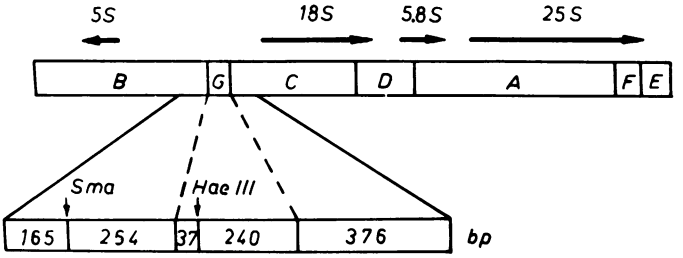

Figure 2. Endonuclease Eco RI restriction map of the *S.cerevisiae* rDNA repeating unit. Above: location of Eco RI fragments A to G. The arrows indicate the position of rRNA genes. Below: expanded map showing the location of Eco RI - Hae III and Sma - Eco RI subfargemnts of fragments G and B. The size of the sequenced part of the Hae III - Eco RI sub-fragment of fragment B and of the segment preceding the 18 S rRNA gene is indicated.

that the DNA primer is elongated by about 45-50 nucleotides (data not shown). Further, following the same approach, we used the 37 nucleotide pairs Eco RI - Hae III subfragment of the Eco RI-G fragment as primer (see Figure 2). After elongation of this fragment by reverse transcriptase in the presence of ddNTP terminators (Figure 3) a 5'-terminus sequence for 37 S pre-rRNA may be

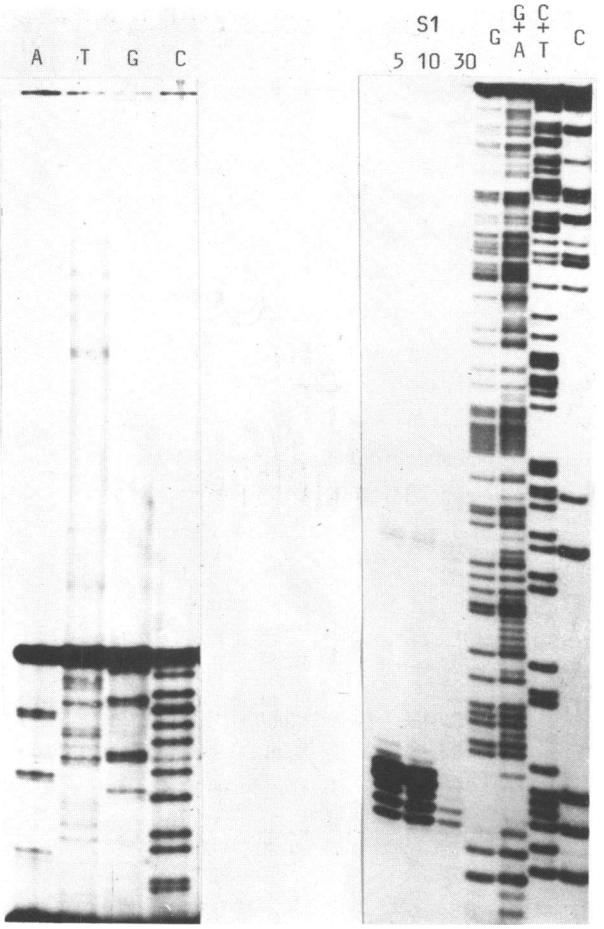

Figure 3 (left). Reverse transcriptase extension of the Eco RI-Hae III subfragment of fragment Eco RI-G hybridized to 37 S pre-rRNA. The hybrids are extended with reverse transcriptase in the presense of a dNTP mixture containing [α-^{32}P]dATP. Sequencing of the extended rDNA subfragment is achieved by addition of ddNTP terminators: A - ddATP; T - ddTTP; G - ddGTP; C - ddCTP.
Figure 4 (right). Nuclease S1 mapping of 37 S pre-rRNA. The Hae III-Eco RI subfragment of fragment Eco RI-B is labelled at its 5'-end with [γ-^{32}P]ATP and hybridized to 37 S pre-rRNA. The hybrids are digested with 5, 10 and 30 units of nuclease S1 and run in parallel with a Maxam and Gilbert (23) sequence ladder of the same rDNA subfragment.

deduced as 5'-$(N)_{2-3}$GCGAAAGCAGUUGAAGACAAGUUCGAAAAGAGUUU... This sequence coincides exactly with the terminal sequence of the Sma - Eco RI subfragment of the Eco RI-B fragment (15, see below). However, the technique used does not allow the unequivocal identification of the initial 5'-terminus nucleotides in 37 S pre-rRNA.

Mapping of the 5'-Terminus of 37 S pre-rRNA and Structure of Adjacent rDNA Regions

For precise mapping of the 5'-terminus of 37 S pre-rRNA we exploited the single-strand specificity of nuclease S1 (26). The Hae III - Eco RI subfragment of the Eco RI-B fragment, ^{32}P-labelled at the Eco RI 5'-end, was isolated. It was hybridized with 37 S pre-rRNA and digested with nuclease S1. After denaturation, the resultant ^{32}P-DNA was sized by acrylamide gel electrophoresis run in parallel with a DNA sequence ladder obtained from the same Hae III - Eco RI subfragment (Figure 4). Identical results were obtained by analysis of the hybrids of 37 S pre-rRNA with the Sma - Eco RI subfragment of the Eco RI-B fragment (data not shown). Two groups of nuclease S1 resistant rDNA fragments are observed. The two major bands of the first group of five bands have a mobility corresponding to nucleotides 47 and 48 (corrected for a single nucleotide, see ref.28) from the 5'-end of the coding strand of the Hae III - Eco RI or the Sma - Eco RI subfragments of fragment Eco RI-B. The two bands of the second group display a mobility which corresponds to nucleotides 78 and 79 of the same subfragments of the Eco RI-B fragment.

The sequence of part of the Hae III - Eco RI subfragment of the Eco RI-B fragment was determined as shown in Figure 5.

The results obtained by the two independent techniques used in this work locate transcription initiation for 37 S pre-rRNA at nucleotides 47-51. Previous studies have identified pppAp at the 5'-end of 37 S pre-rRNA (10) and

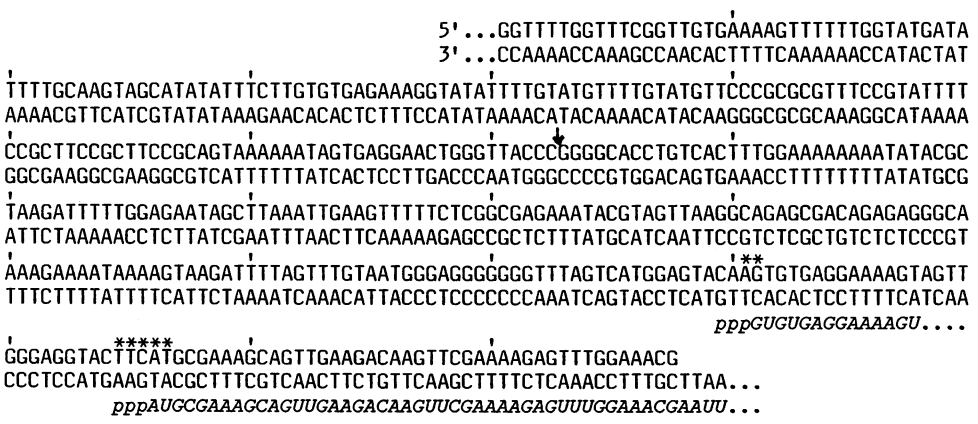

Figure 5. Primary structure of part of the Hae III - Eco RI subfragment of the Eco RI-B fragment and location of 37 S pre-rRNA transcription initiation sites. The numbering (right to left) adopted in this work starts from the Eco RI site in the coding strand. Nucleotides marked by * are identified by nuclease S1 mapping. The arrow indicates the Sma site.

therefore it is most likely that transcription starts at nucleotide 48 (see Figure 5). Thus the 5'-end of 37 S pre-rRNA maps at about 700 bp upstream from the 5'-end of 18 S rRNA. Further, experiments on nuclease S1 resistant fragments indicate the possibility for an additional transcription start site at nucleotides 78 and 79, corresponding to the presence of 5'-terminal A and G in 37 S pre-rRNA (10). The absence of reverse transcriptase extension to this site makes uncertain its role as a major transcription start. Yet, it is noteworthy that at both transcription starts similar flanking sequences (GGAGTAC...TGTGA and GGAGGTAC....TGCGA) can be discerned (non-coding strand).

Comparison of the only known nucleotide sequences preceeding transcription initiation sites of rRNA genes in *Xenopus laevis* (28-30) and *Saccharomyces cerevisiae* (this work) do not reveal any apparent similarities. It should be be noted also that from nucleotide 96 upstream considerable differences exist between the nucleotide sequences found in this work and the corresponding rDNA region published earlier by Valenzuela et al.(14). These differences in the spacer region could be due to the use of different clones of the rDNA repeating unit.

Correspondence should be addressed to: Prof. Dr. A.A.Hadjiolov, Institute of Molecular Biology, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria.

REFERENCES

1. Gilbert,W.(1976) in RNA Polymerase, Losick,R. and Chamberlin,M. Eds., pp.193-205, Cold Spring Harbor Laboratory, N.Y.
2. Rosenberg,M.and Court,D.(1979) Ann.Rev.Genet., 13, 319-353.
3. Maxam,A., Tizard,R., Gilbert,W.and Skryabin,K.G.(1977) Nature, 267, 643-644.
4. Valenzuela,P., Bell,G.I., Masiarz,F.R., DeGennaro,L.J.and Rutter,W.J.(1977) Nature, 267, 641-643.
5. Fedoroff,N.and Brown,D.D.(1978) Cell, 13, 701-716.
6. Benoist,C., O'Hare,K., Breathnach,R.and Chambon,P.(1980) Nucl.Acids Res., 8, 127-142.
7. Hadjiolov,A.A.and Nikolaev,N.(1976) Progr.Biophys.Mol.Biol., 31, 95-144
8. Rungger,D.and Crippa,M.(1977) Progr.Biophys.Mol.Biol., 31, 247-269.
9. Hadjiolov,A.A., Dabeva,M.D., Dudov,K.P., Gajdardjieva,K.C., Georgiev,O.I., Nikolaev,N.and Stoyanova,B.B.(1978) FEBS Symposia, 51, 319-328.
10. Nikolaev,N., Georgiev,O.I., Venkov,P.V.and Hadjiolov,A.A.(1979) J.Mol.Biol., 127, 297-308.
11. Levis,R.and Penman,S.(1978) J.Mol.Biol., 121, 219-238.
12. Reeder,R.,Sollner-Webb,B.and Wahn,H.(1977) Proc.Nat.Acad.Sci.USA,74,5402-5406
13. Niles,E.(1978) Biochemistry, 17, 4839-4844.
14. Batts-Young,B.and Lodish,H.(1978) Proc.Nat.Acad.Sci.USA, 75, 740-744.
15. Valenzuela,P., Bell,G.I., Venegas,A., Sewell,E.T., Masiarz,F.R., DeGennaro,L. Weinberg,F. and Rutter,W.J.(1977) J.Biol.Chem., 252, 8126-8135.
16. Skryabin,K.G., Zakharyev,V.M., Rubtsov,P.M.and Bayev,A.A.(1979) Doklady Akad.Nauk USSR, 247, 1275-1277.
17. Venkov,P.V., Hadjiolov,A.A., Battaner,E. and Sclessinger,D.(1974) Biochem. Biophys.Res.Commun., 56, 599-604.

18. Venkov,P.V.(1979) Molec.Gen.Genet., 175, 111-112.
19. Petes,D.T., Hereford,L.M. and Skryabin,K.G.(1978) J.Bacteriol., 134, 295-305.
20. Glynn,I.M. and Chappell,J.B.(1964) Biochem.J., 90, 147-154.
21. Tanaka,T. and Weissblum,B.(1975) J.Bacteriol., 121, 354-362.
22. Skryabin,K.G., Zakharyev,V.M. and Bayev,A.A.(1978) Doklady Akad.Nauk USSR, 241, 488-490.
23. Maxam,A.M. and Gilbert,W.(1977) Proc.Nat.Acad.Sci.USA, 74, 560-564.
24. Rubtsov,P., Musahanov,M.M., Zakharyev,V.M., Krayev,A.S., Skryabin,K.G. and Bayev,A.A., Nucl.Acids Res. submitted.
25. Dudov,K.P., Dabeva,M.D. and Hadjiolov,A.A.(1976) Anal.Biochem.,76, 250-258.
26. Berk,A.J. and Sharp,P.A.(1977) Cell, 12, 721-732.
27. Levy,S., Sures,I. and Kedes,L.H.(1979) Nature, 279, 737-739.
28. Moss,T. and Birnstiel,M.(1979) Nucl.Acids Res., 6, 3733-3743.
29. Boseley,P., Moss,T., Machler,M., Portman,R. and Birnstiel,M.(1979) Cell, 17, 19-31.
30. Sollner-Webb,B. and Reeder,R.H.(1979) Cell, 18, 485-499.