Volume 9 Number 4 1981 Nucleic Acids Research

The structure of the yeast ribosomal RNA genes. 3. Precise mapping of the 18 S and 25 S rRNA genes and structure of the adjacent regions

A.A.Bayev, O.I.Georgiev*, A.A.Hadjiolov*, 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 29 December 1980

ABSTRACT

The 5'-termini of *Saccharomyces cerevisiae* 18 S and 25 S rRNA are precisely mapped within the sequence of the rDNA repeating unit. The 3'-termini of 25 S rRNA and 37 S pre-rRNA are located within a 548 bp segment of the rDNA repeating unit by the use of a DNA polymerase I extension technique. The analysis of the rDNA sequences at the structural gene boundaries reveals the presence of oligonucleotide repeats which may be involved in transcription or processing control mechanisms. The sequence of rDNA in the transcription termination region is determined and possible mechanisms shaping the 3'-end of 25 S rRNA are discussed.

INTRODUCTION

Generation of mature rRNA molecules in eukaryotes involves several subsequent processing steps (1-3). The mechanism of action of processing enzymes is not yet clear partly due to the absence of information on the sequence of the terminal segments of rRNA and their precursors. In our studies on the structure of the yeast rRNA genes (4-6) we encountered difficulties in the exact location of 25 S and 18 S rRNA termini due to the presence of repeated oligonucleotide sequences in the external boundaries regions of the structural genes.

In the present work the localization of the 5'-terminus of the 18 S and both the 5'- and 3'-termini of the 25 S rRNA genes is identified within the sequence of the respective rDNA regions. The 3'-terminus of 37 S pre-rRNA, the primary transcript of rRNA genes (7,8),is also mapped showing that transcription termination occurs at or a few nucleotides downstream from the 3'-end of the 25 S rRNA gene. The use of a DNA polymerase I extension technique for mapping the 3'-termini of 25 S rRNA and 37 S pre-rRNA is described.

MATERIALS AND METHODS

The recombinant plasmids pY1rA3, pY1rA9 and pY1rB3, containing *S.cerevisiae* rDNA were used (9). Restriction enzymes, polynucleotide kinase and

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

DNA polymerase I from *E.coli* were prepared in our laboratory by standard procedures. Ribonucleases T_1 and U_2 were obtained from Calbiochem AG, Lucern, Switzerland. The [α-^{32}P]dCTP, [α-^{32}P]dGTP, [α-^{32}P]dTTP (400 Ci/mmol) as well as ^{32}P-orthophosphate were products of The Radiochemical Centre, Amersham, UK. The [γ-^{32}P]ATP (about 1000 Ci/mmol) was prepared according to Glynn and Chappel (10).

Plasmid DNA was isolated by the method of Tanaka and Weissblum (11). The rDNA fragments Eco RI-C, Eco RI-B, Eco RI-E and Eco RI-A (9) were purified by sucrose density gradient centrifugation and/or by preparative electrophoresis in 4 % acrylamide gels (12). Sequencing of DNA was carried out according to Maxam and Gilbert (13) with minor modifications (5). Sequencing of RNA was done according to Donis-Keller et al.(14). The sequence of the 5'-terminal oligonucleotide in 18 S rRNA was determined by the method of Sanger et al.(15) and Jay et al.(16). The location of the 5'-end of the 25 S rRNA gene was identified by nuclease S_1 mapping (17) and reverse transcriptase extension (18) as described in details earlier (6).

The 18 S and 25 S rRNA, and 37 S pre-rRNA were isolated and purified from the osmotic-dependent *S.cerevisiae* mutant VY 1160 (19) as described earlier (7).

Mapping of the 3'-end of rRNA by DNA polymerase I extension

The method is based on the observation that DNA polymerase I can synthesize complementary DNA on RNA-DNA hybrids using both strands as primer and template (20-22). Approximately 1 µmole of the rDNA fragment was dissolved in 9 µl 90 % formamide, denatured for 3 min. at 90°C and rapidly chilled on ice. The denatured rDNA was added to a 25 S rRNA or 37 S pre-rRNA pellet (about 10 pmoles) and after addition of 1 µl 20xSSC, the mixture was heated 10 min. at 60°C and annealed for 120 min. at 50°C. The RNA-DNA hybrids were precipitated with 3 vol. of 96 % ethanol, containing 0.3 M Na acetate (pH 6.0). The hybrids were dissolved in 15 µl of 2.2 mM each Tris-HCl (pH 7.5), NaCl, $MgCl_2$ and dithiotreitol. To this solution 1 µl of a mixture of dTTP, dCTP, dGTP and dATP (1 nmol each) was added, followed by 20 µl of a mixture containing [α-^{32}P]dTTP, [α-^{32}P]dCTP and [α-^{32}P]dGTP (40 pmoles each). The volume was brought to 40 µl with distilled water and 10 µl (2 units) of DNA polymerase I were added. The enzyme reaction was carried out for 60 min. at 15°C. The reaction was stopped by the addition of 8.5 µl N NaOH and the samples were heated for 3 min. at 90°C to hydrolyze the RNA chains. After neutralization with HCl, the DNA was precipitated with 3 vol. of 96 % ethanol, containing 0.3 M Na acetate. The labelled DNA products were analysed by electro-

phoresis in 8 % acrylamide - 8 M urea gels.

RESULTS AND DISCUSSION

The boundaries of 5 S rRNA (23,24), 5.8 S rRNA (25,26) and the 3'-ends of 18 S and "20 S" rRNA (27) are the only termini mapped precisely within the *Saccharomyces cerevisiae* rDNA repeating unit. The location of 25 S and 18 S rRNA genes (Figure 1) was established by the blotting technique (25,28,29) and R-loop analysis (30). These results locate the 5'-end of the 18 S rRNA gene within the Eco RI-C fragment. Bell et al.(25) proposed a more precise location between the Xba I and Bgl II restriction sites. The 5'-end of 25 S rRNA maps within fragment Eco RI-A (30), while its 3'-end is located within fragment Eco RI-E (30) with the possibility to overlap with the left end of fragment Eco RI-B. Both rRNA genes are transcribed in a common 37 S precursor to rRNA starting at the right end of the Eco RI-B fragment (6,31).

Nucleotide sequence and mapping of the 5'-terminus of 18 S rRNA

In order to identify the position of the 5'-end of 18 S rRNA within the Eco-RI-C fragment the sequence of the 5'-terminal oligonucleotide generated by ribonuclease T_1 was determined. The resulting hexanucleotide was pUAUCUG. Surprisingly, this hexanucleotide sequence is repeated two fold in the respective region of the Eco RI-C fragment (4). Therefore, more extensive sequence information on 18 S rRNA was needed in order to locate its 5'-end. To this end we analyzed more nucleotides from the 5'-terminus of 18 S rRNA (Figure 2). From this experiment the sequence pUAUCUGGUUG was deduced. Further, the distribution of purine nucleotides in a longer 5'-end segment of 18 S rRNA was analyzed by the method of Donis-Keller et al.(14). The results (data

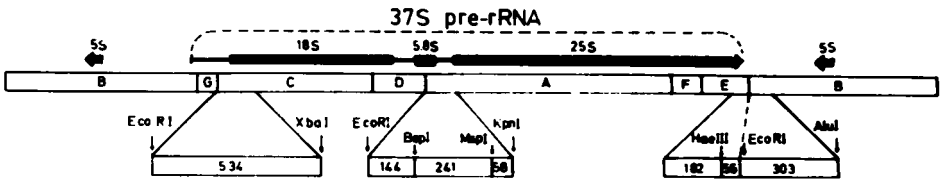

Figure 1. Endonuclease Eco RI restriction map of the *S.cerevisiae* rDNA repeating unit. Above: Arrangement of the Eco RI fragments A to G. The location of rRNA genes is indicated. Below: Expanded maps showing the position of the Eco RI-Xba I subfragment of fragment C, the Eco RI-Kpn I subfragment of fragment A and the analyzed regions of fragments E and B including the Hae III, Eco RI and Alu I restriction sites. The size of the sequenced parts of fragments C, A, E and B is indicated.

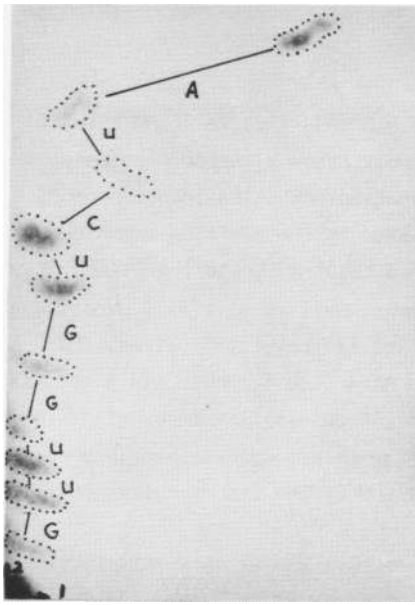

Figure 2. The 5'-terminal nucleotide sequence of *S.cerevisiae* 18 S rRNA.
Purified and dephosphorylated 18 S rRNA is labelled at its 5'-end with [γ-^{32}P]ATP. After partial alkaline digestion the 5'-end products are analyzed by two-dimensional mapping techniques (15, 16). The arrows indicate the direction of the electrophoresis [1] and chromatography [2] runs.

not shown) allow the unequivocal conclusion that the 5'-terminus of 18 S rRNA coincides with the decanucleotide box a2 (Figure 3). It should be noted that the 5'-end hexanucleotide of 18 S rRNA in *S.cerevisiae* is identical with the 5'-terminal pUAUCUG sequence established in *S.carlsbergesis* (32). The identification of this hexanucleotide at the 5'-end of both 18 S rRNA and its immediate precursor ("20 S" pre-rRNA) led these authors to the conclusion that the maturation process involves only the 3'-end of "20 S" pre-rRNA (32). However, our findings leave open the possibility that the

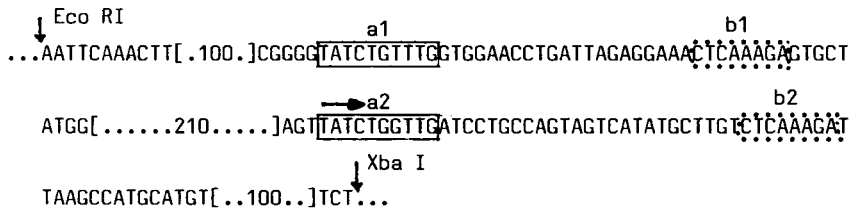

Figure 3. Primary structure (non-coding strand) of part of the Eco RI-Xba I subfragment of the Eco RI-C fragment showing the location of the 5'-end of 18 S rRNA (horizontal arrow). The decanucleotide and octanucleotide repeats located at similar distances from each other are marked by the boxes a1 and b1 and a2 and b2, respectively.

5'-end of "20 S" pre-rRNA is located at the decanucleotide box a1. This possibility is supported further by the presence of another repeated sequence of eight nucleotides (boxes b1 and b2) located at similar distances from the a1 and a2 repeat (see Figure 3). The significance of the observed repeated sequences remains unknown, but it is likely that they are involved in determining the specificity of processing mechanisms leading to mature 18 S rRNA.

The 5'-end of the 25 S rRNA gene

In order to locate the 5'-end of the 25 S rRNA gene we used the subfragment Eco RI-Kpn I of fragment Eco RI-A (30). Preliminary experiments showed that this fragment hybridizes with 25 S rRNA. More precise hybridization experiments revealed that the 5'-end of 25 S rRNA maps within the subfragment Bsp I-Msp I of fragment Eco RI-A (see Figure 1).

The exact location of the 5'-end of 25 S rRNA within the sequence of the subfragment Eco RI-Kpn I (Figure 4) was achieved by: (a) nuclease S_1 protection mapping of the hybrid between 25 S rRNA and the BspI-MspI subfargment of fragment Eco RI-A (Figure 5) and (b) reverse transcriptase extension of the hybrid between 25 S rRNA and the Msp I-Kpn I subfragment (52 nucleotides in the coding strand) of fragment Eco RI-A used as primer (see Figure 5). As can be seen, the nuclease S_1 protection experiment yields five possible 5'-end nucleotides of the 25 S rRNA gene with the predominence of the first three. This heterogeneity reflects more likely partial attack by nuclease S_1 of the terminal nucleotides in the RNA-DNA hybrid (33). The reverse transcriptase extension product gives only one strong band, followed by a weaker next band. After the appropriate corrections (33) both methods coincide to

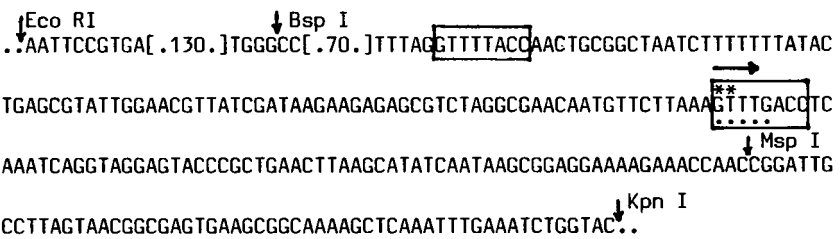

Figure 4. Primary structure (non-coding strand) of part of the Eco RI-Kpn I subfragment of the Eco RI-A fragment showing the location of the 5'-end of 25 S rRNA (horizontal arrow). The nucleotides marked by · and * are identified as corresponding to the 5'-terminus of 25 S rRNA by nuclease S_1 mapping (·) or by reverse transcriptase extension (*), respectively. The octanucleotide repeat located at the 5'-end of 25 S rRNA and 92 nucleotides upstream is boxed.

Figure 5. Identification of the 5'-terminus of the 25 S rRNA gene. Right - Mapping of nuclease S_1 protection products of the hybrids between 25 S rRNA and the Bsp I-Msp I subfragment of fragment Eco RI-Kpn I (see Figs 1 & 4). The hybrids are digested with 5 [a] and 25 [a'] units of nuclease S_1 and run in parallel with a Maxam and Gilbert (13) sequence ladder of the same rDNA subfragment. Left - Reverse transcriptase extension product of the Msp I-Kpn I subfragment of fragment Eco RI-Kpn I hybridized with 25 S rRNA. The horizontal arrows indicate the position of the 5'-terminus nucleotides of 25 S rRNA on the ladder of the Bsp I-Msp I subfragment after correction for the size of the Msp I-Kpn I subfragment (52 nucleotides in the coding strand) used as primer.

show that 25 S rRNA starts with the predominant sequence pGUUUGACCUC... and possibly pUUUGACCUC... (see Fig.4). The observed one nucleotide 5'-end heterogeneity of *S.cerevisiae* 25 S rRNA seems to reflect the situation in vivo since pU... was identified as the 5'-end of 25 S rRNA after hydrolysis with ribonuclease T_1 (34). The 5'-end of the 25 S rRNA gene is located at 235-236 nucleotides downstream from the 3'-end of the 5.8 S rRNA gene (Skryabin et al., in preparation). Inspection of the sequences at the 5'-end boundary of the 25 S rRNA gene fails to reveal any obvious structural features. However, it is noteworthy that the octanucleotide sequence at the 5'-end of 25 S rRNA is repeated 92 nucleotides upstream in the internal transcribed spacer.

The 3'-end of the 25 S rRNA gene and sequence of the transcription termination region

Electron microscopic analysis of hybrids of *S.cerevisiae* 25 S rRNA and rDNA (30) indicates that the 3'-end of 25 S rRNA maps in the Eco RI site region of the fragments Eco RI-E or Eco RI-B (see Figure 1). Therefore we determined the sequence of rDNA on both sides of the Eco RI site encompassing 238 nucleotides of Eco RI-E and 310 nucleotides of Eco RI-B (Figure 6).

To identify the position of the 3'-end of 25 S rRNA we used DNA polymerase I extension of the hybrids between 25 S rRNA and fragments Eco RI-E or

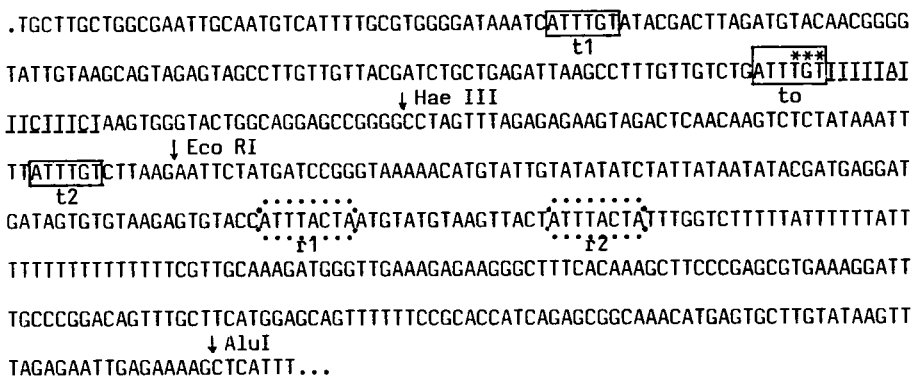

Figure 6. Primary structure (non-coding strand) of the parts of Eco RI-E and Eco RI-B fragments encompassing the 3'-terminus of the 25 S rRNA gene. The nucleotides marked by * are identified as corresponding to the 3'-terminus of 25 S rRNA (see text and Fig.7). The boxes to, t1 and t2 designate a hexanucleotide repeat found at the distal end of fragment Eco RI-E. The octanucleotide repeat r1 and r2 marks the position of a 24 nucleotide insertion not encountered in previous structural studies on a different yeast rDNA clone. Terminal nucleotides found in 37 S pre-rRNA are underlined.

Eco RI-B (see Methods). Negative results were obtained when hybrids of the rDNA fragment Eco RI-B with either 25 S rRNA or 37 S pre-rRNA were analyzed (data not shown) thus showing that the 3'-end of these two rRNA should map within fragment Eco RI-E. The analysis of the DNA polymerase I extension products using a hybrid between 25 S rRNA and fragment Eco RI-E is presented in Figure 7. The results shown in Figure 7 (A) demonstrate the efficiency of the DNA polymerase I extension method. The action of DNA polymerase I on the hybrid Eco RI-E rDNA x 25 S rRNA extends the primer rRNA with a chain of about 100 nucleotides. The accuracy of the DNA polymerase I extension was ascertained by the cleavage of the extension product with restriction endonuclease Hae III yielding the expected segment of 56 nucleotides (see Fig.6) upstream from the Eco RI site (data not shown). The precise location of the 3'-end of 25 S rRNA was achieved by apposition of the DNA polymerase I extension product with a Maxam and Gilbert sequence ladder of the rDNA fragment Eco RI-E (see Fig.5, B). The obtained results locate the 3'-end sequence of 25 S rRNA at the right end of the hexanucleotide box t_o (see Fig.6). The observed heterogeneity in the terminal nucleotides in the DNA polymerase I extension products may reflect an artifact of the method or a genuine heterogeneity of the 3'-end of 25 S rRNA. The major band yields the 3'-terminal sequence ..AUUUGU$_{OH}$ which agrees with previous results obtained with *S.cerevisiae* (34,35) and *S.carlsbergensis* (8,36) 25 S rRNA. Inspection of the determined sequences of fragments Eco RI-E and Eco RI-B reveals that the 3'-terminal hexanucleotide sequence (box t_o) is repeated at about the same distance (84-86 nucleotides) upstream and downstream (boxes t1 and t2). The role of these repeats in shaping the 3'-end of 25 S rRNA remains to be clarified.

The DNA polymerase I extension method was used also to identify the 3'-end of 37 S pre-rRNA (see Fig.7). As can be seen, the major extension product with 37 S pre-rRNA as template coincides with the 3'-end of the 25 S rRNA gene. However, in this case well delimited, longer extension products are present covering a T-rich region of about 15 nucleotides. In studies with *S.carlsbergensis* the sequence $U_{6-8}{}^{AN}{}_{OH}$ was identified tentatively as the 3'-terminus of 37 S pre-rRNA (8). The corresponding sequence ..TTTTTTAT..in the Eco RI-E fragment is found only immediately after the 3'-end of the 25 S rRNA gene. At present we do not know whether the 3'-end of 37 S pre-rRNA is produced by transcription termination or by ribonuclease processing, or both. Our results show that the major site of transcription termination coincides with the 3'-end of the 25 S rRNA gene as found in the case of *Xenopus laevis* (37). Read-through products beyond this point are also present, but the

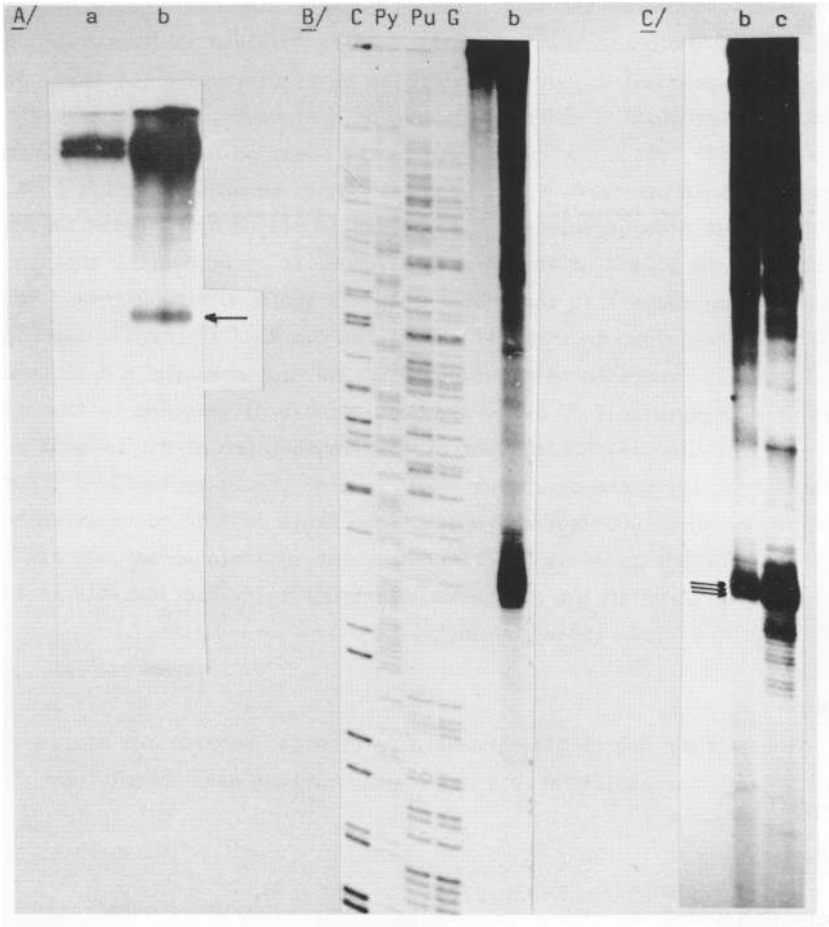

Figure 7. Identification of the 3'-terminus of 25 S rRNA and 37 S pre-rRNA by DNA polymerase I extension.
A/ Acrylamide gel electrophoresis (8 % gel) of the [α-^{32}P]dNTP-labelled reaction product of DNA polymerase I with the Eco RI-E fragment alone (a) and after hybridization with 25 S rRNA (b). The arrow indicates the position of the extension product (about 100 nucleotides) with the Eco RI-E rDNA x 25 S rRNA hybrid.
B/ Mapping of the DNA polymerase I extension product of the Eco RI-E rDNA x 25 S rRNA hybrid (b) run in parallel with a sequence ladder of the same rDNA fragment (non-coding strand).
C/ Comparison of DNA polymerase I extension products of Eco RI-E hybridized with 25 S rRNA (b) or 37 S pre-rRNA (c).

extra U-rich sequences are most likely trimmed by the action of a processing ribonuclease. The participation of the factors involved in transcription termination and processing in vivo remains to be clarified. In our case, 37 S pre-rRNA is accumulated under conditions of cycloheximide block of protein synthesis (7), which may contribute to the observed amounts of different 3'-ends for this pre-rRNA. Comparison with known termination sites (38,39) indicates that a major role in this process is played by a T-rich sequence distal from the 3'-end of the structural gene. It is noteworthy that in the case of *S.cerevisiae* this terminator sequence is found also within a very long T-rich sequence close to the left end of the Eco RI-B fragment (see Fig.6).

Finally, comparison with the reported DNA sequences for a different rDNA clone of *S.cerevisiae* (23) reveals considerable differences in the structure of the Eco RI-Hind III subfragment of the fragment Eco RI-B. The most profound difference is the presence in our rDNA clone of one 24-nucleotides insertion delimited by an octanucleotide repeat (see Fig.6). This observation indicates that further analysis of different yeast rDNA clones may clarify the mechanisms involved in the generation of structural variations within the spacer region of rDNA repeating units.

ACKNOWLEDGMENT

The authors are greatly indebted to Dr M.Kermekchiev for his expert assistance in the isolation of the RNA preparations used in this work.

REFERENCES

1. Perry,R.P.(1976) Ann.Rev.Biochem., 45, 605-630.
2. Hadjiolov,A.A. and Nikolaev,N.(1976) Progr.Biophys.Mol.Biol., 31, 95-144.
3. Hadjiolov,A.A.(1980) in Roodyn,D.B. (Ed.) Subcellular Biochemistry, Vol.7, p.1-80, Plenum press, New York.
4. Skryabin,K.G., Krayev,A.S., Rubtsov,P.M. and Bayev,A.A.(1979) Doklady Akad.Nauk USSR, 247, 1275-1277.
5. Rubtsov,P.M., Musakhanov,M.M., Zakharyev,V.M., Krayev,A.S., Skryabin,K.G. and Bayev,A.A.(1980) Nucl.Acids Res., 8, 5779-5794.
6. Bayev,A.A., Georgiev,O.I., Hadjiolov,A.A., Kermekchiev,M.B., Nikolaev,N., Skryabin,K.G. and Zakharyev,V.M.(1980) Nucl.Acids Res., 8, 4919-4926.
7. Nikolaev,N., Georgiev,O.I., Venkov,P.V. and Hadjiolov,A.A.(1979) J.Mol. Biol., 127, 297-308.
8. Klootwijk,J., DeJonge,P. and Planta,R.J. (1979) Nucl.Acids Res., 6, 27-39.
9. Petes,D.T., Hereford,L.M. and Skryabin,K.G.(1978) J.Bacteriol., 134, 295-305.
10. Glynn,I.M. and Chappel,J.B.(1964) Biochem.J., 90, 147-154.
11. Tanaka,T. and Weissblum,B.(1975) J.Bacteriol., 121, 354-362.
12. Skryabin,K.G., Zakharyev,V.M. and Bayev,A.A.(1978) Doklady Akad.Nauk USSR, 241, 488-490.
13. Maxam,A.M. and Gilbert,W.(1977) Proc.Nat.Acad.Sci.USA, 74, 560-564.

14. Donis-Keller,H., Maxam,A.M. and Gilbert,W.(1977) Nucl.Acids Res., 4, 2527-2538.
15. Sanger,F., Brownlee,G.G. and Barrell,B.G.(1965) J.Mol.Biol., 13, 373-398.
16. Jay,E., Bambara,R., Padmanabhan,R. and Wu,R.(1974) Nucl.Acids Res., 1, 331-353.
17. Berk,A.J. and Sharp,P.A.(1977) Cell, 12, 721-732.
18. Levy,S., Sures,I. and Kedes,L.H.(1979) Nature, 279, 737-739.
19. Venkov,P.V., Hadjiolov,A.A., Battaner,E. and Schlessinger,D. (1974) Biochem.Biophys.Res.Commun., 56, 599-604.
20. Wells,R.D., Flugel,R.M., Larson,J.E., Schendel,P.F. and Sweet,R.W.(1972) Biochemistry, 11, 621-629.
21. Peters,G.G. and Hayward,R.S.(1972) Eur.J.Biochem., 31, 360-366.
22. Kornberg,A.(1974) DNA Synthesis, pp.67-123, W.H.Freeman & Co, San Francisco, USA.
23. Valenzuela,P., Bell,G.I., Venegas,A., Sewell,E.T., Masiarz,F.R.,DeGennaro L.J., Weinberg,F. and Rutter,W.J.(1977) J.Biol.Chem., 252, 8126-8135.
24. Maxam,A.M., Tizard,R., Skryabin,K.G. and Gilbert,W.(1977) Nature, 267, 643-645.
25. Bell,G.I., DeGennaro,L.J., Gelfand,D.H., Bishop,R.J., Valenzuela,P. and Rutter,W.J.(1977) J.Biol.Chem., 252, 8118-8125.
26. Skryabin,K.G., Maxam,A.M., Petes,T.D. and Hereford,L.(1978) J.Bacteriol., 134, 306-309.
27. Skryabin,K.G., Krayev,A.S., Rubtsov,P.M. and Bayev,A.A.(1979) Doklady Akad.Nauk USSR, 247, 761-765.
28. Nath,K. and Bollon,A.P.(1975) Nature, 257, 155-157.
29. Nath,K. and Bollon,A.P.(1977) J.Biol.Chem., 252, 2562-2571.
30. Philippsen,P., Thomas,M., Kramer,A. and Davis,R.W.(1978) J.Mol.Biol.,123, 387-404.
31. Klemenz,R. and Geiduschek,E.P.(1980) Nucl.Acids Res., 8, 2679-2691.
32. De Jonge,P., Klootwijk,J. and Planta,R.J.(1977) Eur.J.Biochem., 72, 361-369.
33. Moss,T. and Birnstiel,M.(1979) Nucl.Acids Res., 6, 3733-3743.
34. Sugiura,M. and Takanami,M.(1967) Proc.Nat.Acad.Sci.USA, 58, 1595-1601.
35. Shine,J., Hunt,J.A. and Dalgarno,L.(1974) Biochem.J., 141, 617-623.
36. Veldman,G.M., Brand,R.C., Klootwijk,J. and Planta,R.J.(1980) Nucl.Acids Res., 8, 2907-2920.
37. Sollner-Webb,B. and Reeder,R.H.(1979) Cell, 18, 485-499.
38. Adhya,S. and Gottesman,M.(1978) Ann.Rev.Biochem., 47, 967-996.
39. Rosenberg,M. and Court,D.(1979) Ann.Rev.Genet., 13, 319-353.