Volume 9 Number 24 1981 Nucleic Acids Research

The structure of the yeast ribosomal RNA genes. 4. Complete sequence of the 25 S rRNA gene from Saccharomyces cerevisiae

O.I.Georgiev, N.Nikolaev and A.A.Hadjiolov
Institute of Molecular Biology, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria, and

K.G.Skryabin, V.M.Zakharyev and A.A.Bayev
Institute of Molecular Biology, Academy of Sciences of the USSR, Vavilov str. 32, Moscow B-312, USSR

Received 19 October 1981

ABSTRACT

The complete nucleotide sequence of the 25 S rRNA gene from one rDNA repeating unit of *Saccharomyces cerevisiae* has been determined. The corresponding 25 S rRNA molecule contains 3392 nucleotides and has an estimated relative molecular mass (M_r, Na-salt) of 1.17×10^6. Striking sequence homology is observed with known 5'- and 3'-end terminal segments of L-rRNA from other eukaryotes. Possible models of interaction with 5.8 S rRNA are discussed.

INTRODUCTION

Determination of the primary structure of rRNA molecules and the respective genes is important for understanding ribosome structure, function, biogenesis and evolution. It is established that in *Saccharomyces cerevisiae* the structural genes for 5 S, 18 S, 5.8 S and 25 S rRNA are organized in one rDNA repeating unit (about 9.1 kb), present in 100 to 120 copies per haploid genome (1). The structure of the yeast rDNA repeating unit is studied in details and all four rRNA genes are mapped within seven Eco RI restriction fragments designated A to G according to their length (2-5). A large part of the rDNA repeating unit has been sequenced, including the 5 S (6,7), 18 S (8) and 5.8 S (2,9) rRNA genes. Further, the 5'- and 3'-ends of the 25 S rRNA gene are now precisely mapped within the Eco RI fragments A and E, respectively (10, 24).

In the present work we report the complete sequence of the 25 S rRNA gene from one rDNA repeating unit of *Saccharomyces cerevisiae*.

MATERIALS AND METHODS

The recombinant plasmids pY1rA9 and pY1rB3, containing the rDNA Eco RI fragments A, E and F, are used (3). Restriction endonucleases are prepared in our laboratories by standard procedures or obtained as a gift from Dr. A.Yanulaitis. The [^{32}P]-orthophosphate is a product of The Radiochemical

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

Centre, Amersham, U.K. The [γ^{32}P]ATP (about 1000 Ci/mmol) is prepared by the method of Glynn and Chappel (11). Plasmid DNA is isolated according to Tanaka and Weissblum (12). Restriction rDNA fragments are purified by preparative electrophoresis in 4 or 6 % acrylamide gels. End labelling is carried out with T_4 polynucleotide kinase (Boehringer-Mannheim) as described by Maxam and Gilbert (13,14). The end-labelled fragments are subjected to strand separation (15) or to secondary restriction endonuclease cleavage. Sequencing of DNA is carried out according to Maxam and Gilbert (13,14).

RESULTS AND DISCUSSION

Sequencing strategy. The entire 25 S rRNA gene is contained in the Eco RI fragments A, F and E (Figure 1). These fragments have estimated lengths of about 2.85, 0.36 and 0.59 kb, respectively. Therefore, first of all, a strategy for the sequencing of fragment Eco RI-A had to be devised. Analysis of the cleavage patterns of fragment Eco RI-A with different restriction endonucleases revealed that digestion with Msp I and Sau 3A provides two sets of 11 and 8 subfragments, respectively, which appear to be convenient for sequencing (see Fig.1). Most of the sequences in fragments Eco RI-A, as well as Eco RI-F and Eco RI-E, were determined in both strands and independently confirmed by the use of overlapping restriction fragments. In some cases the orientation of sequenced rDNA chains was defined by hybridization with 25 S rRNA.

Sequence results. The complete sequence of 25 S rRNA (deduced from the

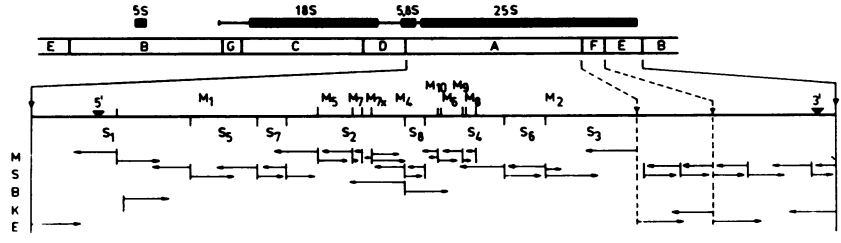

FIGURE 1. Above - Endonuclease Eco RI restriction map of the *S.cerevisiae* rDNA repeating unit. Arrangement of the Eco RI fragments A to G and the location of rRNA genes are indicated.
Below - Expanded map of the 25 S rRNA gene. Restriction sites are indicated by vertical lines: M, Msp I; S, Sau 3A; B, Bgl II; K, Kpn I; E, Eco RI. The subfragments obtained upon digestion of fragment Eco RI-A with endonucleases Msp I [M] and Sau 3A [S] are numbered by size. The sequenced strands are shown as horizontal arrows pointing from the labelled 5'-end.

corresponding rDNA) is shown in Figure 2. The identification of the 5'- and 3'-terminal nucleotides is based on our previous results (10). The total length of 25 S rRNA is 3392 nucleotides, yielding a relative molecular mass (M_r, Na-salt) of 1.17×10^6. This figure is smaller than the ones derived from physicochemical measurements (16,17), but it is in close agreement with results obtained by R-loop and hybridization analyses (4).

The use of the chemical method of Maxam and Gilbert (14) permitted confident reading of sequences extending 200-250 nucleotides from the labelled end. However, the anomalous behaviour of nucleotide 1642 is noteworthy. This nucleotide behaves simultaneously as C and G (data not shown). The absence of a purine band and analysis of the complementary rDNA strand identify this nucleotide as a C. Most likely, this cytidine residue is modified by *E.coli* enzymes in such a way that it is split by the G-specific dimethyl sulfate reaction. Such modification of a cytidine residue in plasmid DNA has not been encountered previously (14) and could be of more general interest.

Homology with other eukaryotic L-rRNA. The 25 S rRNA of *S.cerevisiae* is the first major eukaryotic L-rRNA with a known complete sequence. Previous indirect evidence indicates that strongly conserved regions of homology exist in eukaryotic L-rRNA (18-20). Comparison with published sequence data on the 5'-end of *Xenopus laevis* (21) and the 3'-end of *Neurospora crassa* (22) L-rRNA, reveals that such highly conserved sequences may be present at both ends of eukaryotic L-rRNA molecules. The sequence encompassing nucleotides 2-113 of *S.cerevisiae* L-rRNA shows 81 % homology with the respective sequence of *X.laevis*, while the 3'-end segment (nucleotides 3290-3392) is 74 % homologous with the respective sequence in *Neurospora crassa*.

Interaction with 5.8 S rRNA. It is well known that in the eukaryote ribosome 5.8 S rRNA is hydrogen-bonded to L-rRNA (1,18). Recently, it was reported that 5.8 S rRNA interacts with the 3'-terminal fragment of *Neurospora crassa* L-rRNA and a model of possible complementary sequences was proposed (22). Our preliminary computer search, including the full length of *S.cerevisiae* 25 S rRNA, failed to reveal uninterrupted complementary sequences between L-rRNA and 5.8 S rRNA longer than eight nucleotides. Further, linear alignment of both sequences revealed numerous regions in 25 S rRNA with 35 to 38 % complementarity to 5.8 S rRNA. The best fit is observed for the interaction between nucleotides 5-20 of L-rRNA and the 3'-end half of 5.8 S rRNA. Also, another region of high complementarity involves nucleotides 3265-3333 of L-rRNA and the 5'-end half of 5.8 S rRNA. These results appear to favor a model in which 5.8 S rRNA interacts simultaneously with segments

```
pGUUUGACCUCAAAUCAGGUAGGAGUACCCGCUGAACUUAAGCAUAUCAAUAAGCGGAGGAAAAGAAACCAACCGGAUUGCCUUAGUAACGGCGAGUGAAG  100
CGGCAAAAGCUCAAAUUUGAAAUCUGGUACCUUCGGUGCCCGAGUUGUAAUUUGGAGAGGGCAACUUUGGGGCCGUUCCUUGUCUAUGUUCCUUGGAACA  200
GGACGUCAUAGAGGGUGAGCAUCCCGUGUGGCGAGGAGUGCGGUUCUUUGUAAAGUGCCUUCGAAGAGUCGAGUUGUUUGGGAAUGCAGCUCUAAGUGGG  300
UGGUAAAUUCCAUCUAAAGCUAAAUAUUGGCGAGAGACCGAUAGCGAACAAGUACAGUGAUGGAAAGAUGAAAAGAACUUUGAAAAGAGAGUGAAAAAGU  400
ACGUGAAAUUGUUGAAAGGGAAGGGCAUUUGAUCAGACAUGGUGUUUUGUGCCCUCUGCUCCUUGUGGGUAGGGGAAUCUCGCAUUUCACUGGGCCAGCA  500
UCAGUUUUGGUGGCAGGAUAAAUCCAUAGGAAUGUAGCUUGCCUCGGUAAGUAUUAUAGCCUGUGGGAAUACUGCCAGCUGGGACUGAGGACUGCGACGU  600
AAGUCAAGGAUGCUGGCAUAAUGGUUAUAUGCCGCCCGUCUUGAAACACGGACCAAGGAGUCUAACGUCUAUGCGAGUGUUUGGGUGUAAAACCCAUACG  700
CGUAAUGAAAGUGAACGUAGGUUGGGGCCUCGCAAGAGGUGCACAAUCGACCGAUCCUGAUGUCUUCGGAUGGAUUUGAGUAAGAGCAUAGCUGUUGGGA  800
CCCGAAAGAUGGUGAACUAUGCCUGAAUAGGGUGAAGCCAGAGGAAACUCUGGUGGAGGCUCGUAGCGGUUCUGACGUGCAAAUCGAUCGUCGAAUUUGG  900
GUAUAGGGGCGAAAGACUAAUCGAACCAUCUAGUAGCUGGUUCCUGCCGAAGUUUCCCUCAGGAUAGCAGAAGCUCGUAUCAGUUUUAUGAGGUAAAGCG 1000
AAUGAUUAGAGGUUCCGGGGUCGAAAUGACCUUGACCUAUUCUCAAACUUUAAAUAUGUAAGAAGUCCUUGUUACUUAAUUGAACGUGGACAUUUGAAUG 1100
AAGAGCUUUUAGUGGGCCAUUUUUGGUAAGCAGAACUGGCGAUGCGGGAUGAACCGAACGUAGAGUUAAGGUGCCGGAAUACACGCUCAUCAGACACCAC 1200
AAAAGGUGUUAGUUCAUCUAGACAGCCGGACGGUGGCCAUGGAAGUCGGAAUCCGCUAAGGAGUGUGUAACAACUCACCGGCCGAAUGAACUAGCCCUGA 1300
AAAUGGAUGGCGCUCAAGCGUGUUACCUAUACUCUACCGUCAGGGUUGAUAUGAUGCCCUGACGAGUAGGCAGGCGUGGAGGUCAGUGACGAAGCCUAGA 1400
CCGUAAGGUCGGGUCGAACGGCCUCUAGUGCAGAUCUUGGUGGUAGUAGCAAAUAUUCAAAUGAGAACUUUGAAGACUGAAGUGGGGAAAGGUUCCACGU 1500
CAACAGCAGUUGGACGUGGGUUAGUCGAUCCUAAGAGAUGGGGAAGCUCCGUUUCAAAGGCCUGAUUUUAUGCAGGCCACCAUCGAAAGGGAAUCCGGUA 1600
AGAUUCCGGAACUUGGAUAUGGAUUCUUCACGGUAACGUAACUGAAUGUGGAGACGUCGGCGCGAGCCCUGGGAGGAGUUAUCUUUUCUUCUUAACAGCU 1700
UAUCACCCCGGAAUUGGUUUAUCCGGAGAUGGGGUCUUAUGGCUGGAAGAGGCCAGCACCUUUGCUGGCUCCGGUGCGCUUGUGACGGCCCGUGAAAAUC 1800
CACAGGAAGGAAUAGUUUUCAUGCUAGGUCGUACUGAUAACCGCAGCAGGUCUCCAAGGUGAACAGCCUCUAGUUGAUAGAAUAAUGUAGAUAAGGGAAG 1900
```

FIGURE 2. Nucleotide sequence of 25 S rRNA of *Saccharomyces cerevisiae* derived from the sequence of the 25 S rRNA gene of one rDNA repeating unit. The restriction fragment Eco RI-A terminates at nucleotide 2533. The fragment Eco RI-F is from nucleotide 2534 to nucleotide 2897.

located at both ends of the L-rRNA molecule. Such a model is compatible with the available experimental data on 5.8 S rRNA:L-rRNA interactions (22,23), but further evidence is needed in order to clarify the interaction within the ribosome of these two rRNA molecules.

ACKNOWLEDGEMENTS

The authors are greatly indebted to Dr.K.Dudov for his help in the computer analysis of sequence data and to Dr.R.J.Planta for communicating results from his laboratory prior to publication during the 1981 Cold Spring Harbour Meeting on The Molecular Biology of Yeasts.

REFERENCES

1. Hadjiolov,A.A.(1980) in Subcellular Biochemistry, Roodyn,D.B.(Ed.), Vol.7, pp.1-80, Plenum press, N.Y.
2. 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.
3. Petes,D.T., Hereford,L.M. and Skryabin,K.G.(1978) J.Bacteriol., 134, 295-305.
4. Philippsen,P., Thomas,M., Kramer,A. and Davis,R.W.(1978) J.Mol.Biol., 123, 387-404.
5. Nath,K. and Bollon,A.P.(1977) J.Biol.Chem., 252, 2562-2571.
6. Maxam,A.M., Tizard,R., Skryabin,K.G. and Gilbert,W.(1977) Nature, 267, 643-645.
7. 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.
8. 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.
9. Skryabin,K.G., Maxam,A.M., Petes,T.D. and Hereford,L.(1978) J.Bacteriol., 134, 306-309.
10. Bayev,A.A., Georgiev,O.I., Hadjiolov,A.A., Nikolaev,N., Skryabin,K.G. and Zakharyev,V.M.(1981) Nucl.Acids Res., 9, 789-799.
11. Glynn,I.M. and Chappel,J.B.(1964) Biochem.J., 90, 147-154.
12. Tanaka,T. and Weissblum,B.(1975) J.Bacteriol., 121, 354-362.
13. Maxam,A.M. and Gilbert,W.(1977) Proc.Natnl.Acad.Sci.USA, 74, 560-564.
14. Maxam,A.M. and Gilbert,W.(1980) Methods in Enzymology, 65, 499-560.
15. Szalay,A.A., Grohman,K. and Sinsheimer,R.L.(1977) Nucl.Acids Res., 4, 1569-1578.
16. Maeda,A.(1961) J.Biochem.(Tokyo), 50, 377-385.
17. Loening,U.E.(1968) J.Mol.Biol., 38, 355-363.
18. Cox,R.A.(1977) Progr.Biophys.Mol.Biol., 32, 193-231.
19. Khan,M.S.N., Salim,M. and Maden,B.E.H.(1978) Biochem.J., 169, 531-542.
20. Gourse,R.L. and Gerbi,S.A.(1980) J.Mol.Biol., 140, 321-339.
21. Hall,L.M.C. and Maden,B.E.H.(1980) Nucl.Acids Res., 8, 5993-6005.
22. Kelly,J.M. and Cox,R.A.(1981) Nucl.Acids Res., 9, 1111-1121.
23. Pace,N.R., Walker,T.A. and Schroeder,E.(1977) Biochemistry, 16, 5321-5328.
24. Veldman, G.M., Klootwijk, J., de Jonge P., Leer, R.L. and Planta, R.J. (1980) Nucl.Acids Res., 8, 5179-5192.