Volume 6 Number 7 1979 Nucleic Acids Research

Nucleotide sequence and conserved features of the 5.8 S rRNA coding region of Neurospora crassa

Eric Selker and Charles Yanofsky

Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA

Received 2 April 1979

ABSTRACT

The nucleotide sequence of Neurospora crassa 5.8 S rDNA and adjacent regions has been determined. The deduced 5.8 S rRNA sequence of Neurospora differs from the 5.8 S rRNA sequence of Saccharomyces cerevisiae at 13 of 158 residues. Nine of these differences are clustered in a segment capable of forming a short hairpin secondary structure thought to be involved in the 28 S - 5.8 S rRNA complex[1]. These differences occur in pairs such that the potential secondary structure is preserved.

INTRODUCTION

There are four rRNA species that are constituents of eukaryotic cytoplasmic ribosomes[2]. With the exception of 5 S rRNA, the rRNAs are processed from a large precursor molecule that is transcribed from tandemly repeated rDNA regions[3]. Thus the 5.8 S rRNA of Neurospora crassa is processed from the RNA segment separating 17 S and 25 S rRNA[4,5]. 5.8 S rRNAs of Saccharomyces cerevisiae[6] and several vertebrates[7-11] have been sequenced. The molecules are approximately 160 nucleotides long and their sequence is highly conserved[12]. In this report we present the nucleotide sequence of the DNA region coding for N. crassa 5.8 S rRNA and examine it in relation to a proposed secondary structure for the 5.8 S rRNA of yeast[13].

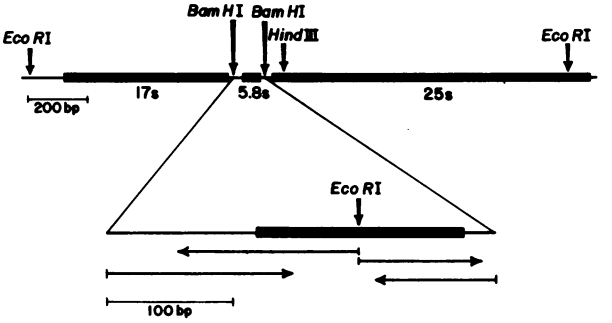
MATERIALS AND METHODS

Plasmid pMF2[5] was the source of DNA used in this study. It consists of an approximately 6000 base pair PstI generated fragment of N. crassa DNA inserted into the plasmid pBR322. DNA sequencing of 5' end-labelled restriction fragments was carried out according to the methods of Maxam and Gilbert[14]. Isolation of plasmid DNA and restriction fragments and DNA electrophoresis were performed as previously described[15]. N. crassa RNA was extracted from germinating conidia[4] as described by Rubin[16].

The RNA was treated with calf intestinal alkaline phosphatase (Boehringer Mannheim), fractionated by electrophoresis through a 10% polyacrylamide gel containing 7M urea[17] and the 5.8 S rRNA eluted as described for DNA fragments. The RNA was then labelled using T4 polynucleotide kinase (PL Biochemicals) and [γ-^{32}P] ATP (ICN Pharmaceuticals)[15]. Hybridization analyses were performed according to the method of Southern[18].

RESULTS AND DISCUSSION

Hybridization of end-labelled 5.8 S rRNA to restriction fragments of plasmid pMF2 confirmed the DNA restriction map of Free et al.[5] in the vicinity of 5.8 S rDNA (Figure 1). 5.8 S rRNA hybridized exclusively




```
          1                                                  50
5'  GTGCCGAAAC TAAACTCTTG ATATTTTATG TCTCTCTGAG TAAACTTTTA
3'  CACGGCTTTG ATTTGAGAAC TATAAAATAC AGAGAGACTC ATTTGAAAAT

                                                            100
    AATAAGTCAA AACTTTCAAC AACGGATCTC TTGGTTCTGG CATCGATGAA
    TTATTCAGTT TTGAAAGTTG TTGCCTAGAG AACCAAGACC GTAGCTACTT

                                                            150
    GAACGCAGCG AAATGCGATA GGTAATGTGA ATTGCAGAAT TCAGTGAATC
    CTTGCGTCGC TTTACGCTAT CCATTACACT TAACGTCTTA AGTCACTTAG

                                                            200
    ATCGAATCTT TGAACGCACA TTGCGCTCGC CAGTATTCTG GCGAGCATGC
    TAGCTTAGAA ACTTGCGTGT AACGCGAGCG GTCATAAGAC CGCTCGTACG

                                                        240
    CTGTTCGAGC GTCATTTCAA CCATCAAGCT CTGCTTGCGT TGGG
    GACAAGCTCG CAGTAAAGTT GGTAGTTCGA GACGAACGCA ACCC
```

Figure 1. 5.8 S rDNA of Neurospora crassa. The position of the 17 S, 5.8 S and 25 S rRNA coding regions relative to the sites of cleavage of the restriction endonucleases used in this study is shown. The DNA segment between the BamHI sites was sequenced as indicated by the arrows. The presumed 5.8 S rRNA coding region is shown in upper case.

to a 302 bp fragment generated by endonuclease BamHI. Segments of both strands of this DNA fragment were sequenced as indicated in Figure 1. The sequence of most of this fragment, including the presumed 5.8 S rRNA coding region and approximately 60 nucleotides preceding and 30 nucleotides following the 5.8 S rDNA, is presented in Figure 1. The sequence corresponding to 5.8 S rRNA (top strand in Figure 1) was recognized by its homology to the sequence of S. cerevisiae 5.8 S rRNA[6]. N. crassa 5.8 S rRNA has been fingerprinted and the pattern obtained following ribonuclease T1 digestion[5] contains oligonucleotides which migrate identically to the 5' and 3' T1 oligonucleotides of yeast 5.8 S rRNA[6]. On this basis it appears that the 5' and 3' ends of N. crassa and S. cerevisiae 5.8 S rRNA are identical. The Neurospora 5.8 S rDNA sequence between nucleotides 59 and 217 matches the yeast 5.8 S rRNA sequence, differing at only 13 of 158 residues (92% sequence conservation). This is consistent with the extent of sequence conservation seen in the other 5.8 S rRNAs which have been sequenced. Thus, the sequences of 5.8 S rRNA from humans[11], rats[7], chickens[11], turtles[8], frogs[10] and fish[9] are all greater than 96% identical and are homologous to the sequence of 5.8 S rRNA of yeast at approximately 70% of the residues. This pattern of conservation of the sequence of 5.8 S rRNAs parallels the phylogenetic tree derived from comparisons of the amino acid sequences of cytochrome c from the same organisms[19].

Figure 2 presents the predicted sequence of Neurospora 5.8 S rRNA arranged in the form of the "cloverleaf" secondary structure recently proposed for 5.8 S rRNA[13]. There are 13 differences from the yeast sequence, each of which could have arisen as a result of a single base pair substitution. Of the 13 differences, 11 do not change the base pairing pattern in the RNA structure. Most of the differences are within a G-C rich arm which is a feature of all proposed 5.8 S rRNA secondary structures[6,7,10,13]. Although 92% of the overall Neurospora and yeast 5.8 S rRNA sequences are identical, only 9 of 18 nucleotides base paired in the G-C rich arm are conserved. Nevertheless, the differences in this region are such that base pairing is conserved. This fact suggests that secondary rather than primary structure is functionally important in this segment of the RNA molecule. Consistent with this view are the findings of Pace et al.[1] indicating that in mouse 5.8 S rRNA the secondary structure of this G-C rich arm may be involved in stabilizing the formation of a 28 S - 5.8 S rRNA complex by coaxial helix stacking.

The base pair differences in the G-C rich arms of yeast and Neurospora

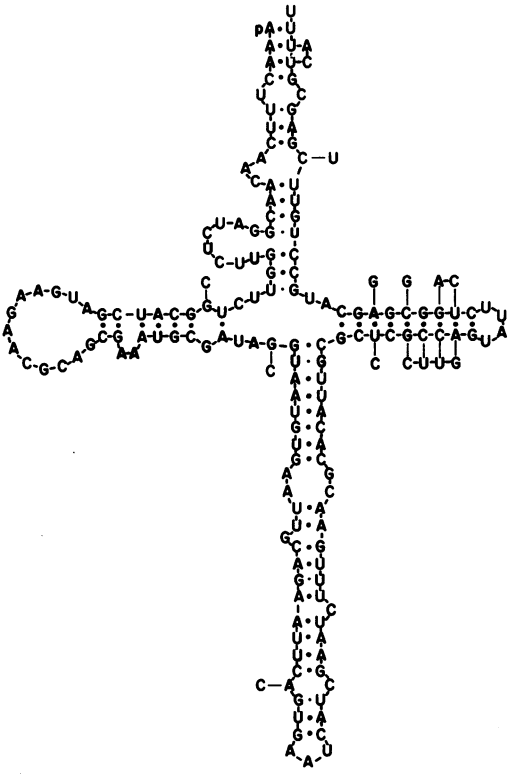


Figure 2. Possible secondary structure of Neurospora crassa 5.8 S rRNA. The 13 residues of the Saccharomyces cerevisiae 5.8 S rRNA sequence which differ from the predicted Neurospora sequence are shown adjacent to the secondary structure.

5.8 S rRNA also pose an interesting evolutionary problem. Single base pair mismatches presumably preceded the double, complementary changes. If base pairing in the G-C rich arm is functionally essential, mismatched intermediates would have been less fit structures. Consideration of the nature of the base changes reveals the interesting fact that of the 9 changes in this region, only 2 were transversions. The remaining 7 transitions could have gone through G-U pair intermediates, as diagrammed.

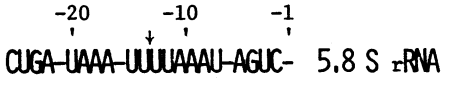
```
G      G ↔ A
|||    ·   ‖
C ↔ U      U
```

G-U pairs, though not as stable as G-C or A-U pairs, are not destabi-

lizing like the other mismatches[20]. In fact, a recent thermodynamic study on the G-C rich arm of yeast 5.8 S rRNA suggests that a G-U base pair in the arm does enhance its stability[21]. The 2 transversions resulting in G-C ↔ C-G conversions presumably went through a mis-paired intermediate. The fact that the rRNA genes are moderately repeated (about 150 copies per genome in both yeast[22] and Neurospora[23]) may explain how such potentially defective intermediates could be tolerated. Presumably most of the 5.8 S rRNA genes of an organism with some mis-paired intermediates would code for perfectly paired structures. A low level of sequence heterogeneity (≈ 0.01) at one position of the 5.8 S rRNA from Xenopus laevis has been detected[10].

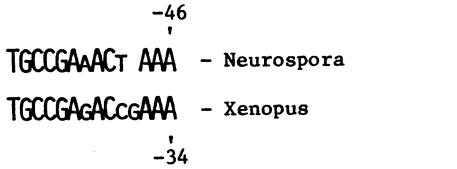
The colinearity of Neurospora 5.8 S rDNA with the yeast 5.8 S rRNA sequence and the similarity of the fingerprints of T1 ribonuclease digests of the two rRNAs[5,6] suggest that the Neurospora 5.8 S rRNA gene does not contain an intervening sequence.

The 5.8 S rRNA is derived from the segment of the rRNA precursor separating the larger rRNAs[3]. It seems likely that the non-coding RNA in this segment is involved in the processing of this molecule. We scanned the sequences surrounding the Neurospora 5.8 S rRNA for unusual features which might be recognized by processing enzymes. The most distinctive sequence feature adjacent to the 5.8 S rRNA coding region is an imperfect symmetrical sequence involving 20 out of 24 nucleotides immediately preceding the 5.8 S rRNA.



(a dash indicates a non-symmetrical nucleotide)

Bell et al.[24] have sequenced the 16 nucleotides preceding the yeast 5.8 S rRNA coding region and 10 of these nucleotides are identical to those preceding the Neurospora 5.8 S rRNA coding region including the hexanucleotide, UUUAAA. Boseley et al.[25] have recently sequenced the 5.8 S rRNA coding region of Xenopus laevis. The sequences flanking the Neurospora and Xenopus 5.8 S rRNA genes are generally different except for the homology:



This occurs 46 nucleotides preceding the presumed Neurospora 5.8 S rRNA region and 34 nucleotides preceding the Xenopus 5.8 S rRNA region. It will be of interest to learn whether special sequences or higher structures such as those implicated in E. coli 16 S rRNA processing[26, 27] are involved in 5.8 S rRNA processing.

ACKNOWLEDGEMENTS

This work was supported by grants from the United States Public Health Service (GM-09738) and the National Science Foundation (PCM77-24333). E.S. is a predoctoral trainee of the United States Public Health Service and C.Y. is a Career Investigator of the American Heart Association. This study was carried out in accordance with the National Institutes of Health Guidelines for Recombinant DNA Research. We thank Drs. Stephen Free and Robert Metzenberg for providing us with the Neurospora rDNA clone, pMF2, and its restriction map, as well as helpful information.

REFERENCES

1. Pace, N.R., Walker, T.A. and Schroeder, E. (1977) Biochem. 16, 5321-5328.
2. Pene, J., Knight, Jr., E., and Darnell, Jr., J.E. (1968) J. Mol. Biol. 33, 609-623.
3. Perry, R. P. (1976) Ann. Rev. Biochem. 45, 605-629.
4. Russell, P.J., Hammett, J.P. and Selker, E.U. (1976) J. Bacteriol. 127, 785-793.
5. Free, S.J., Rice, P.W. and Metzenberg, R.L. (1979) J. Bacteriol. 137, 1219-1226.
6. Rubin, G.M. (1973) J. Biol. Chem. 248, 3850-3875.
7. Nazar, R.N., Sitz, T.O. and Busch, H. (1975) J. Biol. Chem. 250, 8591-8597.
8. Nazar, R.N. and Roy, K.L. (1976) FEBS Lett. 72, 111-116.
9. Nazar, R.N. and Roy, K.L. (1978) J. Biol. Chem. 253, 395-399.
10. Ford, P.J. and Mathieson, T. (1978) Eur. J. Biochem. 87, 199-214.
11. Khan, M.S.N. and Maden, B.E.H. (1977) Nucleic Acids Res. 4, 2495-2505.
12. Erdmann, V.A. (1979) Nucleic Acids Res. 6, r29-r44.
13. Luoma, G.A. and Marshall, A.G. (1978) Proc. Natl. Acad. Sci. 75, 4901-4905.
14. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. 74, 560-564.
15. Bennett, G.N., Schweingruber, M.E., Brown, K.D., Squires, C. and Yanofsky, C. (1978) J. Mol. Biol. 121, 113-137.
16. Rubin, G.M. (1975) In Methods in Cell Biology Prescott, D.M., Ed., pp. 45-64, Academic Press, New York.
17. Maniatis, T., Jefferey, A. and van de Sande, H. (1975) Biochemistry 14, 3787-3794.
18. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
19. Dayhoff, M.O., Park, C.M. and McLaughlin, P.J. (1972) In Atlas of Protein Sequence and Structure. Vol. 5, Dayhoff, M.O., Ed., pp. 7-16. National Biomedical Research Foundation, Washington, D. C.
20. Tinoco, I. Jun., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature N. Biol. 246, 40-41.

21. Lightfoot, D. (1978) Nucleic Acids Res. 5, 3565-3577.
22. Schweizer, E., MacKechnie, E. and Halvorson, H.O. (1969) J. Mol. Biol. 40, 261-277.
23. Chattopadhyah, S.K., Kohne, D.E. and Dutta, S.K. (1972) Proc. Natl. Acad. Sci. 69, 3256-3259.
24. 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.
25. Boseley, P.G., Turyns, A. and Birnstiel, M.L. (1978) Nucleic Acids Res. 5, 1121-1137.
26. Young, R.A. and Steitz, J.A. (1978) Proc. Natl. Acad. Sci. 75, 3593-3597.
27. Dahlberg, A.E., Dahlberg, J.E., Lund, E., Tokimatsu, H., Rabson, A.B., Calvert, P.C., Reynolds, F. and Zahalak, M. (1978) Proc. Natl. Acad. Sci. 75, 3598-3602.