Volume 9 Number 5 1981 Nucleic Acids Research

# The nucleotide sequence at the 3'-end of Neurospora crassa 25S-rRNA and the location of a 5.8S-rRNA binding site

John M.Kelly and Robert A.Cox

National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Received 18 December 1980

ABSTRACT

The sequence of 110 nucleotides adjacent to the 3'-end of Neurospora crassa 25S-rRNA has been derived by chemical sequencing methods. Sequences present between 40 and 85 nucleotides of the 3'-end were found to complement sequences at the 3'-and 5'-ends of 5.8S-rRNA. Interaction was shown to occur between 5.8S-rRNA and a specific 3'-terminal fragment of 85 nucleotides derived from 25S-rRNA. We have also demonstrated that the nucleotide sequence at the 3'-end of N. crassa 5.8S-rRNA (-$UCAUU_{OH}$) is different from the published sequence (-$UUUU_{OH}$) which was derived from rDNA.

## INTRODUCTION

The roles that rRNA species play in protein biosynthesis have yet to be fully defined. One way of identifying regions of an rRNA species that may fulfil a role in either ribosome assembly or function is to seek those sequences that have resisted evolutionary divergence. Thus the rRNA component of the smaller subparticle of eukaryotic ribosomes (S-rRNA) has a nucleotide sequence at the 3'-end that has been conserved during evolution [1,2]. Sequence homologies were found near to the 3'-end of the major rRNA component of the larger subparticle (L-rRNA) of moulds and amphibians[3]. The 5.8S-rRNA component that is found in association with L-rRNA also has a nucleotide sequence that has been highly conserved during evolution; the sequence is identical among mammals[4,5], and the sequences of mammalian and yeast 5.8S-rRNA[6] are at least 70% homologous.

The interaction in vivo between L-rRNA and 5.8S-rRNA survives the extraction procedure using cold phenol, but the complex dissociates under conditions that destabilise base-pairs. A knowledge of the nature and location within L-rRNA of the binding site for 5.8S-rRNA would lead to greater precision in our view of ribosome structure.

We now report the sequence of approximately 110 nucleotides at the 3'-end of N. crassa L-rRNA. This information is of interest not only because it is

one step in the search for conserved sequences, but also because it is an essential step in identifying the 3'-end of the L-rRNA gene. Evidence from Xenopus and yeast rDNA[7,8] suggests this to be close to the sequence signalling termination of transcription. The secondary structure of this 110 nucleotide sequence is also of interest because of its possible role in processing pre-rRNA and in the maturation of L-rRNA. Sequences at the 3'-end of N. crassa 5.8S-rRNA were also identified.

The principal result of the present study is that a fragment of L-rRNA, 85 nucleotides long, carrying the 3'-end, spans a potential binding site for 5.8S-rRNA. A complex formed between these components was detected.

## METHODS

Materials. T4 RNA ligase was purchased from Miles Biochemicals; [5'-$^{32}$P]pCp (2000-3000 Ci/mmol) was purchased from the Radiochemical Centre, Amersham, Bucks.

Preparation and labelling of rRNA. Cytoplasmic RNA was isolated from N. crassa (strain 74A) after disruption of the hyphae with acid-washed sand followed by cold phenol extraction. L-rRNA was isolated on 10% (w/v) to 25% (w/v) sucrose gradients in 0.1M LiCl/0.01M EDTA/0.01M Tris/HCl (pH 7.4)/SDS (0.2% w/v) run at 23,000 rev/min at 20°C in a Beckman SW27 rotor. 5.8S-rRNA was removed from L-rRNA by heating to 60°C for 5 min, followed by rapid cooling. It was isolated after re-running on a sucrose gradient.

The 3'-terminal labelling reaction was carried out in a solution of 50mM Hepes, pH 7.5, 3.3mM dithiothreitol, 15mM $MgCl_2$, 10% dimethyl sulphoxide, 0.01mg serum albumin per ml. The reaction mixture (15µl) contained 120µg of L-rRNA, 20 units of T4 RNA ligase, 0.7nmoles of ATP and 100µC of pCp. Incubation was at 0°C for 16-24h. Labelled RNA was purified on a 1% agarose gel and eluted by homogenising the agarose band in 0.5M ammonium acetate/0.1% SDS/0.1mM EDTA. Labelled RNA was stored at -20°C in deionised sterile water. 5.8S-rRNA was labelled as described by Peattie[9].

Sequencing reactions and gels. The chemical sequencing reactions were carried out as described by Peattie[9]. Fragments were separated on thin 60cm polyacrylamide gels (generally 8% or 12%) at high voltage by electrophoresis for various times, dependent on the degree of separation required.

Isolation of the 85 nucleotide 3'-terminal fragment. A set of oligonucleotide fragments was generated from N. crassa L-rRNA by using one of the chemical sequencing reactions (normally the A>G reaction). Oligonucleotide fragments containing the 3'-terminus were specifically labelled using T4 RNA

ligase and [5'-$^{32}$P]pCp, since only these fragments contain the free 3'-OH group required for the reaction. The labelling reaction was done in 10µl of reaction buffer (as used before) containing 100µg of chemically cleaved 25S-rRNA, 30µC of [5'-$^{32}$P]pCp, 5 units of T4 RNA ligase, 0.3nmoles of ATP. Incubation was for 16-24h at 0°C. The reaction was halted by adding 2 volumes of 10M urea/5mM Tris borate, pH 8.3/0.1mM EDTA/0.05% xylene cyanol/ 0.05% bromophenol blue and then heating to 90°C for 30 sec. This mixture was layered on a 60cm polyacrylamide gel and electrophoresis carried out for sufficient time to allow separation of the required fragments. X-ray film was placed on the gel and exposed for 3-4h. After comparison with a previous sequencing gel, the desired 3'-terminal fragment was identified, then excised from the gel and eluted.

Interaction of 5.8S-rRNA with the 3'-terminal fragment from L-rRNA. 5.8S-rRNA (2µg) was dissolved in 30µl of 0.4M KCl, 6mM MgAc, 10mM β-mercapto-ethanol, 50mM Tris/HCl (pH 7.5), and added to the isolated 3'-terminal fragment which had been freeze dried. The mixture was heated to 65°C for 5 min and cooled to 20°C over 30 min. The reaction mixture was kept on ice for 15 min and dye (5µl) was added prior to layering the solution on an 8% poly-acrylamide gel as described[10]. Electrophoresis was carried out at 4°C, then the gel was dried and autoradiographed.

## RESULTS

Sequence at the 3'-end of L-rRNA. Using direct chemical sequencing methods, it was possible to sequence 110 nucleotides at the 3'-end of *N. crassa* L-rRNA (see Table 1). Typical sequencing gels are shown in Fig. 1. Compression of bands on sequencing gels can be caused by the effect of secondary structure on the rate of migration of oligonucleotide fragments. This effect caused few problems in sequencing L-rRNA, although on one occasion it was noticed to affect the migration of fragments between 26 and 30 nucleotides long. Subsequent inspection of the sequence suggested the possibility of secondary structure involving nucleotides 8 to 27 (see Fig. 3). Cleavage at nucleotide G14 was found to be less efficient than usual, although we cannot explain this. The only nucleotide in the sequence to give a consistently ambiguous result was no. 84. Here a cleavage was found to have occurred irrespective of the specificity of the sequencing reaction. This may signify the presence at this point in the sequence of a methylated base which exists at least for some of the time in a reduced form. Thus the RNA chain would be rendered susceptible to aniline-catalysed cleavage at this

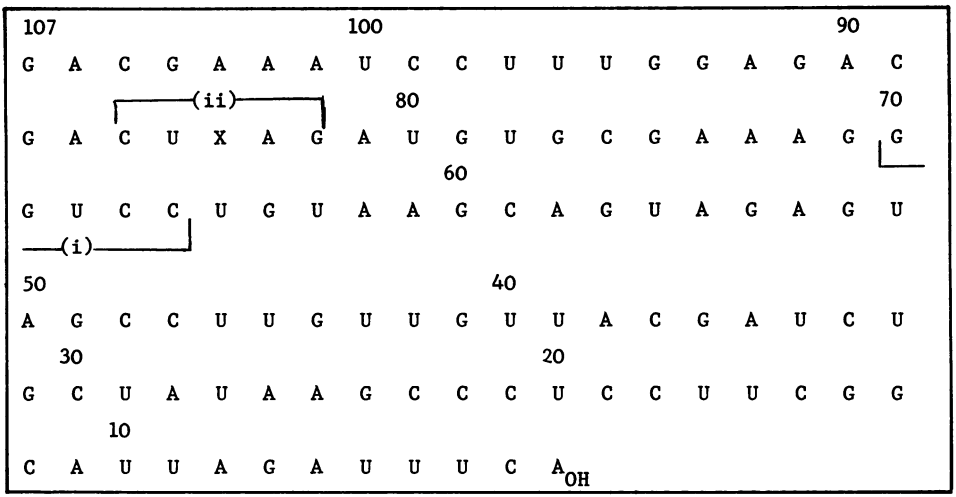


Table 1. Nucleotide sequence at the 3'-end of N. crassa L-rRNA. The nucleotides are numbered from the 3'-end. The nucleotide at position 84 (signified by X) continually gave ambiguous results on sequencing gels. Positions of two restriction sites in rDNA are given: (i) ECORII, (ii) Hinf I.

nucleotide [11].

When the sequences of 100 nucleotides at the 3'-ends of yeast[8] and N. crassa L-rRNA are compared, 80% homology is found. The most striking features are stretches of 37 nucleotides (nos. 29-65) and 15 nucleotides (79-90) which have been completely conserved. The presence of highly conserved sequences in these regions suggest that they are functionally important.

A comparison of the sequence at the 3'-end of L-rRNA from N. crassa and Escherichia coli[12] reveals about 60% homology after taking into account deletions and substitutions. The regions of homology are less extensive than those at the 3'-end of S-rRNA from E. coli[13,14] and several other eukaryotic species[1,2].

Partial sequence of 5.8S-rRNA. We have also obtained sequence data for about half the N. crassa 5.8S-rRNA molecule as read from the 3'-end. They are consistent with the sequence obtained from rDNA[15] with the exception of two nucleotides adjacent to the 3'-end (see Table 2). A group of about 6 nucleotides in 5.8S-rRNA proved difficult to sequence due to compression of bands on the gel. This occurred despite the presence of 8M urea, prior heating to 90°C and electrophoresis at high temperature. The compression of bands on the gel corresponds to the position in the sequence of nucleotides

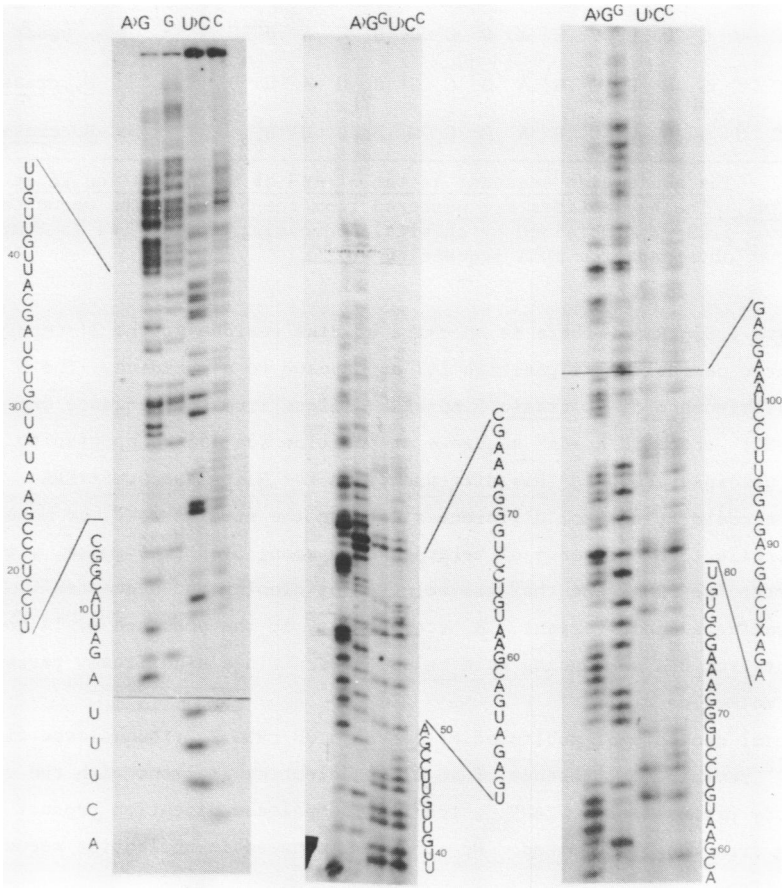


Fig. 1. Autoradiographs of sequencing gels of the 3'-end of N. crassa L-rRNA obtained by chemical sequencing methods. The RNA was labelled at the 3'-end, chemically treated and layered on 12% (left) or 8% (centre and right) polyacrylamide gels as described in Methods. Electrophoresis was carried out at 1.2KV for 4h (left), 8h (centre) and 16h (right). Nucleotides are numbered from the 3'-end.

which form a very stable GC rich helix (nucleotides 137-116). This feature of the secondary structure is common to all known 5.8S-rRNA molecules.

The nucleotide sequences of yeast and N. crassa 5.8S-rRNA are more than 90% homologous. Table 2 shows the nucleotide sequence adjacent to the 3'-end of (i) yeast 5.8S-rRNA, (ii) N. crassa 5.8S-rRNA, (iii) N. crassa 5.8S-rRNA as derived from rDNA sequence data. In this 3'-terminal region the nucleotide sequence of yeast 5.8S-rRNA differs at two points from that of

| 140 | | | | | | | | | | 150 | | | | | | | 158 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| G | C | C | U | G | U | U | U | G | A | G | C | G | U | C | A | U | U | $U_{OH}$ | yeast [6] |
| G | C | C | U | G | U | U | C | G | A | G | C | G | U | C | A | U | $U_{OH}$ | | N. crassa (a) |
| G | C | C | U | G | U | U | C | G | A | G | C | G | U | U | U | $U_{OH}$ | | | N. crassa (b) [15] |

Table 2. The nucleotide sequence at the 3'-end of N. crassa and yeast 5.8S-rRNA. The nucleotides are numbered from the 5'-end. The sequence of N. crassa 5.8S-rRNA obtained by chemical sequencing methods (a) is compared with that obtained from rDNA sequencing (b).

N. crassa 5.8S-rRNA; there is an extra uridine residue at the 3'-terminus and a uridine residue at position 147 as opposed to a cytidine. The 3'-terminal sequence of N. crassa 5.8S-rRNA differs from the sequence derived from rDNA; there is a -CA- sequence at position 154-155 as opposed to a single uridine residue. This discrepancy in the N. crassa 5.8S-rRNA sequence could arise from differences between the strains used, or more likely reflect the presence of heterogeneity among 5.8S-rRNA genes, with the particular rRNA gene that was selected by cloning and sequenced[15] representing a minor variant. A heterogeneity in the sequence would not be apparent from RNA sequencing gels unless present in a significant percentage of the molecules.

Careful analysis of published N. crassa 5.8S-rRNA $T_1$ ribonuclease fingerprints[16] reveals the presence of an oligonucleotide fragment with the migratory properties of $UCAUU_{OH}$, the $T_1$ ribonuclease digestion product arising from the 3'-terminus. No other differences in nucleotide sequence were noticed.

Interaction between 5.8S-rRNA and an 85 nucleotide fragment containing the 3'-end of L-rRNA. Recently, a method has been reported which demonstrated association between yeast L-rRNA and an oligonucleotide fragment from the 5'-end of 5.8S-rRNA[10]. Using a similar approach, we attempted to detect interaction between N. crassa 5.8S-rRNA and a $^{32}P$-labelled fragment from L-rRNA which contained the 3'-terminus. This fragment also contained those nucleotides whose sequence suggests the possibility of complementary base-pairing with nucleotides at the 3'- and 5'-ends of 5.8S-rRNA. The RNA mixture was incubated as described in Methods and subjected to electrophoresis on an 8% polyacrylamide gel which was then autoradiographed (Fig. 2). In those samples which contained 5.8S-rRNA, a slow migrating band could be detected in the gel. The position of this band was consistent with that of a complex formed between 5.8S-rRNA and the L-rRNA fragment. The extent

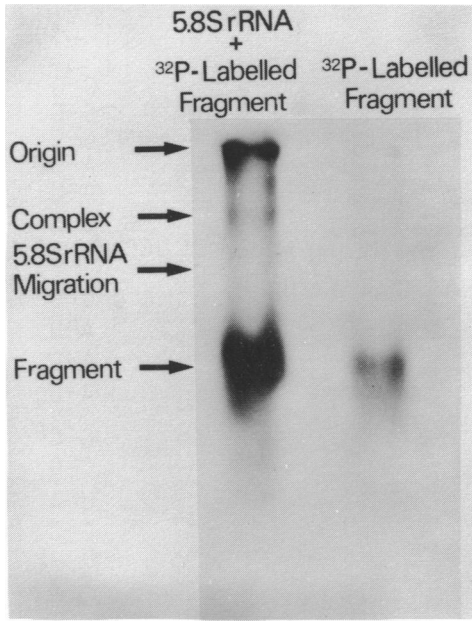


Fig. 2. Reassociation of an 85 nucleotide fragment from the 3'-end of L-rRNA with 5.8S-rRNA. The incubation mixture is described in Methods. After electrophoresis and autoradiography, a band was located in a position consistent with complex formation. A fainter band was also detected nearer to the origin. This band may have arisen from complex formation between two 3'-terminal fragments and one 5.8S-rRNA molecule. This possible explanation is based on the proposal that 5.8S-rRNA contains two binding sites, one at the 3'-end and one at the 5'-end.

of interaction that was observed was quantitatively less than has been demonstrated between intact L-rRNA and 5.8S-rRNA. It may be that other features in intact L-rRNA are necessary to obtain efficient re-association and that other regions also interact with 5.8S-rRNA. However, the interaction appears to be specific and therefore an important factor in ribosomal RNA structure and function.

## DISCUSSION

Secondary structure at the 3'-end of N. crassa L-rRNA. By maximising base-pairing within the sequence of 110 nucleotides at the 3'-end of L-rRNA it is possible to envisage a secondary structure which incorporates four helices and three hairpin loops (Fig. 3). Of these features, helices (i)

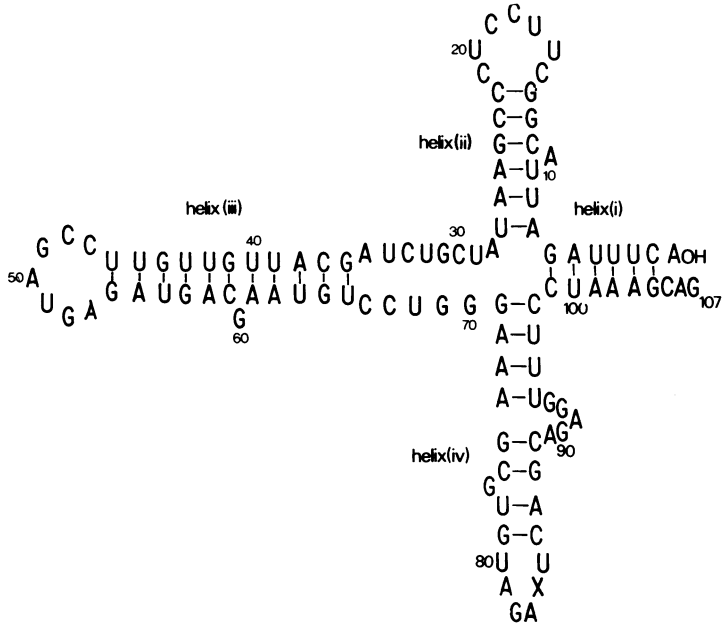


Fig. 3. Secondary structure model for the 3'-end of N. crassa L-rRNA. The nucleotides are numbered from the 3'-end.

and (ii) are the most stable with calculated free energies of -9Kcal and -8Kcal respectively[17]. The other two helices (iii) and (iv) are only marginally energetically favourable with free energies of about -1Kcal. This model serves to illustrate the potential for forming bihelical secondary structure within this 110 nucleotide sequence. Other base-pairing schemes may be possible.

Secondary structure of N. crassa 5.8S-rRNA. Several models have recently been proposed to accommodate the secondary structure of 5.8S-rRNA. Studies using chemical probes[18] suggest that the model proposed by Nazar et al.[19] is a good approximation. N. crassa 5.8S-rRNA can incorporate the general features of this model (Fig. 4). Interaction with L-rRNA probably involves an alteration in at least some of these features. Nucleotides in helix (a) have been implicated in complex formation with L-rRNA in rat[20] and yeast[10]. In this helix instability is generated by the presence of several imperfections in base-pairing. The calculated free energy of helix (a) in N. crassa 5.8S rRNA is -7Kcal.

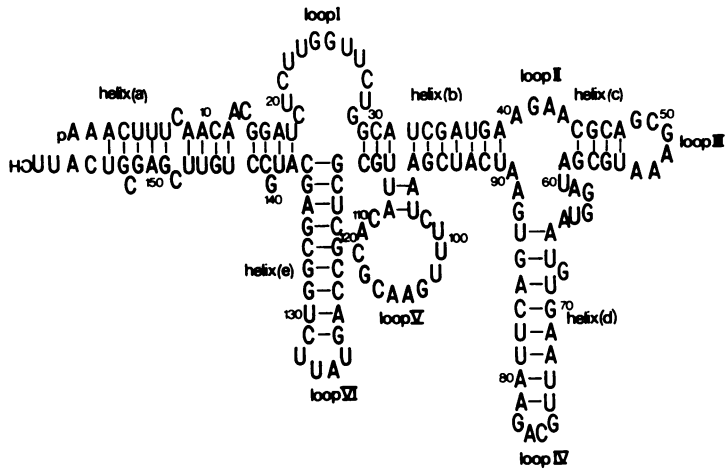


Fig. 4. Secondary structure model for N. crassa 5.8S-rRNA based on that of Nazar et al.[4]. The nucleotides are numbered from the 5'-end. Nomenclature of helical and loop regions are according to reference [18].

5.8S/L-rRNA interaction. An investigation of sequences close to the 3'-end of N. crassa L-rRNA revealed the potential for complementary base-pairing with nucleotides near the 3'- and 5'-ends of 5.8S-rRNA, i.e. within helix (a). From this finding we have been able to propose a model for the structure of this 5.8S-rRNA binding site (Fig. 5). The interaction would arise from the unfolding of features of the secondary structures of L-rRNA and 5.8S-rRNA and the formation of two more stable intermolecular helices with free energies of -18Kcal and -13Kcal.

Consideration of these secondary structure models (Figs. 3-5) has led us to suggest a hypothetical scheme for the processing of 5.8S-rRNA leading to its association with the 3'-end of L-rRNA. Within the large primary transcript product, the 5.8S-rRNA sequence and the nucleotides at the 3'-end of the L-rRNA sequence probably both maintain particular conformations (e.g. as suggested in Figs. 3 and 4). Features within these structures will govern the enzymatically catalysed processing steps. At some stage during processing, the association between the 3'-end of L-rRNA and 5.8S-rRNA is achieved by the unfolding of helix (a) (in 5.8S-rRNA) and helices (ii) and (iii) (in L-rRNA) and the formation of a more energetically favourable junction complex. This interaction may also involve nucleotides from regions of both molecules other than those specified. In this form the

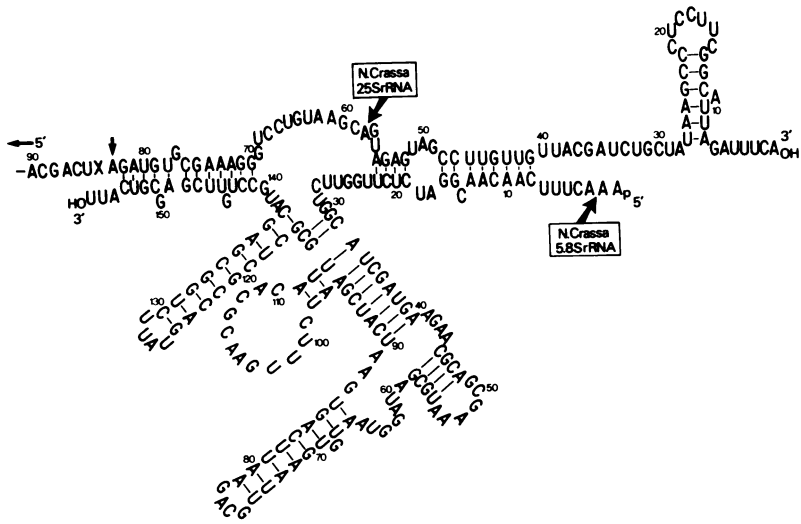


Fig. 5. Proposed interaction between 5.8S-rRNA and the 3'-end of L-rRNA. The nucleotides in 5.8S-rRNA are numbered from the 5'-end, whereas those in L-rRNA are numbered from the 3'-end. Nucleotides 2-7 in L-rRNA are drawn as single-stranded, although they could continue to interact with nucleotides 101-106 as shown in Fig. 3. The arrow ( ) indicates the size of the oligonucleotide fragment used for interaction with 5.8S-rRNA (see Fig. 2).

5.8S/L-rRNA complex would carry out its role(s) in the structure and function of the mature ribosome.

This model of 5.8S/L-rRNA interaction in N. crassa also holds for yeast. A very similar junction complex can be formed. The high degree of sequence homology found at the 3'- and 5'-ends of yeast[6] and N. crassa[15] 5.8S-rRNA is reflected in the presence of highly conserved sequences in L-rRNA[8] in the regions of the proposed binding site (nucleotides 41-83). At one of the regions of interaction (nucleotides 41-56), the sequences are identical and are only different by 3 nucleotides at the other (nucleotides 70-82).

An analogous interaction within prokaryotic L-rRNA. We believe that there is particular significance in the existence of homologies within the stretch of 100 nucleotides at the 3'-ends of N. crassa and E. coli L-rRNA. Nazar[21] has drawn attention to the presence at the 5'-end of E. coli L-rRNA of '5.8S-rRNA-like' sequences. By analogy with the location of 5.8S-rRNA binding site in N. crassa L-rRNA, we sought and found a potential binding site at the 3'-end of E. coli L-rRNA for these '5.8S-rRNA-like' sequences.

A scheme was derived for interaction between the 3'- and 5'-ends of E. coli L-rRNA that has a high degree of homology with that shown in Fig. 5 for the complex formed between N. crassa L-rRNA and 5.8S-rRNA. We also note that an analogous complex can be formed between the 3'- and 5'-ends of human mitochondrial L-rRNA[22].

These observations lend support to the idea that sequences at the 5'-end of prokaryotic L-rRNA carry out the functions during protein biosynthesis which are performed in eukaryotic ribosomes by 5.8S-rRNA. This notion considered together with the location of the 5.8S-rRNA gene, upstream from the 5'-end of the L-rRNA gene, leads to the view that the spacer region between these two genes can be regarded as an intron. This proposal and its implications will be more fully developed elsewhere.

We thank S. Kotecha for skilled technical assistance.

REFERENCES

1. Samols, D.R., Hagenbuchle, O. and Cage, L.P. (1979) Nucleic Acid Res. 7, 1109-1119
2. Azad, A.A. and Deacon, N.J. (1980) Nucleic Acid Res. 8, 4365-4377
3. Cox, R.A. and Thompson, R.D. (1980) Biochem. J. 187, 75-90
4. Nazar, R.N., Sitz, T.O. and Busch, H. (1975) J. Biol. Chem. 250, 8591-8597
5. Khan, N.S.N. and Maden, B.E.H. (1977) Nucleic Acid Res. 4, 2495-2505
6. Rubin, G.M. (1973) J. Biol. Chem. 248, 3860-3875
7. Sollner-Webb, B. and Reeder, R.H. (1979) Cell 18, 485-499
8. Veldman, G.M., Klootwijk, J., de Jange, P. and Planta, R.J. (1980) Nucleic Acid Res. 8, 5179-5192
9. Peattie, D.A. (1979) Proc. Nat. Acad. Sci. USA 76, 1760-1764
10. Nazar, R.N. and Sitz, T.O. (1980) FEBS Lett. 115, 71-76
11. Wintermeyer, W. and Zachau, H.G. (1975) FEBS Lett. 58, 306-309
12. Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc. Nat. Acad. Sci. USA 77, 201-204
13. Young, R.A. and Steitz, J.A. (1978) Proc. Nat. Acad. Sci. USA 75, 3593-3597
14. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Nat. Acad. Sci. USA 75, 4801-4805
15. Selker, E. and Yanofsky, C. (1979) Nucleic Acid Res. 6, 2561-2567
16. Free, S.J., Rice, P.W. and Metzenberg, R.L. (1979) J. Bact. 127, 785-793
17. Tinoco, I., Borer, P.N., Dengler, B., Levine, M.D., Uhlenbeck, O.C., Crothers, D.M. and Gralla, J. (1973) Nature New Biol. 246, 40-41
18. Kelly, J.M. and Maden, B.E.H. (1980) Nucleic Acid Res. 8, 4521-4535
19. Nazar, R.N., Sitz, T.D. and Busch, H. (1975) J. Biol. Chem. 250, 8591-8597
20. Pace, N.R., Walker, T.A. and Schroeder, E. (1977) Biochemistry 16, 5321-5328
21. Nazar, R.N. (1980) FEBS Lett. 119, 212-214
22. Eperon, I.C., Anderson, S. and Nierlich, D.P. (1980) Nature 286, 460-467