Volume 7 Number 8 1979 Nucleic Acids Research

Sequence and secondary structure of Drosophila melanogaster 5.8S and 2S rRNAs and of the processing site between them

George N.Pavlakis*, Bertrand R.Jordan+, Regina M.Wurst* and John N.Vournakis*

*Department of Biology, Syracuse University, Syracuse, NY 13210, USA and +Centre de Biochimie et de Biologie Moleculaire, CNRS, 31 chemin Joseph Aiguier, 13274 Marseille, Cedex 2, France

Received 19 September 1979

ABSTRACT

Drosophila melanogaster 5.8S and 2S rRNAs were end-labeled with ^{32}P at either the 5'or 3' end and were sequenced. 5.8S rRNA is 123 nucleotides long and homologous to the 5' part of sequenced 5.8S molecules from other species. 2S rRNA is 30 nucleotides long and homologous to the 3' part of other 5.8S molecules. The 3' end of the 5.8S molecule is able to base-pair with the the 5' end of the 2S rRNA to generate a helical region equivalent in position to the "GC-rich hairpin" found in all previously sequenced 5.8S molecules. Probing the structure of the labeled Drosophila 5.8S molecule with S1 nuclease in solution verifies its similarity to other 5.8S rRNAs. The 2S rRNA is shown to form a stable complex with both 5.8S and 26S rRNAs separately and together. 5.8S rRNA can also form either binary or ternary complexes with 2S and 26S rRNA. It is concluded that the 5.8S rRNA in Drosophila melanogaster is very similar both in sequence and structure to other 5.8 rRNAs but is split into two pieces, the 2S rRNA being the 3' part. 2S anchors the 5.8S onto the 26S rRNA. There are additional sites of interaction between the 5.8S and 26S rRNA. The order of the rRNA coding regions in the ribosomal DNA repeating unit is shown to be 18S - 5.8S - 2S - 26S. Direct sequencing of ribosomal DNA shows that the 5.8S and 2S regions are separated by a 28 nucleotide spacer which is A-T rich and is presumably removed by a specific processing event. A secondary structure model is proposed for the 26S-5.8S-2S ternary complex and for the presumptive precursor molecule.

INTRODUCTION

The large subunit of eucaryotic ribosomes contains one RNA molecule of approximately 28S, one 5S molecule of 120 nucleotides, and one 5.8S molecule approximately 160 nucleotides long which is base-paired with the 28S rRNA. Several 5.8S molecules from different species have been sequenced (1-3). It is known that the gene for 5.8S rRNA lies within the transcribed spacer that separates the 18S from the 28S rRNA in the rDNA repeating unit in all cases examined thus far (3-11). It is well established that the rRNA is transcribed as a long precursor molecule which is processed to mature 18S, 5.8S and 28S rRNA species by a series of specific endonucleolytic cleavages

(for reviews, see 12, 13). Processing involves re-ligation of RNA fragments in many cases studied and may be a general phenomenon in eukaryotes.

Drosophila melanogaster ribosomes appear to be different than other eukaryotic ribosomes. Instead of having a 5.8S molecule of approximately 160 nucleotides they have two shorter RNAs associated with the large ribosomal subunit. One is 30 nucleotides long and is named 2S rRNA (14, 15). The other is approximately 120 nucleotides long and co-migrates with 5S rRNA on polyacrylamide gels. We refer to this molecule as mature 5.8S rRNA (m5.8S rRNA) based on evidence presented in this paper.

We report here the nucleotide sequences of m5.8S and 2S rRNAs. The secondary structure of these molecules is also studied using partial S1 nuclease digestion of the end-labeled RNAs and identification of the exposed nucleotides on sequencing gels (16, 17). It is found that 2S is homologous to the 3' part of 5.8S rRNAs from other species and that it forms a complex with m5.8S to produce a molecule equivalent in structure to the other 5.8S rRNAs. DNA sequencing of the m5.8S-2S region agrees with this interpretation and reveals that there is a 28 nucleotide long spacer between the coding sequences for the m5.8S and 2S rRNAs. It is likely that this spacer is transcribed and subsequently removed during rRNA maturation (18). In an effort to identify structural features that may be recognized by processing enzymes, the secondary structure of the processing site at which the transcribed spacer is excised has been determined.

MATERIALS AND METHODS

Enzymes. Calf intestine alkaline phosphatase (Boehringer-Mannheim) was further purified both by Sephadex G-75 chromatography and treatment with diethyl pyrocarbonate (Sigma) as described (19). T4 polynucleotide kinase (Boehringer Mannheim), T1 (Sankyo), U2 (R-Plus), P1 (P-L Biochemicals) and Pancreatic (Worthington) ribonucleases were used without further purification. Physarum nuclease, purified according to Pilly et al. (20) was a gift of H. Donis-Keller. B. cereus pyrimidine specific ribonuclease, purified as described (21), was a gift of J. Heckman. S1 nuclease was purified from -amylase powder (Sigma) according to Rushizky et al. (22) with modifications. The DEAE cellulose chromatography was repeated after the CM cellulose step. This enzyme preparation was free of contaminating RNase T1 and T2 activities by two criteria: it generated only 3'-OH nucleotides under the partial conditions used, and after destroying the S1, no T1 or T2 nuclease activity was detectable. S1 was stored in 50 mM NaOAc pH 5.0, 0.2M NaCl, 1mM $ZnSO_4$. This preparation of S1 has no detectable loss

of activity after two years either at -20°C in storage buffer plus 50% glycerol, or at 4°C over a drop of chloroform. T4 RNA ligase was a gift of N. Pace. Calf thymus terminal transferase, EcoRI and PstI were obtained from Boehringer Mannheim. XmaI was a gift of D. Glover.

Purification of m5.8S and 2S rRNA. Total Drosophila melanogaster (strain Oregon R) RNA from adults and larvae was prepared by phenol-chloroform extraction or, alternatively, by guanidinium chloride extraction (see below). The 26S-m5.8S-2S complex was separated from small molecular weight and 18S RNAs by centrifugation on 15%-30% sucrose gradients in NET buffer (0.1M NaCl, 50 mM Tris HCl pH 7.4, 1mM EDTA) run at 70,000 x g for 24 hours. The complex was heat-denatured and run again on a 5% - 20% sucrose gradient to separate 2S and m5.8S from the two pieces of 26 RNA. The fractions of the gradient containing 2S and m5.8S were ethanol precipitated and run on a preparative 8% polyacrylamide-8.3M urea gel (6mm thick), to separate these two molecules.

Polysomal RNA was purified by salt precipitation of the polysomes, phenol or guanidinium chloride extraction, and subsequent sucrose gradient and gel fractionation as above. Alternatively, the guanidinium chloride extracted RNA was purified by preparative gel electrophoresis on an 8% gel. 5S rRNA and tRNA are not precipitated by the guanidinium chloride procedure. This expedites the purification of m5.8S RNA which co-migrates with 5S RNA on polyacrylamide gels (15). Uniformly labeled rRNA was obtained by in vivo labeling of the D. melanogaster KCO cell line and was purified as described (15).

Guanidinium chloride extraction. A modification of the published procedure was used (23; D. Sullivan, personal communication). 11 ml of 8M guanidinium chloride solution was added per gram of cells, and was vortexed at room temperature until the suspension became a homogeneous solution. The RNA was precipitated by addition of 0.5 volume of 95% ethanol and storage at -20°C for a minimum of one hour. The precipitate was recovered by centrifugation, resuspended in 4M guanidinium chloride (0.5 ml per gram starting material) and reprecipitated with ethanol as before. Resuspension in 4M guanidinium chloride (0.25 ml per gram starting material) and ethanol precipitation was repeated. The resultant pellet was extracted three times with 0.15 M NaCl and the supernatants were combined, the salt concentration was adjusted to 0.3M with 3M NaOAc, pH 5.5, and the RNA was precipitated with 3 volumes of ethanol.

5' End-labeling. RNA was dephosphorylated in 50mM Tris-HCl, pH 8.3 with 0.01 mU/pmol alkaline phosphatase at 50°C for 30 min. Following dephosphorylation the RNA was extracted once with an equal volume of phenol-chloroform (1:1 v/v). The phenol phase was back-extracted and the aqueous phases were combined and

ether-extracted 4 times. Remaining ether was removed by evaporation and the RNA was precipitated with 3 volumes of ethanol. 50 to 100 pmols of dephosphorylated RNA were 5' end-labeled with 2 to 3 times molar excess of γ-[^{32}P]-ATP (New England Nuclear, 2000 Ci/mmol) in a 10 to 20 µl reaction containing 50 mM Tris-HCl pH 8.9, 10 mM $MgCl_2$, 10mM dithiothreitol, 5% glycerol at 37°C for 30 min. An equal volume of urea-dyes-EDTA solution (9M urea, 10 mM EDTA, 0.05% Xylene cyanol, 0.05% Bromophenol blue) was added to the reaction mixture and it was loaded on an 8% polyacrlamide-8.3 M urea gel. Radioactive bands, detected by autoradiography, were excised and the RNA was eluted from the gel as described (16).

3' End-labeling. Both m5.8S and 2S have 3'-OH ends and can be ligated directly to 5' [^{32}P]-pCp using T4 RNA ligase (24). To avoid intermolecular ligation and formation of multimeric RNA molecules, we used dephosphorylated RNA. After dephosphorylation, 50 to 100 pmols of RNA were incubated with an equimolar amount of 5' [^{32}P]-pCp (New England Nuclear, 2000 Ci/mmol) in a 10 µl reaction of 50 mM HEPES pH 8.3, 10 mM $MgCl_2$, 5 mM dithiothreitol, at 4°C for 24 hours. T4 RNA ligase was 2 to 3 units per 10 µl reaction, and ATP was added to 4 times molar excess over the RNA.

S1 structure-mapping digestions. S1 digestions were performed in 5 µl reactions containing 40 mM NaOAc pH 4.5, 10 mM $ZnSO_4$, 0.2 M NaCl at enzyme to substrate ratios of 10^{-2} to 10^{-1} unit/µg RNA. Samples were preincubated for at least 10 min at 37°C, digested with S1 for 1 min to 5 min as indicated, and the reactions were stopped with 5 µl S1 stop solution (9 M urea, 10 mM EDTA, 0.5 µg/µl tRNA carrier, 0.05% XC and 0.05% BPB).

Sequencing digestions. Sequencing digestions were performed as described (16, 25). Reactions were incubated at 60°C and were stopped by freezing at -80°C. The reaction volume was 5 µl. *B. cereus* pyrimidine-specific nuclease digestions were stopped by addition of an equal volume of urea and dyes solution (21). The reactions were loaded on to thin (40 x 33 x .028 cm) 10% to 20% polyacrylamide gels and electrophoresed as described (17).

End analysis. Samples of 5'[^{32}P]-labeled RNAs were digested exhaustively with P1 nuclease. 3'[^{32}P]-labeled RNAs were digested with T2 nuclease. The products were identified by PEI-cellulose thin layer chromatography as described (16).

Complete T1 digestions. Complete T1 digestions were performed in 50 mM Tris-HCl pH 7.4, with 0.5 unit T1/µg RNA for 2 hours at 50°C. The products were separated on 20% sequencing gels and identified by autoradiography.

Intermolecular complex formation.
End-labeled 2S or m5.8S molecules were annealed to each other or to 26S rRNA in 5 µl of 0.4 M NaCl, 10 mM Tris pH 7.5, 2.5 mM EDTA (26). 20 ng of end-labeled m5.8S (specific activity 5×10^5 cpm/µg) or 5 ng of end-labeled 2S (specific activity 10^6 cpm/µg) were incubated with approximately equimolar amounts of the non-labeled species at temperatures indicated in the figures. After one hour of incubation, the reactions were chilled on ice, mixed with an equal volume of a saturated sucrose solution in the same buffer and loaded onto a 5% non-denaturing polyacrylamide gel. The gel was run at 4°C until the bromophenol blue migrated 10 cm from the origin. Bands were detected by autoradiography.

DNA preparation and transfer to nitrocellulose filters. Plasmid cDm238 constructed by dA-dT ligation of randomly sheared D. melanogaster embryo DNA to plasmid ColE1, was prepared as described (27). DNA fragments were electrophoresed on 0.5 to 1.5% horizontal agarose gels, transferred to Schleicher and Schuell BA85 nitrocellulose (28) and hybridized with the probes in 5XSSC 50% formamide buffer.

DNA labeling and sequencing. Whole plasmid DNA was digested with PstI, phenol extracted and ethanol precipitated. The fragments were labeled with α-[^{32}P]-ATP using terminal transferase in the presence of cobalt ion as described (29). After ethanol precipitation, fragments were digested with XmaI or HaeIII, separated on a preparative agarose gel and eluted by incubation of the macerated gel slice with a phenol-0.2M Tris pH 8 (1:1 v/v) solution at 60°C for 2 hours. DNA sequencing was performed as described by Maxam and Gilbert (30), and the reaction products were electrophoresed on thin (0.4 mm) gels.

Computer predictions of secondary structure. Computer-generated predictions of secondary structure were obtained using a modified version of the program developed by Pipas and McMahon (31) based on the empirical rules of Tinoco et al (32), run on an IBM 370 computer. The S1 accessibility data were used to generate models of m5.8S, as discussed (17).

RESULTS AND DISCUSSION

Purification of m5.8S and 2S rRNA

The purification of ^{32}P labeled m5.8S and 2S RNAs on polyacrylamide gels is shown in Figure 1. End-analysis of the labeled molecules was routinely performed using both P1 and T1 nucleases for the 5' end-labeled molecules and with T2 and T1 ribonucleases for the 3' end-labeled molecules, followed by either PEI chromatography (16) or gel electrophoresis to separate the products. The 5' terminal nucleotide of m5.8S rRNA is found to be 100% A. However, complete T1 digestion gives two products differing in length by one nucleotide. This suggests that a minor species exists with an additional adenine residue at the 5' end. This may reflect either a property of the processing enzyme(s) or the action of a non-specific exonuclease. There are no minor species detected upon end-analysis of 3' end-labeled m5.8S rRNA. The difference observed at the two ends of this molecule may reflect differential accuracy in cleavage at these two processing sites and supports the possibility that two enzymes are involved. Several other 5.8S RNAs appear to be heterogeneous at the 5' and/or the 3' end (1). The 2S rRNA is found to be equally homogeneous at both ends by end-analysis.

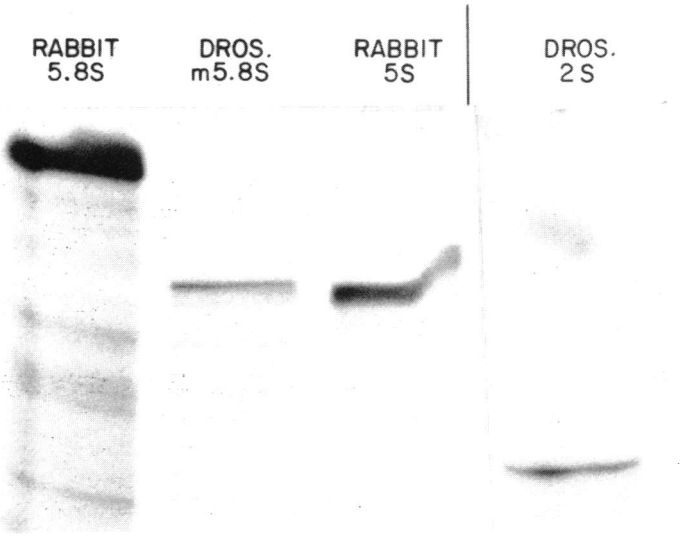

Figure 1. Preparative gel purification of 5' end-labeled D. melanogaster m5.8S and 2S rRNAs. The m5.8S and 2S were run on 8% and 20% polyacrylamide - 8.3M urea gels, respectively. Autoradiographs were obtained by exposing films for 10 sec. Rabbit 5.8S and 5S rRNAs are 5' labeled and run in parallel with m5.8S as markers. The m5.8S molecule is slightly longer than the 120 nucleotide rabbit 5S rRNA.

Sequence of m5.8S rRNA

The m5.8S molecule was completely sequenced from both ends using 5' and 3' end-labeled RNA and the rapid RNA sequencing method (25, 33) with thin polyacrylamide gels (see Materials and Methods). Examples of sequencing gels are shown in Figures 2 and 3. The major species is 123 nucleotides long and is found to be homologous to the 5' part of other sequenced 5.8S molecules (1). The extent of homology with the 5' part of other 5.8S RNA molecules is between 66% and 77% (see Table 1 and Figure 10). Light bands one nucleotide longer than the major species are seen on the sequencing gels of the 5' end-labeled RNA (Figure 2) due to the presence of a small percentage of molecules with one additional adenine residue at the 5' end.

Figure 3 shows an example of a sequencing gel with 3' end-labeled m5.8S RNA. The double bands in the second T1 lane (T(2)) are caused by molecules that have been cleaved twice by T1, once at a random G and again at the penultimate nucleotide which is a G and to which the ^{32}pCp molecule has been ligated. Since T1 generates 3' phosphate oligonucleotides, the radioactive phosphate will remain with the terminal G, producing a second ladder of labeled T1 partial digestion products differing by one nucleotide from the ones that retain the ^{32}pCp molecule ligated to the m5.8S molecule.

The different specificites of pancreatic ribonuclease and B. cereus pyrimidine-specific ribonuclease are also demonstrated in Figure 3. Pancreatic ribonuclease under the conditions used for partial digestion cleaves strongly at Py-A and less strongly at Py-G bonds. Py-Py bonds are cleaved weakly or not at all. B. cereus ribonuclease cleaves at all pyrimidines with the exception of runs of C's (21).

Nucleotide 74 is 2'-O-ribose methylated, resulting in a gap on the gels and cannot be identified by the gel sequencing method. Levis and Penman (18) have shown that the m5.8S RNA of D. melanogaster contains only one methylated sequence, GmpC. Furthermore, all the sequenced 5.8S RNAs, except for the yeast species (1), have a 2'-O-methylated guanosine at the same position in the molecule. We conclude that the methylated nucleotide is a G. This agrees with the DNA sequence (see below). The sequencing method used here cannot distinguish between uridine, pseudouridine and thymidine residues, therefore, other modified bases known to exist in other 5.8S RNAs, may be present.

Structure-mapping of m5.8S rRNA

S1 nuclease cleaves the m5.8S rRNA primarily in two regions; A50-52 and U99, A101, A102 (Figure 2). Other, less susceptible nucleotides also exist

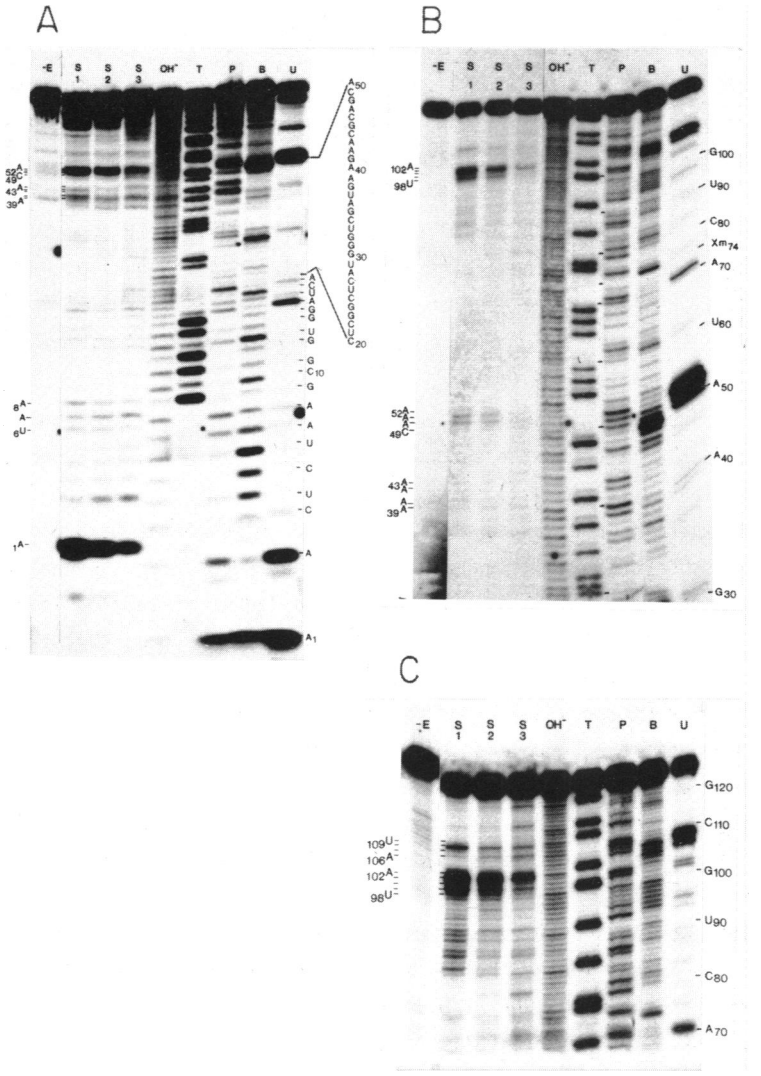

Figure 2. Electrophoretic patterns of partial enzymatic digestions for sequencing and structural analysis of 5' end-labeled m5.8S rRNA on 20% (Panel A),and 10% (panels B and C) polyacrylamide-8.3M urea gels. An identical aliquot from the following reactions was loaded in the same order on each gel, from left to right: -E, minus enzyme control; S(1), S1 nuclease, 0.1 U/µg RNA, 5 min ; S(2), S1 nuclease, 0.01 U/µg RNA, 5 min ; S(3), S1 nuclease, 0.01 U/µg RNA, 1 min ; OH^-, alkaline digestion, 3 min at 90°C; T, T1 nuclease, 5 x 10^{-4} U/µg RNA, 3 min; P, Physarum nuclease, 5 min; B, Bacillus cereus nuclease, 10 min; U, U2 nuclease, 10^{-2} U/µg RNA, 10 min. 20,000 Cerenkov cpm of the 5'-^{32}P-m5.8S rRNA was loaded per slot. The positions of cleavage by S1 are indicated along the left side and the nucleotide sequence is shown along the right side of each panel.

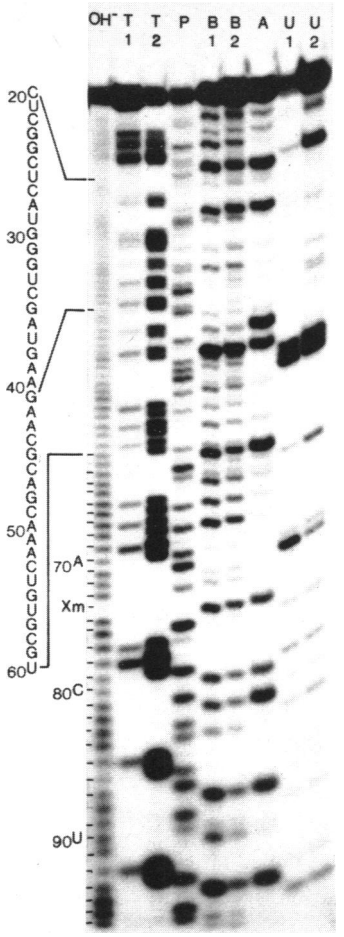

Figure 3. Electrophoretic patterns of partial enzymatic digestions of 3' end-labeled m5.8S rRNA. 20,000 Cerenkov cpm were loaded per slot. Lanes from left to right are; OH^-, alkaline digestion, 3 min 90°C; T(1), T1 nuclease, 10^{-4}U/µg RNA, 3 min; P, Physarum nuclease, 10 min; B(1) and B(2), Bacillus cereus nuclease, 10 min and 20 min, respectively; A, pancreatic ribonuclease, 10^{-4} U/µg RNA 10 min; U(1), U2 ribonuclease, 10^{-2} U/µg RNA, 10 min; U(2), U2 ribonuclease, 10^{-3} U/µg RNA, 10 min.

and are seen in Figures 2 and 6. The mobilities of the S1 partial digestion products are different than those of the products of the sequencing digestions with all the other enzymes and care must be taken to insure the correct assignment of the bands. With 5' end-labeled molecules the S1 bands run slower than the corresponding sequencing bands which have an additional phosphate at the 3' end. For example, in Figure 2A it is seen that (above the first six oligonucleotide products) the S1 bands are always between 1 to 1-1/2 bands slower. With 3' end-labeled molecules, the mobilities are reversed, because the labeled S1 products have more negative charge. For example, in Figure 4B the S1 partial digestion products of the 2S molecule run approximately

Table 1

Percentage of identities among 5.8S molecules

	Hela	X.laevis	Trout	Yeast	N.crassa
Drosophila m5.8S	76.9	76.9	76.9	66.9	66.2
Drosophila 2S	58.1	58.1	61.3	51.6	45.2
Drosophila m5.8S+2S	73.3	73.3	73.9	64.0	62.1
Hela	-	97.0	95.1	72.6	68.9
Hela; 5' 132 positions	-	98.5	95.5	71.2	69.7
Hela; 3' 32 position	-	90.6	93.8	78.1	65.6
Drosophila; positions 118-143	58.3	58.3	54.2	45.8	29.2
Hela; positions 118-143	-	96.1	88.5	65.4	34.6

Comparisons are made of equivalent positions according to the alignment shown in Figure 10.

1/2 to 1-1/2 bands faster than the corresponding sequencing reaction products. Double bands are obtained with S1 nuclease digestion of 3' labeled molecules due to a 3' phosphatase activity that is present in the S1 preparation and can convert ^{32}pCp to $^{32}pC_{OH}$. This activity is probably a property of the S1 nuclease molecule (data not shown). The 3'phosphatase activity can be used as a rapid assay for S1 nuclease during the purification procedure. It is also an internal control in the S1 reaction that provides an estimate of the percentage of the molecules cleaved twice. We do not have a direct estimate of the different rates of cleavage at the phosphodiester bonds (endonucleotytic activity of S1) versus the cleavage of the 3' terminal phosphate (3' phosphatase activity), but the latter appears to be very fast. We can therefore assume that whenever the double bands are not prominent the majority of the cleaved molecules have been hit only once by S1 nuclease. The double bands can be eliminated by alkaline phosphatase treatment of the 3' end-labeled RNA prior to S1 nuclease digestion.

A computer model of the secondary structure of the m5.8S molecule, generated using the S1 susceptibility data (17), is shown in Figure 6. The model has many similarities to that proposed for the rat hepatoma 5.8S rRNA

(34). However, the GC rich hairpin which is the most stable structural characteristic of all the sequenced 5.8S molecules (1-3) is missing in m5.8S rRNA. The 3'part of the molecule, which in mouse 5.8S is known to form a stable complex with 28S rRNA (26), is also missing. We present evidence that the missing part is the 2S RNA molecule. An alternative base-paired structure for the region A43-U60 but with less negative free energy can be shown between nucleotides C44-G48 and C53-G57 leaving nucleotides C49-A52 in a hairpin loop. This alternative structure agrees with the S1 structure mapping data and aligns better with the position of the homologous hairpin in the other 5.8S molecules (see Figure 10).

There are some published reports that are not in agreement with this analysis. The cloverleaf model proposed recently (35) is considerably different from the model presented here for Drosophila and for rabbit 5.8S (G. Pavlakis, unpublished) and is in disagreement with the S1 structure-mapping data.

Sequence of 2S rRNA

Gels displaying the sequence of Drosophila 2S rRNA, obtained with 5' and 3' end-labeled RNA, are shown in Figure 4. The primary sequences of both 2S and m5.8S RNA obtained by rapid RNA sequencing and presented here are in complete agreement with the DNA sequence. This sequence corrects the previously published one (15) which was in error due primarily to the misinterpretation of pancreatic oligonucleotide P9. The 2S RNA sequence is between 45% and 61% homologous to the 3' part of other 5.8S molecules (see Figure 10 and Table 1).

Structure mapping of 2S rRNA

Jordan et al. (15) attributed the failure to obtain partial digests of the 2S molecule to the lack of secondary structure. Computer predictions also suggest that the 2S RNA lacks stable secondary structure in solution. The most stable interaction is between nucleotides 10-12 and 28-30 and the most stable intermolecular interaction among 2S molecules is within the complementary palindromic sequence 5'-CAUAUG-3' (nucleotides 12-17). Since these interactions are unstable, the 2S is expected to be a single-stranded monomer in solution. Structure mapping of the 2S rRNA verifies the conclusion drawn from computer predictions. Figure 4 shows that S1 cleaves all the phosphodiester bonds in the 2S molecule. The lack of stable secondary structure enabled us to use 2S RNA to study the specificity of S1 nuclease under the partial conditions that are used in the structure-mapping experiments (16, 17). S1 nuclease appears to have preference for A and U over C and G. This pattern remains the same

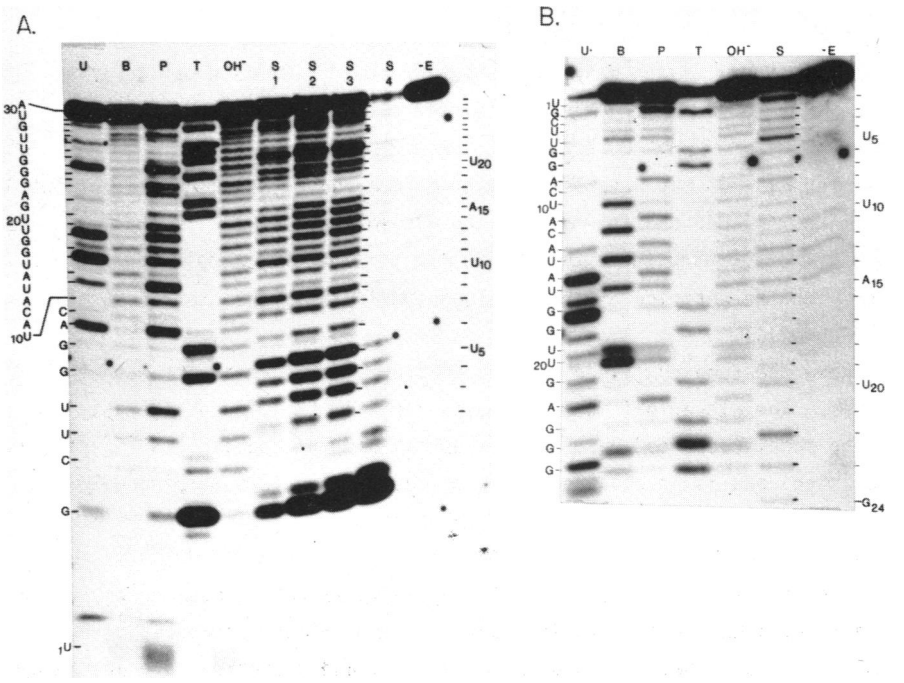

Figure 4. Electrophoretic patterns of partial enzymatic digestions of D. melanogaster 2S rRNA.

A. 5' end-labeled 2S rRNA on a 20% polyacrylamide-8.3M urea gel. Lanes from left to right are: U, U2 ribonuclease, 10^{-2} U/µg RNA, 10 min; B, B. cereus nuclease, 10 min; P, Physarum nuclease, 5 min; T, T1 ribonuclease, 5 x 10^{-4} U/µg RNA, 3 min; OH^-, alkaline digestion, 3 min, 90°C; S(1), S1 nuclease, 0.01 U/µg RNA, 1 min; S(2), S1 nuclease, 0.01 U/µg RNA, 5 min; S(3), S1 nuclease, 0.1 U/µg RNA, 1 min; S(4), S1 nuclease, 0.1 U/µg RNA, 5 min; -E, minus enzyme control.

B. 3' end-labeled 2S rRNA on a 20% polyacrylamide -8.3M urea gel. Lanes from left to right are: U, U2 ribonuclease, 10^{-2} U/µg RNA, 10 min; B, B. cereus nuclease, 10 min; P, Physarum nuclease, 5 min; T, T1 ribonuclease, 5 x 10^{-4} U/µg RNA 3 min; OH^-, alkaline digestion, 3 min, 90°C; S, S1 nuclease, 0.01 U/µg RNA, 3 min; -E, minus enzyme control.

Approximately 20,000 Cerenkov cpm of ^{32}P-end-labeled 2S rRNA were loaded on each slot. Electrophoresis was at 1500-2000 volts until the BPB dye was 21 cm from the origin. The nucleotide sequence is indicated along the left margin and the position of cleavage of S1 nuclease along the right margin of each panel.

over a broad range of conditions (G. Pavlakis, unpublished), including temperature up to 70°C.

Complex formation.

The data in Figure 5A demonstrate that m5.8S and 2S rRNA can hybridize to form a complex with a melting temperature greater than 50°C. This complex migrates on gels similarly to end-labeled rabbit 5.8S RNA. Figure 5B shows

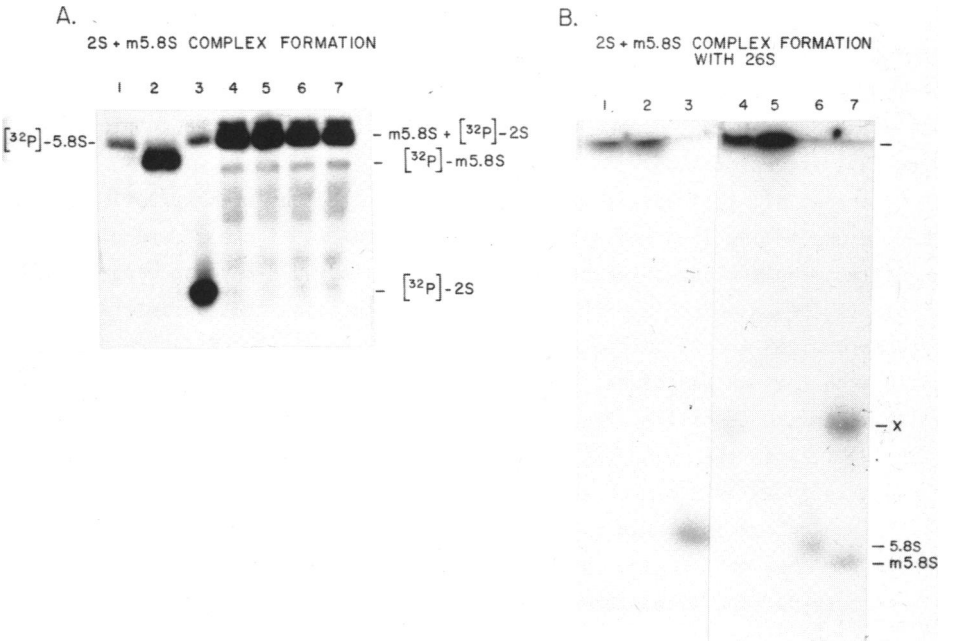

Figure 5. Complex formation of D. melanogaster 2S and m5.8S rRNAs with one another and with 26S rRNA.

A. Autoradiograph following electrophoresis of 2S-m5.8S rRNA complexes on a 5% non-denaturing polyacrylamide gel. Lanes from left to right are: 1. 5'[^{32}P]-rabbit 5.8S rRNA; 2. 5'[^{32}P]- m5.8S rRNA; 3-7.5'[^{32}P]-2S rRNA complexed to m5.8S rRNA at 0°C, 37°C, 40°C, 45°C and 50°C, respectively. Samples were pre-heated for 3 min at 80°C prior to incubation for one hour at the stated temperatures.

B. Autoradiograph following electrophoresis of 2S-m5.8S-26S rRNA complexes on a 5% non-denaturing polyacrylamide gel. Lanes from left to right are: 1. 5'[^{32}P]-2S with m5.8S and 26S; 2. 5'[^{32}P]-m5.8S rRNA with 2S and 26S; 3. 5'[^{32}P]-2S with m5.8S; 4. 5'[^{32}P]-m5.8S with 26S; 5. 5'[^{32}P]-2S with 26S; 6. 5'[^{32}P]-rabbit 5.8S; 7. 5'[^{32}P]-m5.8S. Samples were pre-heated for 2 min at 80°C prior to incubation at 37°C for one hour. The symbol X indicates the position of migration of a species that is probably a m5.8S dimer.

that m5.8S and 2S rRNA can anneal independently and together to the 26S rRNA molecule of Drosophila. These results indicate that m5.8S and 2S have a region of complementarity and each has another region complementary to 26S rRNA. The most stable interaction, as determined by computer analysis, between m5.8S and 2S is shown in Figure 11A and includes nucleotides at the 3' end of m5.8S and at the 5' end of the 2S RNA. This complex is structurally equivalent to the other sequenced 5.8S RNA molecules, all of which have a GC rich hairpin helix at positions homologous to the postulated 5.8S-2S intermolecular interaction (1-3). This arrangement strongly suggests that the 2S molecule is the missing 3' part of the 5.8S molecule and that an additional processing event takes place in Drosophila in order to generate the 2S molecule. Several other facts presented below in addition to sequence homology and the formation of the complex agree with this interpretation.

Pace et al. (26) have shown that one of the regions of interaction between the 160 nucleotide long mouse 5.8S RNA and 28S RNA is at the 3' end of the 5.8S molecule. The analogous interaction in Drosophila would have to occur at the 3' end of the 2S molecule. Figure 5B lane 5, indicates that an interaction between the 26S and 2S molecules is possible. Pace et al. (26) have also postulated that an additional interaction between intact 5.8S and 28S rRNAs must exist since the stability of the complex of 28S with the 3' portion of mouse 5.8S RNA is less than with the intact molecule. Figure 5B shows that the Drosophila m5.8S rRNA can anneal independently to the 26S molecule demonstrating that a region exists within the m5.8S RNA that is complementary to one in 26S. We propose that this region is the 5' part of the m5.8S RNA which does not form stable base-pairing within the m5.8S molecule.

Restriction map of plasmid cDm238 and localization of the rRNA coding regions.

Digestion of plasmid cDm238 DNA with EcoRI gives two fragments, 12Kb and 8 Kb long respectively. The 12 Kb fragment hybridizes strongly with labeled rRNA but does not hybridize with either type 1 or type 2 insertion DNA (36; data not shown). This suggests that cDm238 contains a complete 12 Kb repeating unit of rDNA, and that the 12Kb EcoRI fragment is generated by cleavage at the EcoRI site known to exist close to the 3' end of the 18S coding sequence (37). XmaI has been previously shown to cleave ColEI only once (38). The 18S and 26S coding sequences in D. melanogaster rDNA each contain a single XmaI site (39). Figure 7 (lane 4) shows that four fragments are generated upon digestion of cDm238 with XmaI. Southern transfer (28) and hybridization with labeled rRNA indicates that all the XmaI fragments hybridize with 18S rRNA (lane 6), while only two of them hybridize with 26S rRNA (lane 2). This

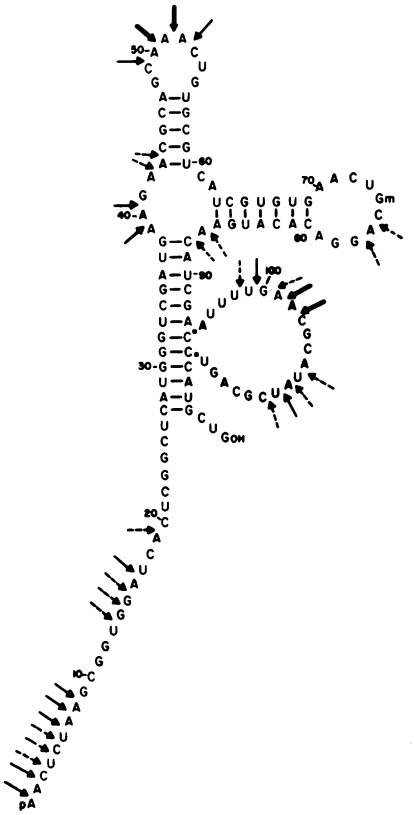

Figure 6. Secondary structure model and S1 nuclease accessibility pattern of Drosophila m5.8S rRNA. The solid arrows indicate positions of strong and the dashed arrows of weak S1 cleavages, respectively.

verifies that plasmid cDm 238 contains a complete 12Kb repeating unit, that the 18S coding regions are represented twice in the cloned fragment up to the XmaI site, and that the relative positions of the regions are: vector-18S-26S-18S-vector. Based on its size and hybridization characteristics, fragment XmaC must contain part of the 18S coding region, the entire "internal transcribed spacer", and part of the 26S coding region. It has been previously shown that the coding regions for the m5.8S and 2S rRNAs are located within the internal transcribed spacer (40). In agreement with this, m5.8S and 2S hybridize with the XmaC fragment (lanes 10 and 14).

Double digestion with nucleases XmaI and PstI followed by blotting and hybridization allow the determination of the orientation of the m5.8S and 2S coding regions. PstI cleaves fragment XmaC once, producing fragments C1 and C2 (figure 7, lane 3). m5.8S hybridizes with both bands (lane 9). Therefore, the PstI site is within the m5.8S coding region, in agreement with the RNA

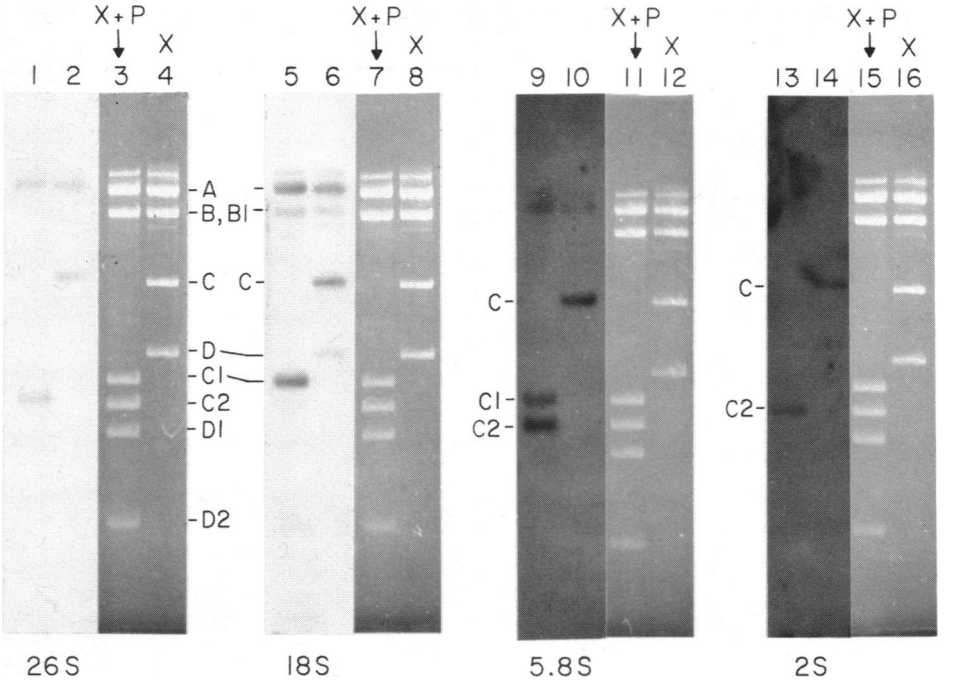

Figure 7. Hybridization of labeled rRNA to cDm238 restriction fragments. The two left lanes of each block are the autoradiograms of the nitrocellulose filters carrying the transferred DNA from the gels shown to the right stained with Ethidium Bromide. Lanes 4, 8, 12 and 16 are XmaI digestions and 3, 7, 11 and 15 are XmaI + PstI digestions. The radioactive rRNA probe is indicated at the bottom of each block.

and DNA sequence data. 2S hybridizes only with fragment C2 (lane 13). 18S hybridizes only with band C1 (lane 5) and 28S only with C2 (lane 3). Thus, the order of the rDNA coding regions is 18S-m5.8S-2S-26S. Knowledge of the position of PstI sites in ColEI (38) and size determination of the XmaI and XmaI + PstI restriction fragments, allows the orientation of the rDNA insert with respect to the vector. Figure 8 shows a restriction map of plasmid cDm238 as well as the orientation of the rDNA coding regions.

Sequence of the m5.8S-2S rDNA region.

The sequence of the m5.8S-2S rDNA region was determined starting from the PstI site that exists within the m5.8S coding region. The standard Maxam and Gilbert sequencing procedure (30) was used after labelling the DNA with

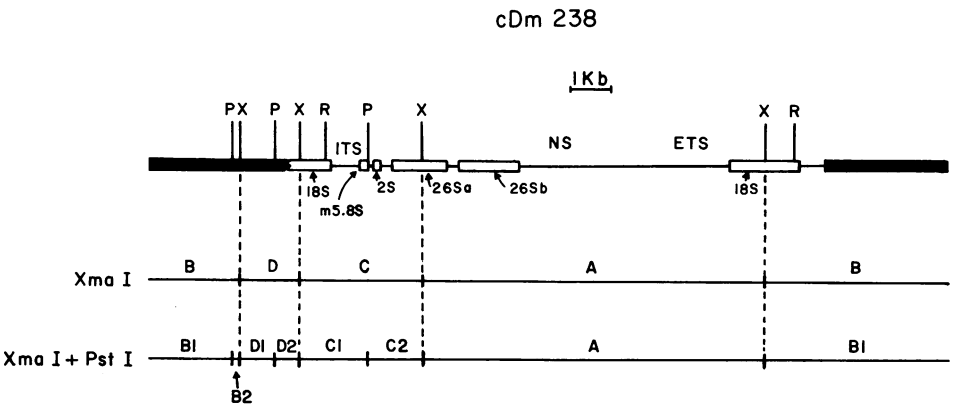

Figure 8. Restriction map of plasmid cDm238. Ribosomal RNA coding sequences are boxed. Vector ColE1 sequences are indicated by a thick line and spacer regions by a thin line. The restriction fragments obtained with XmaI and XmaI + PstI are indicated at the bottom. P = PstI, R = EcoRI, X = XmaI, ITS = internal transcribed spacer, NS = non-transcribed spacer, ETS = external transcribed spacer.

calf thymus terminal transferase and α-[^{32}P]-ATP. The regions coding for the rRNAs are recognized by using the m5.8S and 2S RNA sequences. The DNA sequence thus derived is shown in Figure 9. It includes the complete m5.8S and 2S coding regions and extends into the spacer. The DNA sequence is in complete agreement with the RNA sequence.

The most interesting feature of this DNA sequence is the 28 nucleotide long spacer between the m5.8S and 2S. This spacer is extremely AT rich (82%), and it is likely that it is transcribed and subsequently removed during ribosomal RNA maturation, since a 34S precursor to Drosophila 18S, 26S and 5.8S rRNAs has been demonstrated (18).

These results strengthen the conclusion already drawn that the 2S molecule is the 3' part of the m5.8S and that *in situ* they form a "5.8S complex" similar in structure to other 5.8S rRNAs.

However, the difference between Drosophila and other eucaryotes is not simply an additional cleavage separating m5.8S and 2S, but a whole AT rich region which is removed during processing.

Another interesting feature of the DNA sequence is the extremely AT rich spacers immediately before and after the m5.8S and 2S coding sequences (see Figure 9). These sequences are limited, and a "normal" DNA sequence is

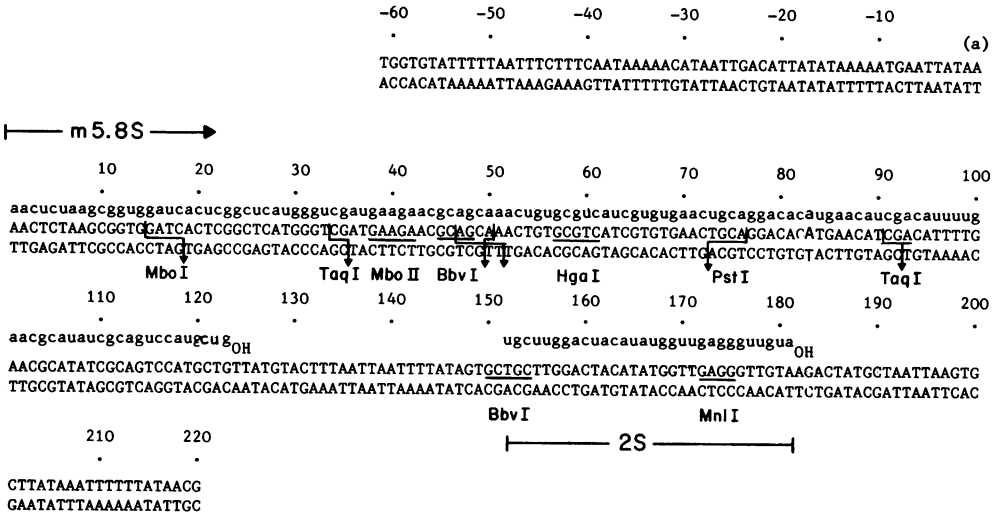

Figure 9. Nucleotide sequence of the m5.8S-2S rDNA region of D. melanogaster. The DNA sequence is in upper case and the m5.8S and 2S rRNA sequences in lower case letters. The computer predicted sites for all known restriction nucleases are indicated. The A-T content is (from 5' to 3'): spacer 85%, m5.8 50%, spacer 82%, 2S 57%, spacer 76%.

encountered within the spacer region as evidenced by the reappearance of bands in the G and C lanes of the sequencing gels (data not shown).

Comparisons of the DNA sequences.

While the coding regions of the DNA are homologous to other species (see Figure 10), no homology can be detected within the sequenced portions of the spacers between Drosophila, N. crassa (3) and X. laevis (41). The A-T content of the spacers also varies. In X. laevis the spacers are G-C rich and it was proposed that they may base-pair before processing (41). We favor the hypothesis that the 5.8S has similar structure in the precursor and that at least part of the processing signal lies in the sequence and/or structure of the mature molecule.

Sequence and structure homologies with other 5.8S rRNAs

Sequence and structure homologies between Drosophila m5.8S and 2S rRNAs and 5.8S rRNAs from other species are shown in Figure 10 and in Table 1. Comparisons are made by aligning the m5.8S rRNA sequence with the first 130 positions and

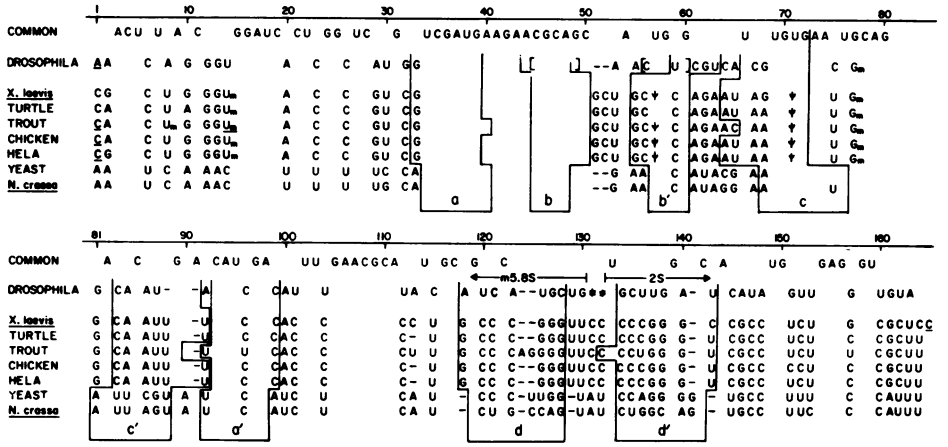

Figure 10 Comparison of nucleotide sequences of Drosophila melanogaster m5.8S and 2S with other 5.8S molecules. The 2S molecule is aligned with the 3' part of the other 5.8S rRNAs. The upper line shows the nucleotides common in all 5.8S RNAs. The four regions of secondary structure designated aa', bb', cc', and dd' are boxed. Brackets indicate the alternative bb' region in Drosophila. Underlined nucleotides are present at less than one mole per mole 5.8S RNA.

the 2S with the last 31 positions, respectively, from the 5' end of the other 5.8S molecules in Figure 10. Five vertebrate and two fungal sequences are shown and compared to the Drosophila sequence. Alignment was arranged to give the maximum number of identities. The nucleotides common to all eight species are indicated along the top of the figure. The sequence of X. laevis 5.8S rDNA (41) differs by a dinucleotide (positions 51-52 in figure 10) from the RNA sequence of Khan and Maden (42) which is based on homology of RNA fingerprint patterns to those of Nazar et al. (34) for rat hepatoma 5.8S rRNA. If this dinucleotide is not present, as suggested by the DNA sequence, then the Drosophila, yeast, Neurospora and Tetrahymena (G. Pavlakis, unpublished) sequences align better with the others in Figure 10. The comparisons in Figure 10 and Table 1 indicate that the 3' part of the 5.8S molecule, which is the 2S rRNA in Drosophila, diverges slightly faster than the 5' part during evolution. Homology of the Drosophila sequence to the others is between 66%-77% in the 5' 130 positions, and only 45%-61% in the 3' 31 positions, as seen in Table 1. This appears to be a general rule with the exception of yeast. Comparison of the Hela and yeast sequences shows more homology at the 3' end. However, if one compares

positions 118-143 (Figure 10, Table 1), a region which can be shown to form the very stable "GC-rich" helix in all 5.8S molecules (see below), it is seen that this region diverges faster than the rest of the 5.8 molecule. It appears, therefore, that regions exist within the 5.8S sequence that are less evolutionarily conserved than others.

There are four regions of potential secondary structure that can be shown to exist in all sequenced 5.8S rRNAs at equivalent positions in their sequence. They are labeled aa', bb', cc', and dd' respectively, in Figures 10 and 11A. The four double stranded regions are incorporated into the secondary structure model of Nazar et al. (34) for the rat hepatoma 5.8S molecule. The dd' double-stranded region in Drosophila is formed by a junction complex between the 3' end of m5.8S and the 5' end of 2S RNA. S1 structure mapping (Figure 2) and complex formation (Figure 5A) data presented here agree with the existence of the four helical regions. Yeast 5.8S does not have a 2'-O-methylated G at position 77 and the hairpin loop included at the end of the cc' helix has only four nucleotides in the proposed model of Nazar, Sitz and Busch (34), rather than

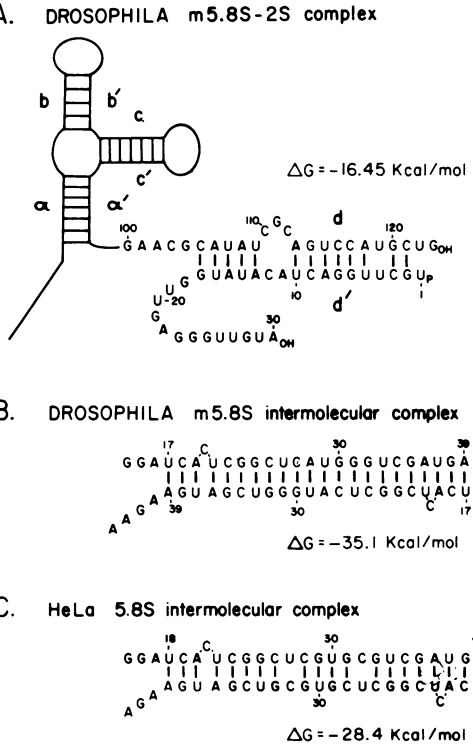

Figure 11 A. Intermolecular complex between D. melanogaster m5.8S and 2S rRNA molecules. The most stable interaction is shown. The four double stranded regions conserved in all the 5.8S rRNA molecules are designated aa', bb', cc', dd' as in Figure 7.
B. A possible intermolecular interaction between two m5.8S molecules that can explain the formation of band X in Figure 5B.
C. A possible intermolecular interaction between two mammalian (Hela) 5.8S rRNA molecules similar to the one for Drosophila. This complex could explain the results of Pace et al. (26).

10 nucleotides as in all other 5.8S rRNAs. Since 5.8S rRNA is found in the ribosome as a complex with 28S rRNA, through its 3' and possibly 5' ends, the secondary structure of these portions of the 5.8S molecule will be different when studied as a free molecule in solution than when studied complexed to 28S rRNA.

As mentioned above, the region 118-143 diverges more rapidly during evolution than the rest of the 5.8S RNA. This region can always be folded into a very strong hairpin which is the most stable structural characteristic of all 5.8S molecules and is referred to as region dd' in Figures 10 and 11A. This hairpin has, in the case of yeast, an experimentally estimated free energy of -15.3 Kcal/mol (43) which agrees well with a computer estimate of -16.2 Kcal/mol. We propose that the sequence per se in this part of the molecule is not of importance but that the secondary structure is critical for the biological function of the 5.8S molecule. This extent of divergence is not found in other double-stranded regions within the molecule. For example, helical region aa' is highly conserved in sequence for all 5.8S molecules. In this case the primary sequence may be important, as well as the secondary structure, for the function of the molecule. The regions of the molecule exposed to S1 show also a variable degree of sequence conservation: there are some exposed nucleotides that are not conserved, and others that are, e.g. the sequence -AAGAA- at positions 40-45 (see Figures 6 and 10).

Intermolecular Complexes

A band designated X can be seen in Figure 5B (lanes 4 and 7) that migrates slower than the m5.8S. This band appears only when m5.8S is heat-denatured and then annealed. Pace et al. (26) observed the same effect with mouse 5.8S rRNA. They proposed that this band is either a dimer caused by inter-molecular interactions or a different structure with a very much more expanded radius of gyration. A computer search for intermolecular interactions supports the first hypothesis. A very stable interaction 22 base-pairs long can be shown between two m5.8S molecules at nucleotides 17-39 which form a long palindrome (Figure 11B). Gibb's free energy of formation calculated for this structure indicates that it is more stable than the sum of the intramolecular interactions for the best structure predicted by the computer. Therefore, re-annealing at high temperature would favor the dimer. There are only two differences between the Drosophila and the mammalian (Hela) sequences in this region (Figure 10). Therefore, a similar intermolecular interaction is possible among mammallian

5.8S molecules as is illustrated in Figure 11C.

As shown in Figure 9, the PstI site within the m5.8S rDNA sequence occurs at nucleotides 72-77 leaving a longer part of the coding sequence with the fragment C1 than the C2. However, the autoradiogram shown in Figure 7 lane 9 indicates a stronger hybridization signal for the C2 fragment than the C1. The C2 fragment contains the whole 2S rRNA and part of the 26S rRNA coding region. We reason therefore, that under the condition of the hybridization experiment (no nuclease digestion of the DNA-RNA hybrids), m5.8S RNA base-pairs with complementary sequences present on the 2S and 26S rDNA. This prompted us to try to locate the sites of these interactions more precisely (work in progress). The influence of the m5.8S dimer on this result is uncertain.

Processing of Drosophila 5.8S rRNA

The m5.8S and 2S rRNA nucleotide sequences and Drosophila rDNA sequence presented here, suggest an unusual RNA processing mechanism that is found only in Drosophila at the present time. The most probable secondary structure (obtained by computer modeling) of the presumptive 5.8S RNA precursor molecule is shown in Figure 12. The spacer sequence is partially double-stranded and forms a hairpin loop. The nucleotides in the predicted loop form a palindrome -UUUAAUUAAUUU-. The helix at which the processing enzymatic cleavage takes place includes the direct repeat -UGCUG- at positions 119-123 and 149-153. The role, if any, of the loop, helix and direct repeat in the processing step is unknown. However, it is suggested that the secondary structure due to the presence of these elements is important for the processing at this site. This model, which has the cleavage site in a double-stranded region, is similar to the "looped-out" model proposed by others (44, 45) for processing of primary transcripts of messenger RNA.

There may be similarities between the 5.8S processing site and one known to exist in the middle of the 26 rRNA precursor molecule in Drosophila. The Drosophila 26S rRNA is also split into two pieces as it is in several other species (46), and 100-150 nucleotides are specifically removed from the rRNA precursor, according to electron microscopic evidence (9, 47).

There is no 153 nucleotide 5.8S species detectable in Drosophila, therefore, no ligation occurs following the excision step during the removal of the transcribed spacer. The possible 181 nucleotide long precursor has not been detected as yet. Absence of an intact 160 nucleotide long 5.8S molecule from the ribosomes of D. melanogaster suggests (but does not prove) that all the rDNA repeating units have a spacer at that position. Further studies are needed to elucidate the sequence of the processing events. There are no detectable

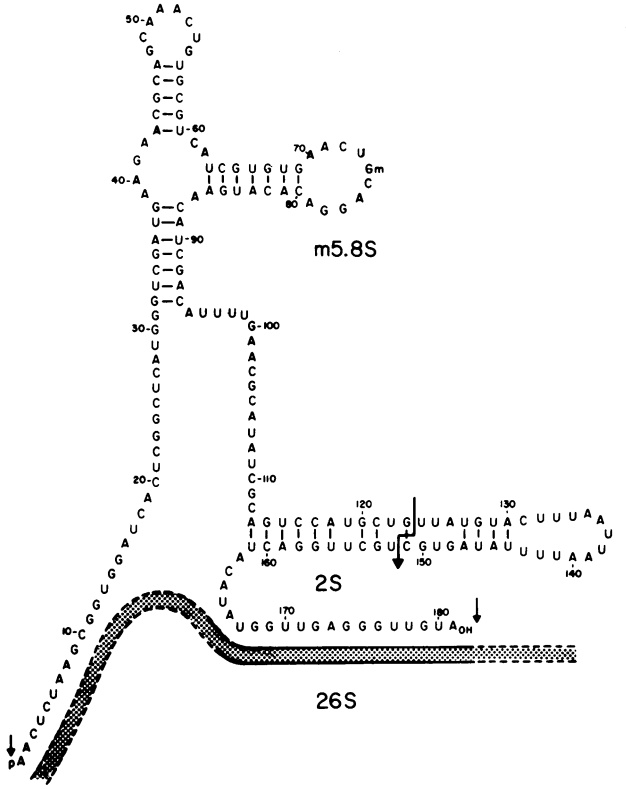

Figure 12. A computer generated secondary structure model for the precursor of the m5.8S and 2S RNAs. The 5' end of the m5.8S and the 3' end of the 2S are shown to interact with the 26S. The 3 processing sites are indicated by arrows. Processing between m5.8S and 2S specifically removes the 28 nucleotide long spacer.

differences between the RNA from adults and larvae.

No spacer analogous to the 28 nucleotide sequence in Drosophila has been found in either the X. laevis or N. crassa (41, 3) sequenced rDNA genes coding for 5.8S rRNA. We hypothesize, therefore, that the spacer sequence was inserted into the 5.8S rDNA gene of Drosophila at a time following the divergence of the vertebrates and invertebrates. A pre-existing enzymatic mechanism was able to excise the spacer, producing 5.8S molecules similar in structure to those of other species, but missing an apparently non-functional loop. Alternatively, the spacer may have existed in ancestral sequences and was eliminated during evolution from those species studied so far, but not from Drosophila. Therefore, both the spacer and the enzymatic mechanism for its removal were maintained in Drosophila. Further studies of invertebrate ribosomal RNA genes are needed to distinguish between these alternatives. A simple first approach for screening many different species is to detect the size of their 5.8S rRNA.

Making the assumption that an intact 160 nucleotide long 5.8S molecule reflects a continuous 5.8S coding region on the rDNA, the observation that some invertebrates have intact 5.8S molecules (G.N.P., A. Trout and D. Sullivan, unpublished) could mean that the insertion of the spacer is relatively recent in evolution.

Since other species may have a 5.8S molecule in two pieces (11), we propose that these two pieces be named 5.8Sa for the 5' part (which we name m5.8S in this paper) and 5.8Sb for the 3' part (which is the 2S molecule), in accordance with the terminology used for the two pieces of the 26S rRNA molecule.

A preliminary account of this work has been presented at the XIth Miami Winter Symposium.

Abbreviations used: Py, pyrimidine; BPB, bromophenol blue; XC, xylene cyanol.

ACKNOWLEDGEMENTS

We thank Drs. D. Sullivan, G. G. Brownlee and D. Glover for materials and other help, and Lauren Reiser and R. Jourdan for technical assistance. B.R.J. thanks D. Glover for hospitality while on sabbatical and acknowledges MRC grants G975/660/C and G97B/455/C to D. Glover. The work was supported by grants from the N.I.H. (GM22280), the Alton-Jones Foundation and the Syracuse University Senate Research Fund to J.N.V., and a long-term EMBO fellowship to B.R.J.

NOTE ADDED IN PROOF.

After this manuscript was submitted we learned that E.O. Long and I.B. Dawid (submitted for publication) have detected a possible nuclear precursor to m5.8S and 2S which is approximately 500 bases long, but no 123 bases long species were detected, suggesting that the generation of the m5.8S and 2S species is a late processing event occurring in the cytoplasm (ref. 15). Also, a T should be inserted between positions 100-200 in the nucleotide sequence shown in Figure 9.

REFERENCES

1. Erdmann, V. A. (1979) Nucl. Acids Res. 6, r29-r44.
2. Nazar, R. and Roy, K. L. (1978) J. Biol. Chem. 253, 395-399.
3. Selker, E. and Yanofsky, C. (1979) Nucl. Acids Res. 6, 2561-2567.
4. Spiers, J. and Birnstiel, M. L. (1974) J. Mol. Biol. 87, 237-256.
5. Walker, T. A. and Pace, N. R. (1977) Nucl. Acids Res. 4, 595-609.
6. Cory, S. and Adams, J. M. (1977) Cell 11, 795-805.
7. Frankel, G., Cockburn, A. F., Kindle, K. L. and Firtel, R. A. (1977) J. Mol. Biol. 109, 539-558.
8. Bell, G. I., DeGennaro, L. J., Gelfand, D. H., Bishop, R. J., and Valenzuela, P. (1977) J. Biol. Chem. 252, 8118-8125.

9. Pellegrini, M., Manning, J. and Davidson, N. (1977) Cell 10, 213-224.
10. Free, S. J., Rice, P. W. and Metzenberg, R. L. (1979) J. Bact. 137, 1219-1226.
11. Renkawitz, R. and Gerbi, S. A. (1979) Molec. gen. Genet. 173, 1-13.
12. Perry, R. P. (1976) Ann. Rev. Bioch. 45, 605-629.
13. Fedoroff, N. V. (1979) Cell 16, 697-710.
14. Jordan, B. R. (1974) FEBS Lett. 44, 39-42.
15. Jordan, B. R., Jourdan, R. and Jacq, B. (1976) J. Mol. Biol. 101, 85-105.
16. Wurst, R., Vournakis, J. N. and Maxam, A. (1978) Biochem. 17, 4493-4499.
17. Pavlakis, G. N., Lockard, R. E., Vamvakopoulos, N. C., Rieser, L. A., RajBhandary, U. L. and Vournakis, J. N. (1980) Cell, in press.
18. Levis, R. and Penman, S. (1978) J. Mol. Biol. 121, 219-238.
19. Efstratiadis, A., Vournakis, J. N., Donis-Keller, H. Chaconas, G., Dougall, K. D. and Kafatos, F. C. (1977) Nucl. Acids Res. 4, 4165-4174.
20. Pilly, D., Niemeyer, A., Schmidt, M. and Bargetzi, J. P. (1978) J. Mol. Biol. 253, 437-445.
21. Lockard, R. E., Alzner-DeWeerd, B., Heckman, J. E., MacGee, J., Tabor, M. W. and RajBhandary, U.L. (1978) Nucl. Acids Res. 5, 37-56.
22. Rushizky, G. W., Shaternikov, V. A., Mozejko, J. H., and Sober, H. A. (1975) Biochem. 14, 4221-4226.
23. Cox, R. A. (1968) Methods in Enzymology 12B, 120-129.
24. England, T. E. and Uhlenbeck, O. C. (1978) Nature 275, 560-561.
25. Donis-Keller, H., Maxam, A. M. and Gilbert, W. (1977) Nucl. Acids Res. 4, 2527-2531.
26. Pace, N. R., Walker, T. A. and Schroeder, E. (1977) Biochem. 16, 5321-5328.
27. Glover, D. M., White, R. L., Finnegan, D. L. and Hogness, D. S. (1975) Cell 5, 149-155.
28. Southern, E. M. (1975) J. Mol. Biol. 98, 503-517.
29. Roychoudhury, R., Jay, E. and Wu, R. (1976) Nucl. Acids Res. 3, 863-877.
30. Maxam, A. M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
31. Pipas, J. M. and McMahon, J. E. (1975) Proc. Nat. Acad. Sci. USA 72, 2017-2021.
32. Tinoco, I., Borer, D. N., Dengler, B., Levine, M., Uhlenbeck, O. C., Crothers, D.M. and Gralla, J. (1973) Nature N. B. 246, 40-41.
33. Simoncits, A., Brownlee, G. G., Brown, R. S., Rubin, J. R. and Guilley, H. (1977) Nature 269, 833-836.
34. Nazar, R. N., Sitz, T. O., and Busch, H. (1975) J. Biol. Chem. 250, 8591-8597.
35. Luoma, G. A. and Marshall, A. G. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 4901-4905.
36. Wellauer, P. K., Dawid, I. B. and Tartof, K. D. (1978) Cell 14, 269-278.
37. Glover, D. M. and Hogness, D. S. (1977) Cell 10, 167-176.
38. Dougan, G., Saul, M., Warren, G. and Sherratt, D. (1978) Molec. gen. Genet. 158, 325-327.
39. Glover, D. M. (1977) Proc. Natl. Acad. Sci. USA 74, 4932-4936.
40. Jordan, B. R. and Glover, D. M. (1977) FEBS Lett. 78, 271-274.
41. Boseley, P. G., Tuyns, A. and Birnstiel, M. L. (1979) Nucl. Acids Res. 5, 1121-1137.
42. Khan, M. S. N. and Maden, B. E. H. (1977) Nucl. Acids Res. 4, 2495-2505.
43. Lightfoot, D. (1978) Nucl. Acids Res. 5, 3565-3577.
44. Klessig, D. F. (1977) Cell 12, 9-12.
45. Tonegawa, S., Maxam, A. M., Tizard, R., Bernard, O. and Gilbert, W. (1978) Proc. Nat. Acad. Sci. USA 75, 1485-1489.

46. Shine, J. and Dalgarno, L. (1973) J. Mol. Biol. 75, 57-72.
47. Wellauer, P. K. and Dawid, I. B. (1977) Cell 10, 193-212.