Volume 9 Number 8 1981 Nucleic Acids Research

The nucleotide sequence at the transcription termination site of ribosomal RNA in Drosophila melanogaster

Radha K.Mandal[+] and Igor B.Dawid

Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20205, USA

Received 24 February 1981

ABSTRACT

The sequence of 300 nucleotides surrounding the transcription termination site for ribosomal RNA in *Drosophila melanogaster* has been determined. The precise position of the RNA 3'-end was located by sizing rDNA fragments protected from S_1 nuclease by hybridization with cytoplasmic 28S rRNA or with nuclear 38S pre-rRNA. Next, the sequence of the first 10 nucleotides at the 3'-end of 28S and 38S rRNA was determined directly. The results indicate the absence of 3'-terminal processing of the 38S pre-rRNA. The nucleotide sequence from a position 27 nucleotides upstream to 150 nucleotides downstream from the termination site was identical in three uninterrupted rRNA genes and one interrupted gene, except for a single base change in the spacer region of one clone. The terminal 100 nucleotides of *Drosophila* 28S rRNA have more than 60 per cent homology with the corresponding region of *Xenopus* 28S rRNA and yeast 26S rRNA.

INTRODUCTION

In the fruitfly *Drosophila melanogaster*, as in other eukaryotic organisms, ribosomal RNA genes are tandemly repeated, each repeat unit being composed of alternating gene and nontranscribed spacer (NTS) regions (1-3). The 28S, 18S and 5.8S rRNA molecules are transcribed as parts of a large precursor molecule which is subsequently processed in several steps to yield the mature rRNA species (4, 5). The structure and function of rRNA genes are of special interest as they are transcribed by a distinct class of RNA polymerase (polymerase I) and are under specific control during development (6-8). The transcription initiation and termination sites have been analyzed at the nucleotide level in *Xenopus laevis* (9), and yeast (10), and the initiation site has been reported for the mouse (11).

Drosophila rDNA is particularly interesting in that many of its repeating units are interrupted in the 28S RNA coding region by ribosomal insertions (1). In *Drosophila*, interrupted rRNA genes do not contribute significantly to the production of rRNA (12,13,14). In an effort to analyze

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

in detail regions of D. melanogaster rDNA that may be involved in the control of its expression we report here the location and sequence of the site of transcription termination. The sequence of the transcription initiation site of these genes has been reported recently (15), and the present work describes the termination site in interrupted and uninterrupted rRNA genes.

MATERIALS AND METHODS

Isolation of DNA and restriction fragments. The recombinant plasmids (pDmr Y22 and a56) and phages (λDmr 312 and 326) used in this study and their purification have been previously described (15,16). After digestion of DNA with restriction enzymes under conditions recommended by the suppliers (BRL and New England Biolabs), DNA fragments were separated by electrophoresis on 0.8% agarose (Seakem agarose, Marine Colloids) or 5% polyacrylamide gels and extracted from agarose gels (17) or polyacrylamide gels (18). In some cases, the end-labeled fragments were isolated from low-melting agarose (Seaplaque, Marine Colloids) as described by Wieslander (19).

DNA labeling and sequencing. Labeling at the 5' end was achieved by using γ-^{32}P-ATP and T_4 polynucleotide kinase as described (18). Labeling at recessed 3' end was done by the appropriate α-^{32}P-deoxynucleotide and AMV reverse transcriptase. In a final volume of 50 µl, 5 pmol of 3' ends was incubated with 250 µCi of α-^{32}P-deoxynucleotide (3,000 Ci/mmol) in 10 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 50 mM KCl, 1 mM dithiothreitol at 37°C for 1 hr. DNA sequencing was carried out by the base-specific chemical cleavage method of Maxam and Gilbert (18). The chemically cleaved fragments were electrophoresed in 0.4 mm thick 20%, 14% and 8% polyacrylamide gels containing 8M urea.

RNA isolation and sequencing. Nuclear and cytoplasmic RNA was isolated from Drosophila embryos as described (12). The precursor RNA was purified by SDS-sucrose gradient centrifugation. The heavy side of the 38S region from several gradients was pooled and rerun. 28S rRNA was purified by SDS-sucrose gradient centrifugation of cytoplasmic RNA. The RNAs were labeled at the 3' end by ^{32}P-pCp and T_4 RNA ligase (20). The 3'-terminal RNA fragments were isolated by hybridization to nitrocellulose-bound HhaI spacer fragment of rDNA which includes the terminal coding segment (15). The end-labeled RNAs were sequenced by the method of Donis-Keller et al. (21), the alkaline ladder being produced by incubation with 0.1 M NaOH at 37°C for 4 min instead of using bicarbonate buffer. The digests were run in a 0.4 mm thick 20% polyacrylamide gel containing 8M urea. For 3'-terminal nucleo-

tide determination, complete alkaline digests (0.3 M KOH at 37°C, 16 hr) of the end-labeled RNAs were electrophoresed on Whatman 3MM paper in pyridine-formate buffer, pH 3.5, with 2', 3'-ribonucleotides as markers (22).

S_1 nuclease mapping. Labeled DNA was hybridized with nuclear pre-rRNA or cytoplasmic 28S RNA in 70% formamide, 0.3 M NaCl, 40mM Tris-HCl (pH 8.0), 1 mM EDTA at 41°C. The mixture was diluted 10 fold with S_1 digestion buffer (50 mM NaOAc, pH 4.75; 0.15 M NaCl; 0.5 mM $ZnSO_4$) containing 10 μg/ml of native pBR322 DNA as carrier and treated at 37°C for 1 hr with 125-1000 Miles units of S_1 nuclease. The digestion was stopped by adding EDTA (5 mM), SDS (0.1%) and 10 μg of *E. coli* tRNA. DNA was subsequently phenol extracted, precipitated with ethanol, and electrophoresed in alkaline agarose gels or in 14% sequencing polyacrylamide gels.

RESULTS

Nucleotide sequence surrounding the transcription termination site. For sequencing the transcription termination site, we used primarily the cloned rDNA fragment pDmr Y22 which contains an 11 kb EcoRI fragment derived from two adjacent ribosomal genes (see ref. 16 and Figure 1). Previous electron microscopic and restriction-hybridization data located the 3' end of the 28S rRNA gene near a HindIII site (16). We mapped several other restriction sites in this region. Figure 1 shows the restriction map of the relevant part of the rDNA and the strategy used for sequencing. The sequence of about 150 nucleotides in either direction from the transcription termination site was determined by sequencing either both strands or in both directions (Figure 1). Figure 2 presents the sequence of the noncoding strand in this region. In this sequence, the position of transcription termination (determined as described below) is marked by a vertical arrow. Some features of this sequence deserve comment. The transcription termination site is not at a T cluster (see Discussion). Only a tetranucleotide dyad symmetry is present, located 10 nucleotides upstream from termination. The terminal hexanucleotide sequence ATTCGA is present in opposite polarity 20 nucleotides upstream. The hexanucleotide TTGTAA is directly repeated three times within 100 nucleotides.

Determination of the 3' terminus by S_1 protection. The five kb HaeIII fragment containing the terminal region of the 28S rRNA gene and the entire NTS was labeled at its 3' ends. After denaturation, samples of the labeled fragment were hybridized separately with cytoplasmic 28S RNA and nuclear 38S pre-rRNA. Aliquots of the reannealed samples were then digested with

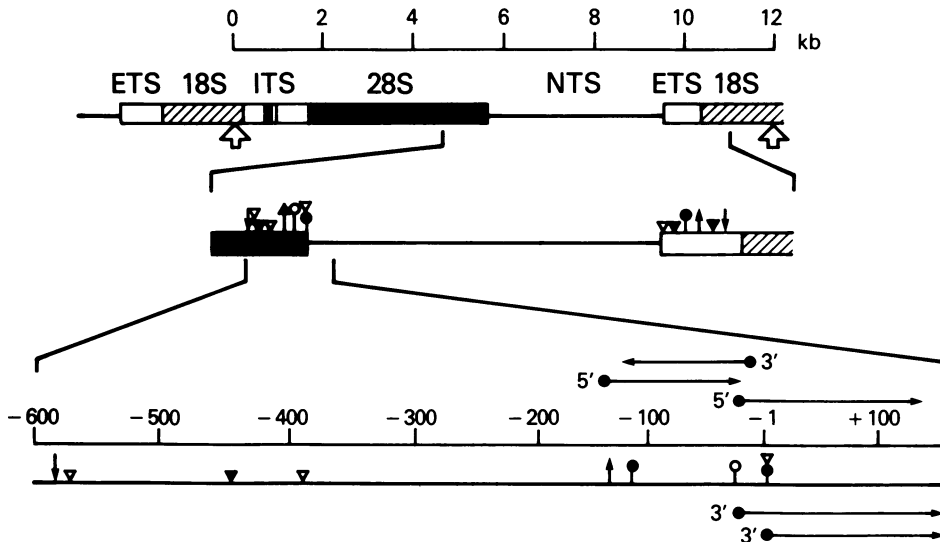

Figure 1. Strategy for sequencing the rRNA transcription termination region in pDmr Y22. The termination region is shown enlarged with sites for the restriction endonucleases: HindIII (⚲), HinfI (⚲), TaqI (▽), RsaI (▼), HaeIII (↓) and HhaI (↑). The EcoRI site in the rDNA repeat unit is shown by ⇧. Base pairs are numbered such that those in the 28S RNA gene are negative starting with the 3' terminal nucleotide as -1. The arrows indicate the direction of sequencing of restriction fragments labeled at their 3' or 5' ends, as indicated.

increasing amounts of S_1 nuclease, and the protected fragments were electrophoresed in an alkaline agarose gel. Figure 3 shows the results with only one concentration of S_1. The protected fragment of the coding strand is slightly larger than the HaeIII-HindIII marker fragment, which has a size of about 550 nucleotides. This experiment also shows that the fragments protected by 28S and 38S RNA have about the same length. When the rRNA was omitted from the hybridization mixture, no protected fragment was found (data not presented). Thus the 3' termini of the 28S and 38S RNA genes are located slightly downstream of the HindIII site (see Figure 1).

To locate the 3' end precisely, it was necessary to protect a smaller fragment that could be run on a sequencing gel. For this purpose the coding strand of the HindIII - HaeIII spacer fragment was labeled at the 3' end with α-^{32}P-dATP and reverse transcriptase, which labels only the 5' overhanging HindIII site, but not the HaeIII site. Part of this labeled DNA was cut with AluI and hybridized with 28S RNA. After digestion with S_1 nuclease

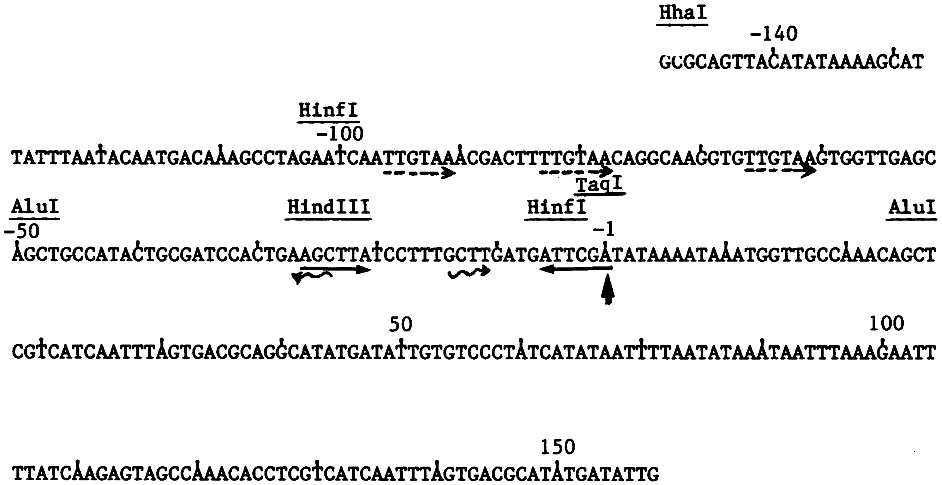

Figure 2. Nucleotide sequence of the non-coding strand of rRNA transcription termination region in pDmr Y22. The 3'-terminal nucleotide is pointed out by a vertical arrow, and has position -1. The positions of the restriction sites determined experimentally are indicated. The inverted repeats are shown by solid arrows, direct repeats by broken arrows and the dyad symmetry by wriggled arrows.

and denaturation, the protected fragment was run in parallel with the original labeled fragment cleaved by base-specific chemical reactions (Figure 4). The most prominent band protected at the highest amount of S_1 (lane 4) was two nucleotides larger than the HindIII-HinfI fragment (lane 1). When a correction is made for the different migration of an S_1-generated and a chemically cleaved fragment (which is one nucleotide shorter) the last transcribed nucleotide is found to be T. This maps the 3'-terminal nucleotide at a position 26 nucleotides downstream from the HindIII site. A similar experiment was carried out with 38S pre-rRNA and resulted in identification of the same terminal nucleotide as for 28S RNA (see also below).

Sequencing the 3' termini of 28S and 38S rRNA. The position of transcription termination was also obtained by direct determination of the sequence of several nucleotides at the 3' end of both 38S and 28S rRNA. The RNAs were labeled at the 3' end with $5'$-^{32}P-pCp and T_4 RNA ligase. To purify a homogeneous short terminal fragment free from any degradation products formed during the labeling procedure, the labeled RNA was hybridized with the HhaI spacer fragment of pDmr Y22 which contains the 3'-terminal 148 nucleotides of the coding strand (see Figure 1). After RNase treatment, the

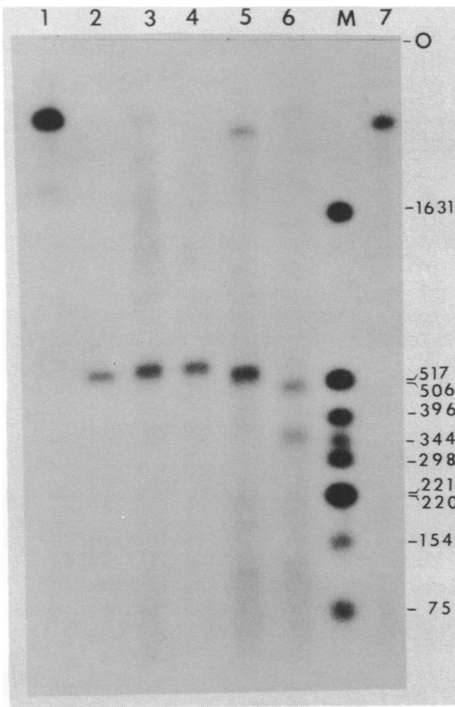

Figure 3. Protection of 3'-^{32}P-HaeIII spacer fragment of pDmr Y22 by hybridization with cytoplasmic 28S rRNA and nuclear 38S pre-rRNA against digestion with S_1 nuclease. The 5 kb HaeIII fragment containing the terminal region of the 28S rRNA gene and nontranscribed spacer (see Fig. 1) was labeled at its 3' ends. The labeled DNA was hybridized separately with 28S rRNA and with 38S pre-rRNA. Aliquots were digested with S_1 nuclease and electrophoresed in a 1.5% alkaline agarose gel. Lanes 1, 7- labeled HaeIII fragment; 2 - purified HaeIII - HindIII fragment 3'-labeled at HindIII site; 3 - HaeIII fragment protected from S_1 digestion (1000 u) by 28S RNA; 4 - HaeIII fragment protected from S_1 digestion (1000 u) by 38S RNA; 5-labeled HaeIII fragment cut by HindIII; 6 - HaeIII fragment cut by HinfI; M - pBR322 HinfI digest as marker.

hybridized RNA fragment was eluted and subjected to sequencing by partial enzymatic cleavage as described under Methods (Figure 5). The identical sequence 5'-GPyPyPyGAPyGAPyPyPyG-3' was derived for both 28S and 38S RNA. Since both alkaline and enzymatic cleavage at the first phosphodiester bond between the 3'-terminal nucleotide of RNA and 5'-^{32}P-pCp gives rise to an unlabeled product, the 3'-terminal nucleotide could not be read from this gel. To determine the 3'-terminal nucleotide, aliquots of the labeled fragment were totally digested with alkali and the resulting 2', 3'-nucleotides were separated by high-voltage paper electrophoresis (22). This procedure yielded a single radioactive spot which comigrated with AMP, showing that the terminal nucleotide for both 38S and 28S RNA is A. The 3'-terminal A corresponds to the T in the coding strand of DNA as shown by arrows in Figures 2 and 4. Thus direct RNA sequencing is in agreement with S_1 protection in identifying the 3' terminus for 28S and 38S rRNA.

The sequence from the HindIII site to 150 nucleotides into the nontranscribed spacer, spanning the termination site, was also determined in one interrupted rRNA gene cloned in a plasmid (pDmr a56) and two uninterrupted

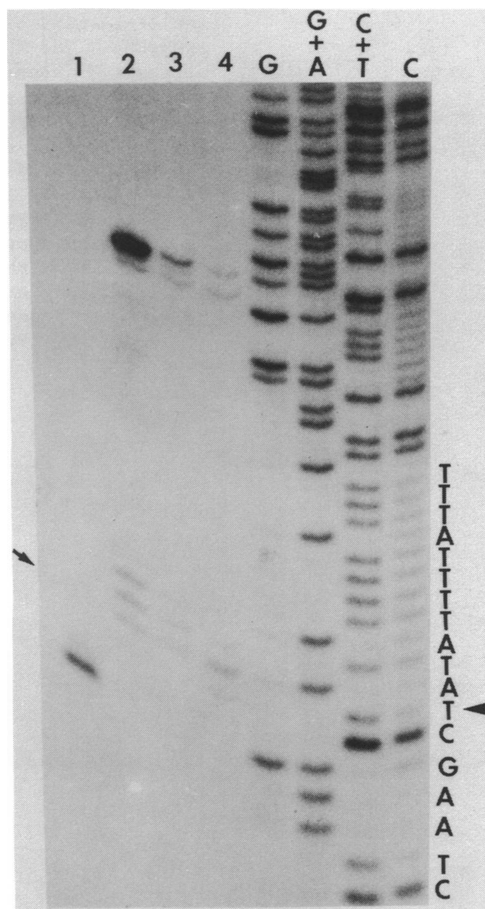

Figure 4. Localization of the transcription termination site by sizing of the S_1-protected fragment on a sequencing gel. The coding strand of pDmr Y22 HindIII - HaeIII spacer fragment spanning the termination site was labeled at its 3'-end with α-^{32}P-dATP and reverse transcriptase, and sequenced. In parallel, part of the terminally labeled fragment was cut with AluI, hybridized with 28S rRNA and digested with different levels of S_1 nuclease. The protected fragments were denatured and electrophoresed in a 14% sequencing gel, along with the fragments cleaved by base-specific chemical reactions. The arrows mark corresponding positions of the fragment protected at the highest level of S_1 used, after correction being made for 1 nucleotide displacement in migration of the chemically cleaved fragments. Lane 1, HindIII - HaeIII fragment cut with HinfI; 2, 3, 4, HindIII - AluI fragment protected from 125, 500, and 1250 units of S_1, respectively, by hybridizing with 28S rRNA; G, G + A, T + C, and C, fragments generated by base-specific reactions. The bands at a higher position in lanes 2,3,4 correspond to the HindIII-AluI fragment which is not fully digested by S_1 nuclease, possibly due to renaturation.

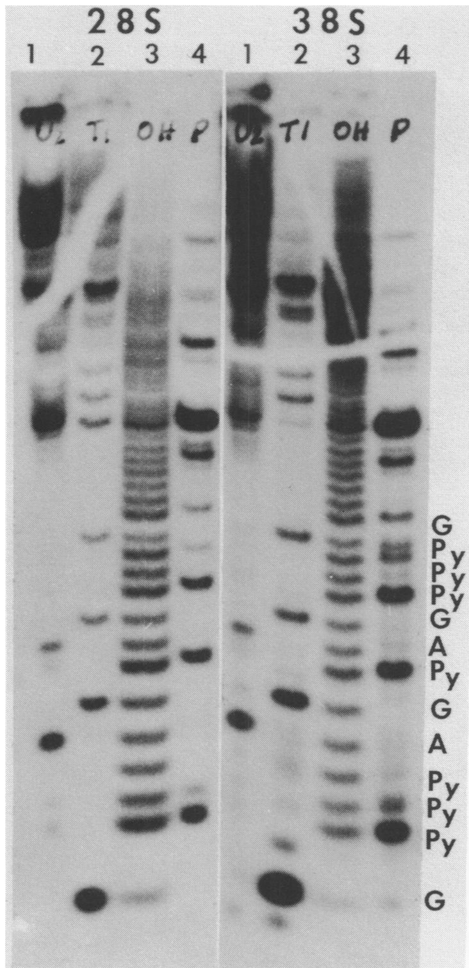

Figure 5. Nucleotide sequence at the 3' termini of 28S rRNA and 38S pre-rRNA. The RNAs were labeled at their 3' ends with 5'-^{32}P-pCp and T_4 RNA ligase, hybridized with HhaI-HhaI spacer fragment from pDmr Y22 (Fig. 1) and eluted from the filters as described in the text. Aliquots of the purified 3'-terminal fragments were partially cleaved and run on 20% polyacrylamide sequencing gels. Lane 1, U_2 digest; 2, T_1 digest; 3, alkaline ladder; 4, pancreatic RNase digest. Py denotes pyrimidine.

rRNA genes cloned in lambda (λDmr 312 and 326) (for a description of these clones see ref. 15 and 16). These sequences are identical except for position 91 (see Figure 2) in a56 where T is replaced by G. This mutation in the spacer is not likely to have functional significance.

DISCUSSION

The results presented in this paper establish the 3'-terminal sequences of the 28S rRNA and 38S pre-rRNA in *D. melanogaster* and give the sequence of a region of 310 nucleotides in rDNA centered at the termination position.

Three conclusions derive from these results.

There is no processing at the 3' end of Drosophila pre-rRNA. This conclusion is based on protection of DNA fragments against digestion with S_1 exonuclease by hybridization with 28S or 38S RNA, and more firmly on direct sequencing of the 14 3'-terminal nucleotides of both RNAs (Figure 5). The absence of 3'-terminal processing has also been suggested for *Xenopus laevis* pre-rRNA on the basis of S_1 protection experiments (9). In contrast, 3'-terminal processing has been indicated for two other organisms. Veldman et al. (10) reported that yeast pre-rRNA is longer by 7 nucleotides than the mature 26S rRNA, and Hamada et al. (23) reported that 45S pre-rRNA, processing intermediates, and mature 28S rRNA in rat liver and mouse hepatoma have different 3' termini. It might be argued that the 8 kb pre-rRNA molecule that we studied is not the primary transcript and has already been processed at its 3' end. We believe this to be unlikely because this RNA still retains di- and triphosphate residues at its 5' end (4), and because it is quite comparable in kinetic properties (4) to the yeast and rat pre-rRNAs that apparently contain 3'-terminal nucleotides that are not retained in mature rRNA. Therefore, these results suggest a basic difference in the 3' ends of pre-rRNAs of different organisms.

The termination sequence in D. melanogaster is unusual. Many transcripts from prokaryotes and eukaryotes terminate in a stretch of U residues (24, 25). The examples for this generalization include the rRNA of yeast (10). In *X. laevis*, the pre-rRNA appears to terminate just upstream of a stretch of 4 T residues in the noncoding strand of the DNA (9). *D. melanogaster* rDNA differs in that the termination sequence does not include a stretch of U's in the RNA and is not followed by a stretch of T's in the noncoding strand for a considerable distance (Figure 2). The termination region is, however, AT rich, in agreement with observations in some other systems (25). Since the entire nontranscribed spacer in *D. melanogaster* is AT rich the significance of the base composition in the region immediately downstream from termination cannot be assessed.

Termination regions generally contain a dyad symmetry in the RNA sequence immediately preceding termination (24,25). This is also true of *X. laevis* rDNA, but the dyad symmetry in this case involves only 4 nucleotides (9). A 4 nucleotide dyad symmetry is also present in *D. melanogaster* rDNA at positions -11 to -14 and -24 to -27 (Figure 2). The hairpin loop that could form on the basis of this symmetry is not stable and the significance of this structure is unknown.

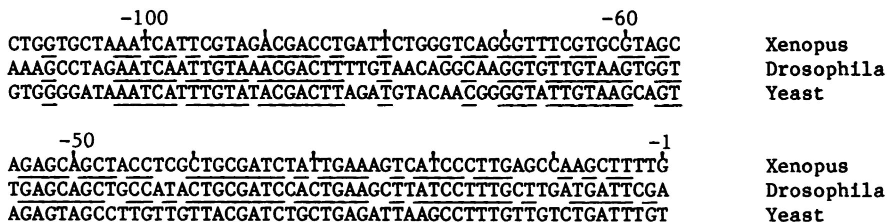

Figure 6. Sequence comparison of 3'-terminal regions of 28S RNA from *Drosophila melanogaster*, *Xenopus laevis* (9), and yeast (10). Homologous nucleotides are underlined. The numbering is according to Figure 2 for *Drosophila*; *X. laevis* rRNA terminates at position -6, while yeast 26S RNA terminates at -1, but pre-rRNA continues to position +7.

There is sequence homology between the 3'-terminal regions of different rRNAs. Figure 6 gives an alignment of the terminal regions of the rRNAs of *X. laevis*, yeast and *D. melanogaster*. There is considerable homology between these sequences, as already noted for *X. laevis* and yeast by Veldman et al. (10). It has long been known that rRNA sequences are highly conserved (26). Within this context of conservation the 3'-terminal region is not particularly stable; a much more conserved region has been characterized in the interior of 28S rRNA (27). In the spacer regions downstream from the termination site there is no detectable homology between the rDNAs of *X. laevis*, yeast and *D. melanogaster*.

ACKNOWLEDGMENT

Thanks are due to Eric Long and Martha Rebbert for suggestions and technical help.

+ On leave from: Biochemistry Department, Bose Institute, Calcutta 700009, India

REFERENCES

1. Long, E.O. and Dawid, I.B. (1980) Annu. Rev. Biochem. 49, 727-764
2. Wellauer, P.K. and Dawid, I.B. (1974) J. Mol. Biol. 89, 397-407
3. Pavlakis, G.N., Jordan, B.R., Wurst, R.M. and Vournakis, J.N. (1979) Nucleic Acids Res. 7, 2213-2238
4. Levis, R. and Penman, S. (1978) J. Mol. Biol. 121, 219-238
5. Long, E.O. and Dawid, I.B. (1980) J. Mol. Biol. 138, 873-878
6. Biswas, B.B., Ganguly, A. and Das, A. (1975) In Progress in Nucleic Acid Research and Molecular Biology, Cohn, W.E. ed., Vol. 15, pp. 145-184, Academic Press, New York

7. Brown, D.D. and Littna, E. (1966) J. Mol. Biol. 20, 81-94
8. Gurdon, J.B. and Woodland, H.R. (1969) Proc. Royal Soc. B173, 99-111
9. Sollner-Webb, B. and Reeder, R.H. (1979) Cell 18, 485-499
10. Veldman, G.M., Klootwijk, J., DeJonge, P., Leer, R.J. and Planta, R.J. (1980) Nucleic Acids Res. 8, 5179-5192
11. Urano, Y., Kominami, R., Mishima, Y. and Muramatsu, M. (1980) Nucleic Acids Res. 8, 6043-6058
12. Long, E.O. and Dawid, I.B. (1979) Cell 18, 1185-1196
13. Jolly, D.J. and Thomas, C.A. (1980) Nucleic Acids Res. 8, 67-84
14. Long, E.O., Rebbert, M.L. and Dawid, I.B. (1980) Cold Spring Harbor Symp. Quant. Biol. 45, in press.
15. Long, E.O., Rebbert, M.L. and Dawid, I.B. (1981) Proc. Natl. Acad. Sci. USA, in press.
16. Dawid, I.B., Wellauer, P.K. and Long, E.O. (1978) J. Mol. Biol. 126, 749-768
17. Tabak, H.F. and Flavell, R.A. (1978) Nucleic Acids Res. 5, 2321-2332
18. Maxam, A.M. and Gilbert, W. (1980) In Methods in Enzymology, Grossman, L. and Moldave, K. eds., Vol. 65, pp. 499-560, Academic Press, New York
19. Wieslander, L. (1979) Anal. Biochem. 98, 305-309
20. England, T.E., Bruce, A.G. and Uhlenbeck, O.C. (1980) In Methods in Enzymology, Grossman, L. and Moldave, K. eds., Vol. 65, pp. 65-74, Academic Press, New York
21. Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527-2538
22. Brownlee, G.G. (1972) in Laboratory Techniques in Biochemistry and Molecular Biology, Work, T.S. and Work, E. eds., Vol. 3, pp. 1-265, Elsevier - North Holland, New York
23. Hamada, H., Kominami, R. and Muramatsu, M. (1980) Nucleic Acids Res. 8, 889-903
24. Korn, L.J. and Brown, D.D. (1978) Cell 15, 1145-1156
25. Adhya, S. and Gottesman, M. (1978) Annu. Rev. Biochem. 47, 967-996
26. Sinclair, J.H. and Brown, D.D. (1971) Biochemistry 10, 2761-2769
27. Rae, P.M.M., Kohorn, B.D. and Wade, R.P. (1980) Nucleic Acids Res. 8, 3491-3504