Volume 8 Number 1 1980 Nucleic Acids Research

Nuclear RNA transcripts from Drosophila melanogaster ribosomal RNA genes containing introns

D.J.Jolly* and C.A.Thomas,Jr.

Department of Cellular Biology, Scripps Clinic and Research Foundation, La Jolla, CA 92037, USA

Received 23 July 1979

ABSTRACT

The transcription of ribosomal genes in a cell line (Kc) derived from female *D.melanogaster* , has been investigated using hybridization probes prepared from restriction fragments of a cloned rDNA repeat with a 5 kb type I [Wellauer et al. (1978) Cell 14, 269-278] *intron*. Gels, of nuclear RNA that have been transferred to diazotized paper and hybridized with labelled *intron* sequences, reveal both large (1-10 kb) transcripts and a discrete 325 base species. Berk-Sharp experiments [(1977) Cell 12, 721-732] reveal large transcripts that are homologous to *intron* sequences and extend into 28S sequences as well. However, while the abundance of 28S transcripts is about 50,000 copies per nucleus [Clark et al. (1977) Genetics 86, 789-800], these long transcripts are only present at 1-2 copies per nucleus and the 325 base species is only 10 times more abundant. In view of the fact that female cells have about 400 rDNA genes, 49% of which have type I *introns*, one must conclude either that transcription rarely occurs from the genes containing *introns* (the majority) or these transcripts are processed unusually rapidly. Transcripts homologous to the "non-transcribed spacer" region have been found, but their abundance is no higher.

INTRODUCTION

The discovery that genes in eukaryotes exist in coding pieces or *exons*, interspersed by noncoding pieces or *introns*[1,2], has altered our picture of gene structure, transcription and the control of gene expression. The first example of an *intron* to be found was in *D. melanogaster* ribosomal DNA (rDNA) [3,4,5]. Subsequently the list of interrupted genes has grown to include such diverse types as mammalian globin [6,7], chicken ovalbumin [8,9], yeast transfer RNA (tRNA) [10], protozoan ribosomal RNA (rRNA) [11,12], and yeast mitochondrial rRNA [13]. In most of these cases it has been shown, or inferred, that the *intron* is transcribed and subsequently processed out by RNA splicing events. In addition, in the cases so far examined [14,15] the *intron* does not occur elsewhere in the chromosome.

In contrast the *intron* in *D. melanogaster* rDNA has not been shown to be transcribed, and conceivably does not have to be, since there are a large number of rDNA repeats without *introns* [16]. On the X chromosome, 49% of the 28S rDNA copies contain type I *introns*, 16% have type II, while the remainder have none. About 16% of the rDNA copies on the Y chromosome have type II *introns* only, while the remainder have no detectable *introns* [17,18]. Type I sequences are repeated many times elsewhere in the chromosome [19]. A further tantalizing point is that not all genes have *introns*, including *Xenopus* rDNA [20] and all *D. melanogaster* structural genes so far characterized [21,22,23]. For all these reasons, the ribosomal *introns* in *D. melanogaster* appear atypical and the status of their transcription worth investigating in some detail.

To do this we have used two extremely sensitive hybridization techniques which have been recently developed [24,25] to screen nuclear RNA (nRNA) for molecules homologous to a cloned DNA *intron* fragment. The nRNA was purified from Kc tissue culture cells [26] which were derived from female (XX) tissue and hence should have only type I *introns* in about 49% of their rDNA repeats. Various probes were prepared from subfragments of cDm103 [4] that correspond to *intron*, 28S, 18S and spacer sequences. These were used to measure the size of the homologous nuclear transcripts and to estimate their abundance.

MATERIALS AND METHODS

Growth of cells and isolation of RNA. Kc *Drosophila* tissue culture cells were maintained in D20 medium [26] without serum. 18 and 28S rRNA were prepared as described by Spear [27] after labelling at 20 μCi/ml of ^{3}H-uridine (New England Nuclear) for 24 hours. Cells were heat shocked, when desired, for one hour at 37°C and pulse labelled during this period with 50 μCi/ml of ^{3}H-uridine [28].

Nuclear RNA from heat shocked and normal cells was prepared by the methods of Levis and Penman [29,30]. Typically cells were scraped off two 1000 ml T-flasks in 10 ml medium, and poured over two volumes of 10mM Tris, 100 mM NaCl, 5 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.2 in the form of crushed ice. The cells were spun down, washed, then resuspended in 2 ml of 30mM Tris-HCl, 100 mM NaCl, 5 mM $MgCl_2$, 5 mM $CaCl_2$, 0.5% Triton X-100, pH 8.3, freshly treated with 10 μl diethyl pyrocarbonate. After 2' on ice, the nuclei were spun down, washed in the same buffer, then digested with 100 μg RNase-free DNase I (Sigma) in 1 ml of 10 mM Tris-HCl, 0.5M NaCl, 50mM

$MgCl_2$, 25 µg/ml polyvinyl sulfate, 5mM N-ethyl maleimide, pH 7.4, at 37°C for 15 seconds. Sodium dodecyl sulfate (SDS) was added to 0.5% and the RNA precipitated with 2 volumes of ethanol. The precipitate was taken up in 1ml of 10mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 0.5% SDS, pH 7.4, and digested with 250 µg proteinase K (E.M. Laboratories) at 37°C for 30'. An equal volume of 10 mM Tris-HCl, 100 mM NaCl, 10 mM EDTA, 0.5% SDS, pH 7.4 was added, the solution extracted twice with a 50:50 phenol: chloroform mix, the RNA precipitated with 2 volumes of ethanol, and resuspended in water.

Isolation and restriction of bacterial plasmids. Bacteria carrying plasmids cDm103 (provided by Dr. G. Rubin) and pRp2 (from a clone library of *D. melanogaster* DNA fragments generated by terminal *Eco*R1 digestion and screened for homology to rRNA) which provided probe E, were grown, the plasmids amplified by chloramphenicol addition, and the plasmid DNA purified as described by Clewell and Helinski [31]. Restriction enzyme *Eco*R1 was purified by a modification of the method of Bingham et al. [32]. *Bam*H1 was a gift of Dr. S. Smith and *Hind*III was purchased from Bethesda Res. Labs, (Rockville, Md). Approximately 50 µg of plasmid DNA was digested with *Eco*R1 plus *Bam*H1 or *Eco*R1 plus *Hind*III in 100 mM Tris-HCl, 75mM NaCl, 7 mM $MgCl_2$, 2mM β-mercaptoethanol, pH 7.4. DNA fragments were isolated by running the appropriate digest on 1.0 cm diameter preparative agarose tube gels in Tris-acetate buffer [33]. The DNA bands were cut out after staining lightly with ethidium bromide, and the DNA eluted from the agarose by the perchlorate glass fiber filter method of Chen and Thomas [34].

In vitro labelling of nucleic acids. DNA hybridization probes were labelled by nick translation [35] in 50 mM Tris-HCl, 5 mM $MgCl_2$, 10 mM β-mercaptoethanol, 100 µg/ml gelatin with 3 unlabelled deoxynucleotide triphosphates at 20 µM and one α-^{32}P-deoxynucleotide triphosphate (400 Ci/mmole, New England Nuclear) at 5 µM. Samples (0.1-0.5 µg) were incubated with 2 ng/ml DNase I for 15' at 37°C, then 2-4 units of DNA polymerase I (Boerhinger-Mannheim) and incubation continued at 15°C for 2 hours. This gave specific activities of 5×10^7 to 10^8 cpm/µg. Labelled 28S RNA sequences were prepared by reverse transcription followed by a DNA copying reaction. 0.5 µg of 28S rRNA was transcribed into cDNA using AMV reverse transcriptase primed with random oligonucleotides from DNase I digested calf thymus DNA [36]. The reaction (50 µl) took place at 42°C with 5 units added reverse transcriptase, in 50 mM Tris, 8 mM KCl, 10mM $MgCl_2$, 7 mM dithiothreitol, pH 7.8 with 400 µM dGTP, dCTP and dTTP, and 40 µM dATP. After 60' the reaction was heated to 100°C for 3', the de-

natured protein removed by centrifugation and the nucleic acid precipitated by adding one tenth volume 1.5 M sodium acetate, 40 µg *E. coli* tRNA and 2.5 volumes 95% ethanol. The pellet was washed once in 70% ethanol, 0.1M sodium acetate, dried and resuspended in 30 µl nick translation buffer with 50 µM dGTP, dCTP and dTTP and 4 µM α-^{32}P-labelled dATP. Four units of DNA polymerase I were added and the reaction, which proceeds by virtue of priming by the turnarounds left by reverse transcriptase [37], incubated at 19°C for 90'. The turnaround was subsequently broken by treatment with S1 [38].

Purity of restriction fragments. Probes B and C (see Fig. 1) were checked for cross-contamination by hybridization to Southern blots of cDm 103 cut by *Hind*III and *Bam*H1. There was no detectable contamination as judged by the autoradiographic responses.

RNA gels and blots. RNA samples were electrophoresed after heating to 65°C in 1.5% agarose gels, 2 mM thick, in Tris-borate buffer containing no denaturing agent [39]. After electrophoresis, channels containing tritiated markers were cut off and subsequently fluorographed [40]. The portion of the gel that contained samples of interest was soaked in 50 mM NaOH for 5' at room temperature, then transferred to diazotized cellulose using potassium phosphate buffers at pH 6.5 [25]). The diazotized cellulose was prepared from Whatman 540 paper and 1[(M-nitrobenzyloxy)methyl] pyridinium chloride by the method of Alwine et al. [25].

Berk-Sharp experiments [24]. 5 µg of RNA in water and 0.1 µg of cDm103, cut with *Eco*R1 in 1 mM Tris, 0.1 mM EDTA, pH 8 were mixed, lyophilized, taken up in 10 µl of 10 mM Pipes, 1 mM EDTA, 0.4M NaCl, 80% formamide, pH 6.4 and sealed in a 10 µl "microcap" pipette. The microcap was heated to 65°C for 5' to denature the nucleic acids, then incubated at 56°C to allow DNA-RNA hybridization. After two hours the microcap was broken open and the contents delivered into 100 µl of 0.25M NaCl, 0.03M sodium acetate, 1 mM $ZnSO_4$, pH 4.6 containing 50 µg/ml each of native and heat-denatured calf thymus DNA. An amount of S1 nuclease previously calibrated to render ^{32}P-labelled denatured λ DNA 95% TCA soluble in 10' in this system was added, the sample incubated at 45°C, and 25 µl aliquots withdrawn at 0', 3', 10' and 30'. In some subsequent experiments, all components were proportionately increased and most samples taken at 5'. The aliquots removed were added to 10 µl 125 mM NaOH, 20% glycerol, 0.1% bromphenol blue and loaded on alkaline agarose gels [41]. Unlabelled λ DNA digested with *Hind*III was also run, as a marker. The gels (13 cm x 26

cm) were run at 50 ma for about seven hours. After electrophoresis the gels were neutralized and blotted onto nitrocellulose sheets (Schleicher and Schull) [42].

Hybridizations to blots and autoradiographic detection. RNA blots were pretreated for 12 hours in 50% formamide, 5 x SSC (SSC is 0.15M NaCl, 0.015M sodium citrate, pH 7.4), 1% glycine, 1 mg/ml heat-denatured salmon-sperm DNA, 0.02% BSA, polyvinyl pyrollidone and Ficoll and then hybridized to about 5 x 10^5 cpm of probe per channel, in the same buffer without glycine. Hybridizations were conducted in sealed plastic bags at 42°C with shaking. For a sheet 13 cm x 15 cm, 10 ml solution was used. After hybridization for 24 hours, the filter sheet was washed 5 times for 60' with 50% formamide, 5 x SSC at 42°C, blotted dry and exposed to Kodak XR5 X-ray film using Dupont lightning plus intensifying screens, at -78°C.

Southern blots from Berk-Sharp experiments were pretreated 4-6 hours in 6 x SSC, 10 mM $NaPP_i$, 0.02% BSA, polyvinyl pyrollidone, and Ficoll [43] plus 250 µg/ml denatured salmon sperm DNA. Hybridization was for 36 hours at 68°C in the same buffer with about 5 x 10^5 cpm of probe per channel. If cold λ markers were on the blot, nick translated λ DNA was also included. Sealed plastic bags and shaking were employed as in the RNA blots. After hybridization the blots were washed in the bags twice with the hybridization buffer without salmon sperm DNA at 68°C for 45', once with 2 x SSC at 68°C for 10' and twice with 2 x SSC at 37°C for 10'. Hybridization was detected by autoradiography, as for the RNA blots.

Scans of autoradiographs. Autoradiographs were scanned using a Joyce-Leoble densitometer. Relative areas under the scans were calculated by cutting out the area desired and weighing.

RESULTS

Hybridization probes. In order to be able to detect transcripts of different parts of rDNA repeats, purified DNA and RNA fragments corresponding to these parts were prepared. The sizes and locations of these fragments with respect to cDm103 and the chromosomal transcription unit are shown in Fig. 1. The coding and restriction map of cDm103 show it to be equivalent to an rDNA repeat with the most common type of *intron* (type I, 5 kb long) [16,44]. The fragments were labelled to high specific activity and used as hybridization probes in subsequent experiments.

Hybridization to nRNA blots with probes B (*exon*) and C (*intron*). Probes B and C are approximately the same size, were labelled in an iden-

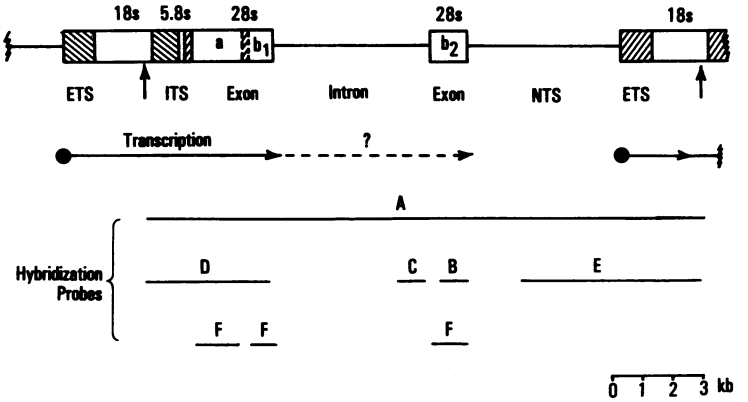

Figure 1. The map shown represents a *D. melanogaster* ribosomal DNA repeat equivalent to that cloned as an *Eco*R1 fragment in cDm103 [4] as supplemented by work of others [3,17]. The 18 and 28S rRNAs are processed from longer transcripts known to be initiated to the left of the external transcribed spacer (ETS) at the left of the 18S gene [17,48], then proceed through the internal transcribed spacer (ITS) and the 28S gene, and terminated at the nontranscribed spacer (NTS). Little is known about the transcription of those repeats that contain an *intron* within the 28S gene (*exon*). The present work made use of various cloned probes (A-E) and cDNA homologous to 28S ribosomal RNA (F). Probe A is cDm103 (used without separating the *col*E1 vehicle); probe B is a 900 bp *Hind*III fragment covering much of 28S b_2 (*exon*); probe C is a 800 bp *Bam*HI fragment that lies entirely within the *intron*; probe D is a 4200 bp *Eco*R1-*Hind*III fragment that includes 5.8S, 28S a and 28S b sequences; probe E is a 5900 bp *Eco*R1 fragment cloned in pML21 (used without separation from the vehicle) that contains 18S, ETS and NTS sequences. Probe F was prepared by making a cDNA copy of cDNA which in turn was made from purified 28S RNA.

tical fashion but correspond to mature 28S and *intron* sequences, respectively. Hence, using probe B as a standard it seemed possible to make an estimate of, or at least set an upper limit on the abundance of nRNA molecules bearing probe C sequences. Since it has been reported that heat shocking *Drosophila* cells blocks rRNA transcript processing at some point [45], nRNA from heat shocked cells (+hs nRNA) was examined, along with normal nRNA (-hs nRNA). Samples of nRNA of both types were loaded onto an agarose gel in sets of 3 slots; the slots contained amounts of RNA varying in the ratio 100:10:1. The samples were electrophoresed, the RNA transferred to diazotized cellulose and identical half-gel blots hybridized to probes B and C. This technique not only gives the relative abundances of the two sequences but also length characterizations of RNA

molecules in which they occur. The results are shown in Fig. 2. For probe B strong hybridization occurred in all channels; as expected hybridization occurred to bands at 7800 bases (33S), and 1900 bases (28S b). In addition there appeared to be hybridization to areas in the gel corresponding to 3700, 600 and 315 nucleotides. The 3700 base hybridization should correspond to 28S intact RNA, but the significance of the small species is unclear. There is no great difference between the +hs and -hs samples except that there appears to be more RNA in the +hs sample despite both being derived from the same number of cells. For probe C, a very different rather weaker pattern of hybridization appeared. Most of the detectable hybridization occurred in the most heavily loaded channels at a position corresponding to 325 ± 25% nucleotides which may be transcribed off the rDNA *intron*, or homologous sequences elsewhere in the genome [19]. Moreover, with -hs nRNA, definite bands could be detected all the way up the gel to a size of about 10,000 bases - considerably larger than the 33S rRNA precursor. These species were not clearly seen in +hs nRNA. Since this hybridization is considerably weaker than for the probe B experiment, it is possible that such higher molecular weight bands are present in the probe B experiment, but overshadowed by hybridization to the 33S species. From the amount of RNA loaded and the amount of probe B used, it was estimated that probe B (and hence also probe C) was in excess during hybridization; this was supported by the fact that most of the radioactive probe was still in solution after the hybridization experiment. Thus it was possible to estimate the relative abundances of probe B and C sequences in nRNA by densitomer scans of the autoradiographs. A single autoradiographic exposure was used, the probe B experiment scanned in the most lightly loaded channel and the probe C experiment in the most heavily loaded, so that the autoradiograph blackening would remain an approximately linear function of radioactivity to which it was exposed. The areas under the whole channel scans were calculated, corrected for the hundred-fold RNA loading factor and the ratio of the amounts of bound radioactive probe C to that of probe B calculated. For -hs nRNA this was 3.3×10^{-4}; for +hs nRNA it was 3.5×10^{-4}. Thus probe C (*intron*) sequences exist at a much lower level than probe B (*exon*) sequences. In addition most of the *intron* sequences are concentrated in the 325 base fragment, while the higher molecular weight RNA species, where *exon* and *intron* might conceivably exist on the same molecule, occur at at least an order of magnitude lower frequency than this fragment. Hence such molecules are

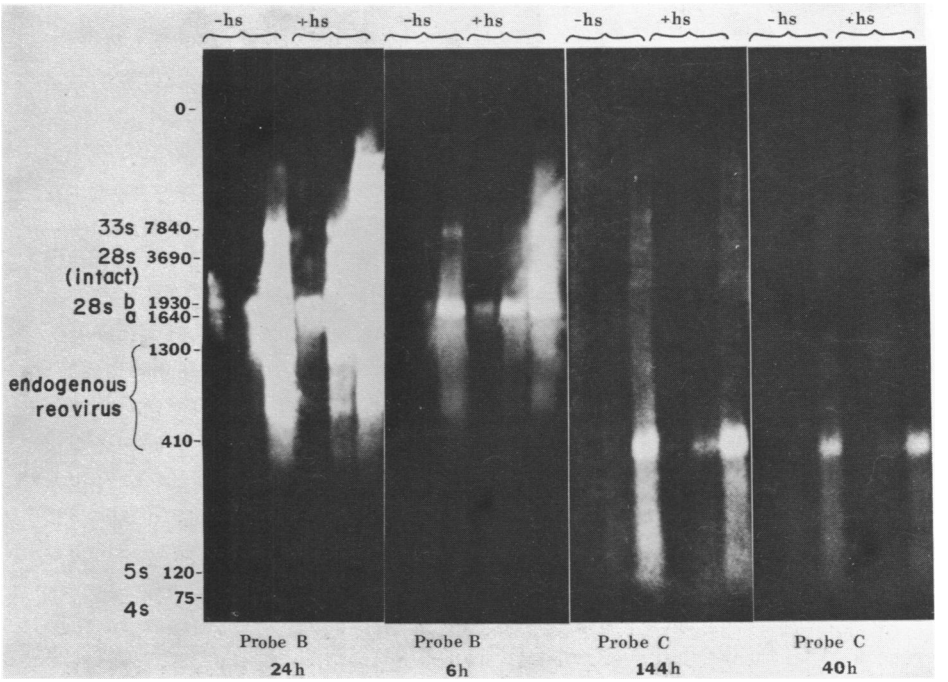

Figure 2. RNA species homologous to probes B and C. Nuclear RNA from Kc cells that had (+hs) and had not (-hs) experienced a heat shock treatment was denatured and electrophoresed on 1.5% agarose gels: three channels in each experiment were loaded with 0.25, 2.5 and 25 µl, respectively, of an RNA sample that contained about 100 µg/ml. Duplicate half-gels of this format were transferred to diazotized cellulose [25] and hybridized respectively to ^{32}P-labelled probe B (mainly 28S sequences) and probe C (only *intron* sequences). Autoradiographic exposures with DuPont Lightning-Plus screens were made on Kodak XR-5 at -70°C for 6 hours, 24 hours, (*exon*) 36 hours and 144 hours (*intron*) then developed with X-ray developer. The lengths of markers are shown in bases and their positions were identified by electrophoresing samples of ^{3}H labelled 28S, 18S and total RNA in adjacent channels on the same gel, then detecting these bands by fluorography [40]. The markers at 1300 and 410 base are from the total RNA sample and their lengths were determined independently by their mobilities on urea gels using 28, 18, 5 and 4S markers. These were used as standards in mobility versus log (molecular length) plots to estimate molecular lengths of observed RNA species. The major bands hybridizing to probe B were at 7800 and 1900 bases, with weak or fuzzy bands at 3700, 600 and 315 bases. The major band hybridizing to probe C was at 325 bases with weak or fuzzy bands at greater or equal to 10,000 bases, and at 9000, 4350 and 2300 bases. The estimated errors range from 10 to 25% depending on the proximity of markers.

either transcribed rarely or broken down extremely rapidly.

Berk-Sharp experiments detecting *intron* transcripts. In the previous section it was shown that *intron* transcript fragments exist, but that the majority of these are small and are not part of large molecules in which rRNA and *intron* sequences have been co-transcribed. However, some hybridization did occur to larger molecules which might be long precursor transcripts containing *introns*. To confirm the existence of *intron* transcripts and to look for such large precursors, a modification of the types of experiments first performed by Berk and Sharp [24] was devised. Plasmid cDm103, cut with *Eco*R1 and denatured, was incubated with nRNA under conditions where DNA-RNA hybrids but not DNA-DNA hybrids form [46]. Such hybridized samples were digested with S1 nuclease, so that the only remaining intact DNA is that class of sequences which was totally matched to RNA molecules. These DNA sequences were then electrophoresed on alkaline gels [41], transferred to nitrocellulose [42] and detected by hybridization to labelled probes for all or part of the cDm103 sequence. The important point is that the DNA fragments must totally and contiguously match the RNA molecules that saved them; no looping-out or mismatching is allowed since S1 will cut the DNA at such sites. The implication then for fragments that are detected by hybridization to labelled probe C (*intron*) is that if they are longer than 5 kb, there is no way to fit them into the map of cDm103 (see Figure 1), without overlapping onto neighboring 28S b_1 or b_2 sequences. Thus fragments longer than 5 kb detected with probe C imply that some 28S sequences are co-transcribed with *intron* sequences from this type of rDNA repeat. Transcripts detected from repeats with shorter *introns* will appear as shorter fragments due to mismatch with cDm103. Similar arguments and implications apply to other probes also. An experiment of this type detecting with probe C is shown in Fig. 3. As before both +hs and -hs RNA were used and *E.coli* tRNA was used as a control. The extent of digestion in the tRNA, +hs and -hs experiments was not totally uniform for unknown reasons; however, the experiments show completely consistent results over a wide range of digestions. No bands at all were seen in the tRNA experiment. In the -hs and +hs experiment, however, a number of bands were observed. The major species occurred at 1,790 bases, but another seven bands at least were observed, including three well over 5000 bases at 6,100, 7,300 and 10,000 bases (see Table I). The band at 6,100 was seen only in the -hs RNA. Thus, by the logic discussed above, some 28S sequences are co-transcribed with

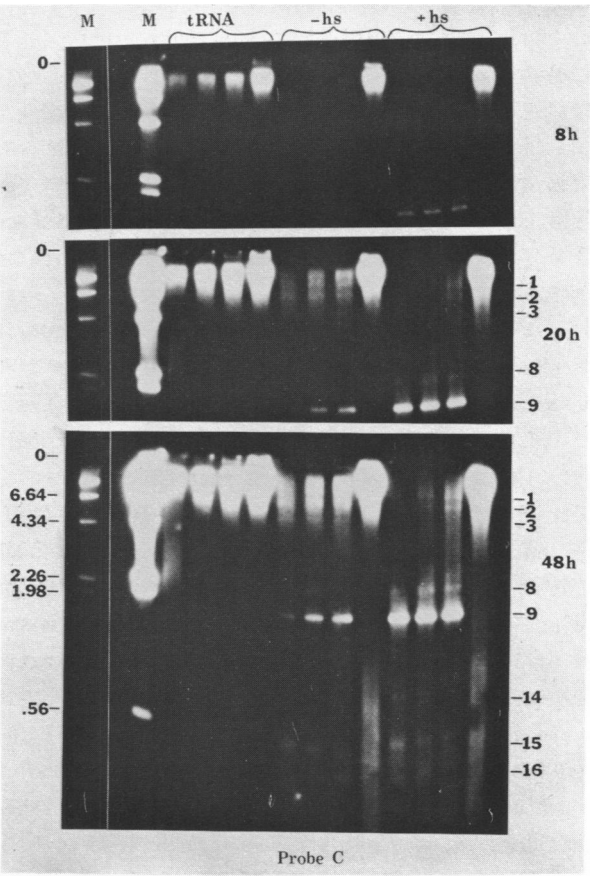

Figure 3. Berk-Sharp experiments with probe C, the *intron* fragment. This figure gives three different autoradiographic exposures of the same experiment exposed for the times shown. These show the electrophoretic patterns of those denatured DNA chains, homologous to probe C, that were saved from S1 digestion by hybridization to nuclear RNA. An *Eco*R1 digest of cDm103 was hybridized with three separate samples of RNA in 80% formamide - conditions that destabilize DNA:DNA hybrids yet allow DNA:RNA hybrids to form [46]: *E. coli* tRNA, nuclear RNA from Kc cells that received no heat shock (-hs) and the same RNA from heat shocked cells (+hs). Each sample was treated with S1 endonuclease for 0, 3, 10 and 30 minutes, the resulting aliquots denatured in 50 mM NaOH and then electrophoresed in agarose gels containing 30 mM NaOH [41]. The gel was neutralized, transferred to nitrocellulose filters as described by Southern [42], and hybridized to ^{32}P-probe C. The markers in the two left most lanes (M) with the lengths shown in kb.p. are different autoradiographic exposures of cold λ/*Hin*dIII fragments hybridized with ^{32}P labelled λ DNA.

intron sequences. However, these results are not immediately compatible with the results from the RNA blotting experiments. There the major species was 325 bases, here that is a minor component. Probable explanations for this are discussed later.

Berk-Sharp experiments with other parts of the rDNA repeat. The results in the above experiment suggested that a fairly complex set of transcripts of the *intron* existed, some of which included rRNA sequences. To investigate this further, the experiment was repeated, with the modification that different probes were used in the final hybridization-detection step. These are the probes described in Fig. 1. The results of such experiments are shown in Fig. 4. A catalogue of the bands consistently observed, number 1 through 16, is given in Table I. The pattern appears complex but is generally the same with +hs and -hs nRNA. A number of points emerge. Firstly, there is ample evidence for long transcripts that cover the whole cDm103 sequence from the 18S portion through 28S a and b_1, the *intron* and 28S b_2 (band 1, probe A, B, C, D). Secondly, all fragments that might be expected from known RNA precursors from 33S downwards [29,44] are seen (see Fig. 5). Thirdly, there appears to be some transcription off the "nontranscribed spacer" (band 4, probe E). Fourthly there are some good candidates for processing intermediates including *intron* and *exon* sequences, notably the 6.1 and 1.79 kb bands detected with probes B and C (bands 3 and 9). The 1.79 kb band (no. 9) is a major species detected with probe B, yet does not fit into any previously proposed precursor processing schemes. Fifthly, there are a number of species detected that do not fit into any obvious transcription-processing scheme, e.g., band 2 (probe C) and many of the shorter bands (e.g., bands 13 and 14, probe B). The species detected and their approximate locations with respect to the rDNA repeat are summarized in Fig. 5.

DISCUSSION

Previous investigations of rRNA transcription and processing have used pulse labelling techniques [29,48]. These experiments detected as the earliest precursor a 33-34S molecule which was identified as the ETS, 18S ITS and 28S sequences (see Fig. 1) on one molecule. No evidence of longer precursors (i.e., including *intron* sequences) was found. The 33S molecule was thought to be cut at the junction of the 18S and ETS sequences and then again at the 18-ITS junction. This gave mature 18S plus a pre-28S molecule which was subsequently processed to give 28S a and b,

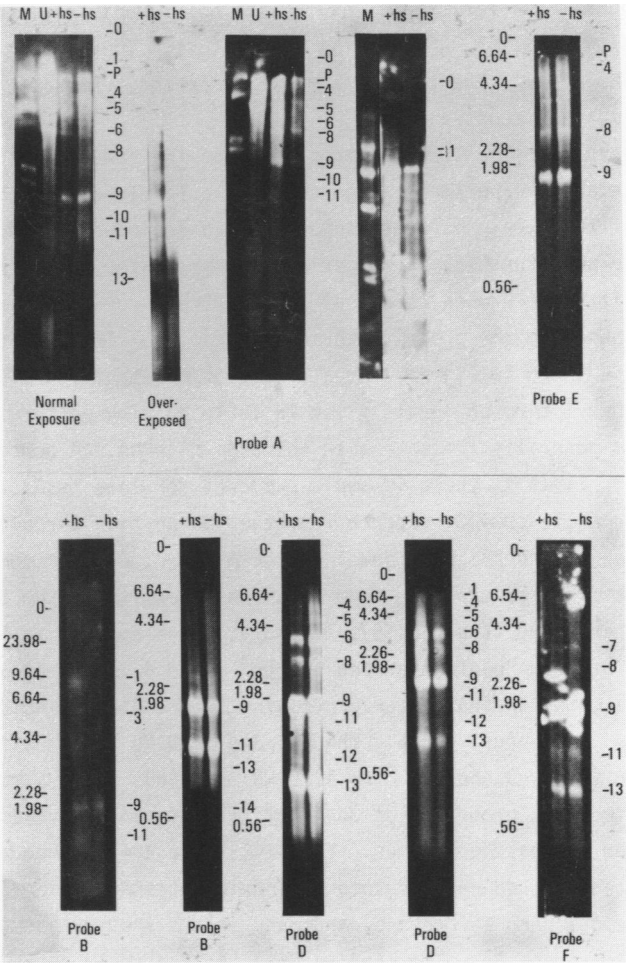

Figure 4. Berk-Sharp experiments with probes A, B, D, E and F. These autoradiographs show the electrophoretic patterns of those denatured DNA chains homologous to the 5 different probes mentioned. These chains were saved from S1 digestion by hybridization to nuclear RNA from Kc cells that had (+hs) and had not (-hs) experienced heat shock. The experiments were performed as described in Figure 3, using a standard S1 treatment of 5'. The nitrocellulose filters were hybridized with probe A (3 gels), D (2 gels), B (2 gels) E and F. All gels contained λ/*Hind*III markers that are shown in some autoradiographs in channel M; in others thay are not shown but the fragment lengths in kb are marked to the left. Two gels probed with A contained samples without S1 digestion (U). The numbers (1-14) refer to observable bands that are summarized in Table 1. The letter "p" refers to the band, detectable with probes A and E, which corresponds to the renatured (or un-denatured) *col*E1 vehicle. It appears in channels without S1 digestion and in samples incubated with tRNA.

	Probes					
Band	A	B	C	D	E	F
1	10,000 i	10,000 i	10,000 i	10,000 w	-	-
2	-	-	7,320 i	-	-	-
3	-	6,080 i	6,100 w	-	-	-
4	5,000 w	-	-	5,000 i	5,100 w	-
5	4,150 w	-	-	4,135 i	-	-
6	3,100 i	-	-	3,340 i	-	3,400*w
7	-	-	-	-	-	2,900*w
8	2,424 i	-	2,416 i	2,670 s	2,650 i	2,690*w
8	1,740 s	1,790 s	1,790 s	1,750 s	1,730 s	1,650 s
10	1,410 i	-	-	-	-	-
11	1,100 i	1,234 s	-	1,200 w	-	1,160 s
12	-	-	-	1,000 w	-	-
13	840 i	900 w	-	800 s	-	860 s
14	-	645*i	645 i	-	-	-
15	-	-	467 i	-	-	455 w
16	-	270*w	340*w	270*w	-	265*w

Table 1. Summary of molecular length measurements from experiments in Fig. 3 and 4. Sixteen significantly different molecular length classes were observed. Bands are marked "s", "i", or "w" to indicate strong, intermediate and weak. All fragments have been confirmed in two or more experiments except those marked *. In these cases bands were clearly seen in one experiment and not contradicted by others. The estimated error in length is ± 10%, except for band 1, where the non-linearity in the mobility-molecular weight plots may double this. The lengths in nucleotides, calculated using denatured λ/*Hind*III fragments as standards, are listed under the probe by which they were detected.

5.8S, and 2S RNA molecules. All or most of this occurred in the nucleus. However, pulse chase experiments are much less sensitive than the use of radiolabelled cloned hybridization probes described here. Thus, in the Berk-Sharp experiments, virtually all the fragments expected from the pulse-chase results have been detected (see Fig. 5), suggesting that the outline of transcription and processing based on these results is correct, as far as it goes. In addition, RNA species corresponding to a complete read-through of 18S, 28S and *intron* sequences have also been observed at the level of not greater than 1 to 2 copies per cell (see below). Some possible *intron* processing intermediates also occur at the same low

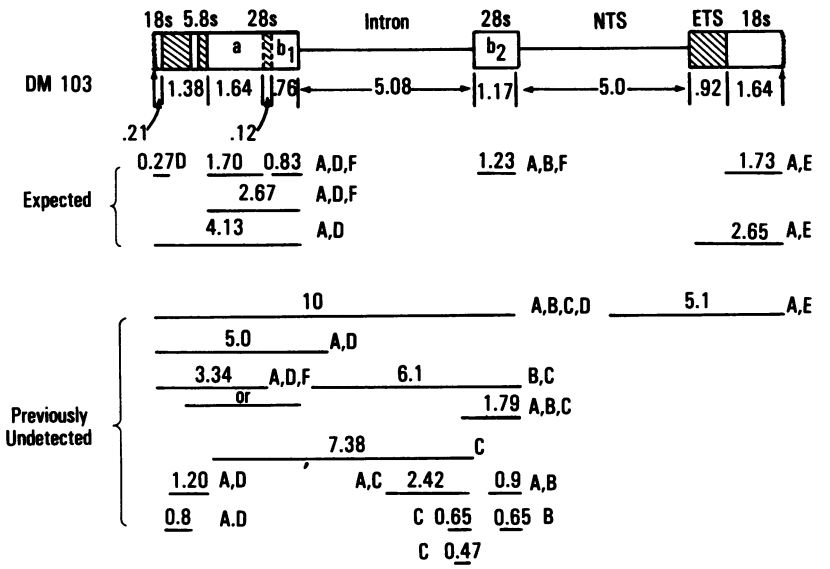

Figure 5. Summary of results. The map of cDm103 as in Fig. 1 is shown with the lengths of the various regions in kb [4,17]. The DNA fragments that are protected from S1 by nuclear transcripts are shown as lines in their probable location defined by the probes by which they were detected (A-F). The number above is their length in kb.

levels. A possible processing intermediate that contains both *intron* and *exon* sequences is the RNA species that saves (from S1) a DNA chain 6100 bases long that hybridizes to both probes B and C. By assigning the position it has in Fig. 5, this chain ends just at the 28S b_1 *intron* junction, just where a cut might be expected. Of course, it cannot be said with absolute certainty that the two 6100 base sequences detected by probes B and C are one and the same, nor that its position is as exactly shown. However, it is the simplest and most reasonable explanation, especially since there are no previous examples of processing of the 3' end (the 28S b_2 end) of any previous rRNA precursors [29].

The situation is further complicated by the existence of a considerable number of rDNA repeats with type I *introns* of lengths other than 5000 bases. These occur in lengths which are approximate multiples of 500 bases, from 500 to 6000 bases; however, there is no internal sequence repetition [16,17]. The 500 base segments appear to be "removed" progressively from the left or b_1 end of the *intron*, so that all type I *introns*

have the 500 bases next to the 28S b_2, but only those with 5000 base or longer *introns* have the 500 bases next to the 28S b_1 sequence in these repeats. Thus the 1.79 kb fragment which is detected by probes B and C is interesting since a likely location of this is covering the entire 28S b_2 sequence and extending into the *intron* for 500 bases. This may represent transcripts off repeats with short *introns* or mean that the 500 base multiples have a significance in transcript processing. This fragment is of further interest because it seems to be a significant part of the sequences detected with probe B (band 9) although this may be somewhat misleading, as discussed below. Other unexpected fragments found include a number of small fragments which appear to represent bits of 28S sequence (notably 28S b_2). Their significance is not at all clear. It should be noted that probes A through E, though not F, are not strand specific and could conceivably in some cases be detecting transcripts off the DNA strand opposite to the one from which rRNA is transcribed. In addition it is possible to imagine configurations of DNA-RNA hybrids for which the specificity of S1 has not been exactly described, in particular a DNA single strand paired with two RNA molecules which are contiguous. An example could be at the *intron* - 28S b_2 junction. However, S1 is a DNase as well as an RNase and cuts DNA opposite to nicks in DNA. In addition, Berk and Sharp [24] reported that S1 cuts R-loops at the DNA-RNA branch point and that they therefore assume that branching discontinuities in any duplex structure will be a site for S1 attack.

There is also some information concerning the processing of the ITS, 5.8S, 28S a and 28S b sequences in the Berk-Sharp experiments. The 1.2 and 0.8 kb fragments detected by probes A and D probably mean that, after the 18S sequence is removed from the 33S precursor, the ITS including the 5.8S sequence is removed as a 1.2 kb piece, then 0.8 kb is removed from the left end to expose one end of the 5.8S molecule. Presumably the remaining 200 or so bases are cleaved from the other side. In addition, the fact that 4.13 kb, 2.67 kb and 3.34 kb precursors can all be obtained from the region suggests that there may well be some variation in the order in which some processing cuts are made. Such variation could well be tolerated if the processed parts remain hydrogen bonded together after covalent bond breakage. A further unexpected outcome from these experiments is that there appears to be transcription off the "nontranscribed spacer" (probe E, band 4). Such a result is not unprecedented [49] and may be the result of read-through from preceding repeats. The question

now arises - at what level do the RNA molecules corresponding to the DNA fragments detected exist. The Berk-Sharp experiments are a rather indirect measurement of this, and as discussed below, there is reason to believe that they have some built-in biases. Thus the best estimate of the abundance of the *intron* transcripts comes from the RNA blotting experiments. The autoradiographic scans gave a ratio of probe B (*exon*) to probe C (*intron*) sequences of 3000:1. One can estimate that there are 5000 to 50,000 rRNA molecules (including 33S and 28S) per nucleus [51]. Taking the larger number as the number of probe B (28S b_2) sequences detected, then there are 10-20 *intron* sequence molecules per nucleus. However, at least 90% of these are the 325 base fragment detected so that any precursor molecules with *intron* and *exon* sequences occur at the level of 1-2 molecules/nucleus. However, this calculation brings out some difficulties in comparing the RNA blot and Berk-Sharp experiments for the *intron*. If the 325 base RNA molecule is the majority species of *intron* RNA, why is it not more prominent in the Berk-Sharp experiments? A number of explanations are possible. The 325 base RNA species detected with probe B may in fact be somewhat mismatched, and such hybrids would then be destroyed by S1. Alternatively if there are a number of RNA molecules of different lengths in excess competing for cDm103 DNA, then it is easy to see that, due to strand migration short molecules will eventually be competed off and "locked out" of the DNA by longer molecules emcompassing the short molecule sequence. Such effects are probably rather important in detecting the long transcripts that are undoubtedly present in very low levels, and suggest that the Berk-Sharp technique is not at all quantitative in such situations.

Finally, it should be noted that experiments rather like these have been attempted for *D. melanogaster* embryo RNA [50]. RNA driven liquid hybridization experiments failed to detect *intron* transcripts with a sensitivity of 10 molecules/cell. As discussed above, the experiments in this paper were probably more sensitive, but the intriguing possibility remains that the embryonic nRNA in those experiments did not include *intron* sequences. Although the tissue culture line used here was embryo derived [26], as a permanent cell line it is obviously differentiated from primary embryo cells; indeed it has some neuron-like properties [52]. Thus it may be that the existence of *intron* transcripts is linked to differentiation. It would be of interest to investigate this by examining nRNA from embryos, larvae and adults or from different tissues or imaginal discs.

ACKNOWLEDGMENTS

This work was supported by a fellowship from the Charles A. King Trust (to DJJ) and the NIH GM 25531. We thank Dr. G. Rubin for the strain bearing cDm103, Dr. S. Smith for *Bam*HI and Dr. J. Beard for AMV reverse transcriptase. We benefitted from many conversations with Kaoru Saigo and thank Carton Chen for his photographic assistance.

*Present address: Department of Pediatrics, University of California at San Diego, La Jolla, CA 92093, USA.

REFERENCES

1. Gilbert, W. (1978) Nature 271, 501.
2. Crick, F.C. (1979) Science 204, 264-271.
3. Pellegrini, M., Manning, J. and Davidson,N. (1977) Cell 10, 213-224.
4. Glover, D.M. and Hogness, D.S. (1977) Cell 10, 167-176.
5. Weelauer, P.K. and Dawid, I.B. (1977) Cell 10, 193-212.
6. Tilghman, S.M., Tiemeier, D.C., Seidman, J.G., Peterlin, B.M., Sullivan, M., Maizel, J.V. and Leder, P. (1978) Proc. Natl. Acad. Sci. USA 75, 725-729.
7. Jeffreys, A.J. and Flavell, R.A. (1977) Cell 12, 1097-1108.
8. Dugaiczyk, A., Woo, S.L.C., Lai, E.C., Mace, M.L., Jr., McReynolds, L. and O'Malley, B.W. (1978) Nature 274, 327-328.
9. Breatnach, R., Mandel, J.L. and Chambon, P. (1977) Nature 270, 314-319
10. Knapp, G., Beckmann, J.S., Johnson, P.F., Fuhrman, S.A. and Abelson, J. (1978) Cell 14, 221-236.
11. Wild, M.A. and Gall, J.C. (1979) Cell 16, 565-573.
12. Campbell, G.R., Littau, V.C., Melera, P.W., Alfrey, V.G. and Johnson, E.M. (1979) Nucl. Acid Res. 6, 1433-1447.
13. Sanders, J.P.M., Heyting, C., Verbect, M.P., Meijlink, F.C.P. and Borst, P. (1977) Molec. Gen. Genet. 157, 239-262.
14. Garapin, A.C., Lepennec, J.P., Roskam, W., Perrin, F., Cami, B., Krust, A., Breatnach, R.,Chambon, P. and Kourilsky, P. (1978) Nature 273, 349-358.
15. Miller, H.I., Konkel, D.A. and Leder, P. (1978) Nature 275, 772-774.
16. Wellauer, P.K., Dawid, I.B. and Tartof, K.D. (1978) Cell 14, 269-278.
17. Wellauer, P.K. and Dawid, I.B. (1978) J. Mol. Biol. 126, 769-782.
18. Tartof, K.D. and Dawid, I.B. (1976) Nature 263, 27-30.
19. Dawid, I.B. and Botchan, P. (1977) Proc. Natl. Acad. Sci. USA 74, 4233-4237.
20. Wellauer, P.K., Dawid, I.B., Brown, D.D. and Reeder, R.H. (1976) J. Mol. Biol. 105, 461-468.
21. Lifton, R.P., Goldberg, M.L., Karp, R.W. and Hogness, D.S. (1977) Cold Spring Harb. Symp. Quant. Biol. 42, 1047-1052.
22. Finnegan, D.J., Rubin, G.M., Young, M.W. and Hogness, D.S. (1977) Cold Spring Harb. Symp.Quant. Biol. 42, 1053-1063.
23. Craig, E.A., McCarthy, B.J. and Wadsworth, S.C. (1979) Cell 16, 575-588.
24. Berk, A.J. and Sharp, P.A. (1977) Cell 12, 721-732.
25. Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5350-5354.
26. Echalier, G. and Ohanessian, A. (1970). In Vitro 6, 162-172.
27. Spear, B.B. (1974) Chromosoma 48, 159-179.

28. McKenzie, S.L., Heinikoff, S. and Meselson, M. (1975) Proc. Natl. Acad Sci. USA 72, 1117-1121.
29. Levis, R. and Penman, S. (1978) J. Mol. Biol. 121, 219-238.
30. Levis, R. and Penman, S. (1977) Cell 11, 105-113.
31. Clewell, D.B. and Helinski, D.R. (1969) Proc. Natl. Acad. Sci.USA 62, 1159-1166.
32. Bingham, A.H.A., Sharman, A.F. and Atkinson, T. (1977) FEBS Lett. 76, 250-256.
33. Hayward, G.S. and Smith, M.G. (1972) J. Mol. Biol. 63, 383-395.
34. Chen, C.W. and Thomas, C.A., Jr., submitted to Anal. Biochem.
35. Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) Proc. Natl. Acad. Sci. USA 72, 1184-1188.
36. Taylor, J.M., Illmensee, R. and Summers, J. (1976) Biochim. Biophys. Acta 442, 324-330.
37. Rougeon, F. and Mach, B. (1976) Proc. Natl. Acad. Sci. USA 73, 3418-3422.
38. Wickens, M.P., Buell, G.N. and Schimke, R.T. (1978). J. Biol. Chem. 253, 2483-2495.
39. Peacock, A.C. and Dingman, C.W. (1968). Biochemistry 7, 668-674.
40. Bonner, W.M. and Laskey, R.A. (1974) Eur. J. Biochem. 46, 83-88.
41. McDonnell, M.W., Simon, M.N. and Studier, F.W. (1977) J. Mol. Biol. 110, 119-146.
42. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
43. Denhardt, D.T. (1966) Biochem. Biophys. Res. Commun.23, 641-646.
44. Dawid, I.B., Wellauer, P.K. and Long, E.O. (1978). J. Mol. Biol. 126, 749-768.
45. Mirault, M-E., Goldschmidt-Clermont, M., Arrigo, A.P. and Tissieres, A. (1977) Cold Spring Harb. Symp. Quant. Biol. 42, 819-827.
46. Casey, J. and Davidson, N. (1977) Nucl. Acid. Res. 4, 1539-1552.
47. Carlson, J., Ott, G. and Sauerbier, W. (1977) J. Mol. Biol. 112, 353-357.
48. Jordan, B.R., Jourdan, R. and Jacq, B. (1976) J. Mol. Biol. 101, 85-105.
49. Rungger, D., Crippa, M., Trendelenburg, M.F., Scheer, V. and Franke, W.W. (1978) Exp. Cell Res. 116, 481-486.
50. Long, E.O., Dawid, I.B. and Wellauer, P.K. (1978). Yearbook of the Carnegie Institution of Washington, Dept. of Embryology 77, 86-93.
51. Clark, S.H., Strausbaugh, L.D. and Kieffer, B.I. (1977). Genetics 86, 789-800.
52. Berger, E., Ringler, R., Alahiotis, S. and Frank. M. (1978) Develop. Biol. 62, 498-511.