Volume 16 Number 10 1988 Nucleic Acids Research

Analysis of the *Drosophila* rDNA promoter by transient expression

David C.Hayward+ and David M.Glover*

Cancer Research Campaign, Eukaryotic Molecular Genetics Research Group, Department of Biochemistry, Imperial College of Science and Technology, London, SW7 2AZ, UK

Received March 3, 1988; Revised and Accepted April 22, 1988

ABSTRACT

We have examined the expression of the bacterial gene chloramphenicol acetyl transferase (CAT) under the control of the *Drosophila* rDNA promoter following transfection into *Drosophila* tissue culture cells. Constructs having an entire NTS, corresponding to approximately 3640 base pairs of upstream rDNA sequence, or constructs with 306 base pairs of upstream sequence respectively, are transcribed at 5 fold or 2 fold higher levels than a construct with 43 base pairs of upstream DNA. In co-transfection experiments, the construct with the entire NTS competes for transcription 20 fold more effectively than the construct with 306 base pairs of upstream sequence. Constructs having either 72 base pairs or 60 base pairs of upstream rDNA sequences, on the other hand, are transcribed very much less efficiently than constructs with either 306 bp or with only 43 bp of upstream DNA. These sequences, which reduce levels of rDNA transcription in the absence of additional upstream DNA, lie in a region in which the rDNA promoter differs from its duplications within the NTS.

INTRODUCTION

Three types of repeat sequence have been described within the "non-transcribed" spacer (NTS) sequences of *Drosophila melanogaster* rDNA. These are approximately 240 bp, 330 bp, and 95 bp in length, the 330 bp repeats being comprised of elements of the 240 bp and 95 bp repeats (1,2). The repeating 240 bp unit contains sequences homologous to a region around the major transcription start site for the 40S rRNA. Kohorn and Rae (3,4) have determined that the sequence between 43 base pairs upstream and 4 base pairs downstream of the major transcription start site is the minimum required for efficient transcription *in vitro*. The reiterations of part of this sequence in the 240 bp repeats of the NTS can serve to initiate transcription both in the cell free system derived from Kc cell extracts (5) and also *in vivo* (6,7).

Transcription proceeds through the 18S and 28S genes to produce a rRNA precursor molecule of approximately 8kb. This molecule undergoes endonucleolytic processing to produce the mature 18S, 5.8S, 2S, and 28S rRNAs. Nuclease S1 analysis of total RNA and "nuclear run-on" assays have shown that there is no defined termination site for transcription in the *Drosophila* NTS (8) implying that transcription continues through the "non-transcribed" spacer

up to the promoter of the next rDNA unit. The 3' end of the rRNA precursor precursor is thus now believed to be an RNA processing site.

An alternative system has been used by Grimaldi and DiNocera (9) to examine the upstream requirements for rDNA expression. They have transfected rDNA into *Drosophila* tissue culture cells, and have shown that the rDNA promoter will drive the transient expression of the bacterial gene for chloramphenicol acetyl transferase (CAT) when introduced into Schneider line II (SL2) cells. They showed that a plasmid having the entire NTS would effectively compete for transcription with one having only 180 bp of upstream sequences. In a recent extension of this work, these authors have shown that the activity of the rDNA promoter correlates with the numbers of 240 bp repeats (10). These results are broadly in agreement with data from competition experiments examining the template activity of *Xenopus* rDNA having varying amounts of upstream sequences (11-13). In this case, the DNA of the NTS several kb upstream of the transcription initiation site is found to enhance expression from the rDNA promoter. The sequences responsible for enhanced transcription are the 60/81 bp repeats which contain homology with the region from -114 to -72 with respect to the transcription start site; and the "Bam island" reiterations of the transcription start site.

In this paper, we use the approach of Grimaldi and DiNocera (9) to re-examine the requirement for the sequences immediately upstream of the transcription start site for transcription *in vivo*. We find that whereas a construct having sequences up to 43 base pairs upstream of the transcriptional start site efficiently directs the expression of CAT, constructs with an additional 17 or 29 base pairs (between positions -43 and -72) express CAT inefficiently. Efficient expression is restored by the addition of longer upstream sequences.

MATERIALS AND METHODS

Construction of rDNA-CAT plasmids

Plasmid clones containing the rRNA gene transcription initiation site with 43, 60 and 72 bp of upstream sequence were a generous gift from Dr. Bruce Kohorn. These plasmids were derived from pDmr275c2, and their construction is described in detail by Kohorn and Rae (4). These plasmids were digested with *Eco*RI and *Bam*HI and the rDNA fragments gel purified, digested with *Taq*I, which cuts once at a position 32 bp downstream from the start site of transcription, and ligated into pEMBL8+ which had been cut with *Eco*RI and *Acc*I. Recombinant colonies containing the *Eco*RI/*Taq*I fragments, i.e. the transcription initiation site, were identified by colour selection, hybridisation with labelled NTS sequences, and by sequence analysis of the single stranded form of the plasmids. These plasmids thus contain 35 bp of sequence downstream of the transcription start site and either 72, 60 or 43 bp of upstream sequence and were named pDm-72, pDm-60 and pDm-43 respectively (Fig. 1).

The 4.7 kb *Hin*dIII/*Bgl*II fragment containing the NTS, the transcription initiation site, and part of the 18S gene, was purified from an *Eco*RI/*Hin*dIII/*Bgl*II digest of pDm238 (14). This fragment was then cloned

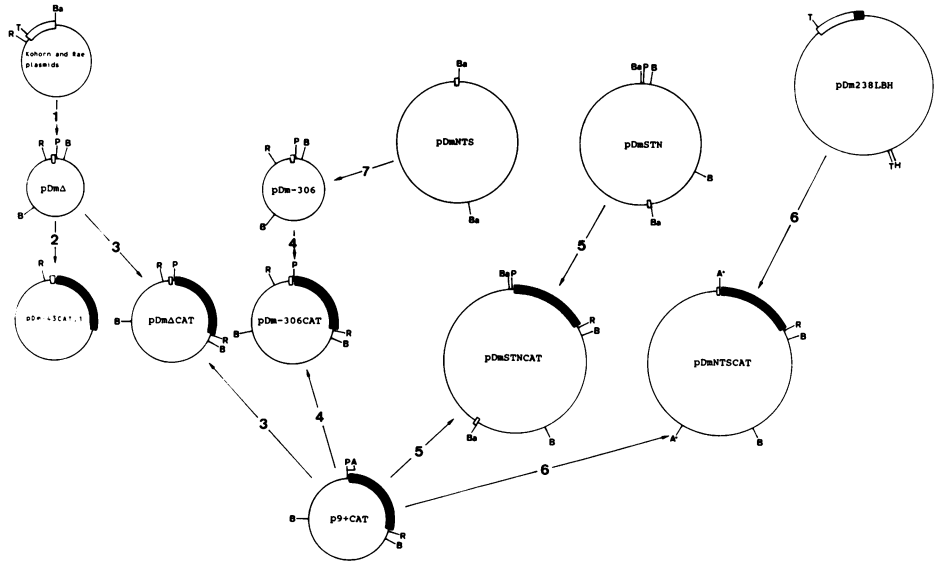

Figure 1. The Construction of the rDNA-CAT Plasmids. Closed blocks represent the sequences from pSV0*Bgl*II containing the CAT gene, except in the case of pDm238LBH where the closed block represents the 5' end of the 18S rRNA gene. Open blocks represent external transcribed spacer sequences of rDNA. Restriction endonuclease cleavage sites are :- R, *Eco*RI; B, *Bgl*II; H, *Hin*dIII; P, *Pst*I; Ba, *Bam*HI; T, *Taq*I; A, *Acc*I. The symbol A* represents the *Acc*I site of p9+CAT which was destroyed when the *Taq*I fragment from pDm238LBH was ligated into it. The stages of plasmid construction were as follows (see text for greater detail): 1. The *Eco*RI/*Bam*HI fragments from the Kohorn and Rae plasmids were digested with *Taq*I and the small fragment inserted between the *Eco*RI and *Acc*I sites of pEMBL8+. 2. The end repaired *Bgl*II/*Bam*HI fragment containing the CAT gene from pSV0*Bgl*II was ligated into the end repaired *Pst*I site of pDm-43T to give pDm-43CAT.1. 3. The *Pst*I/*Bgl*I fragment containing the CAT gene from p9+CAT was ligated with the two large fragments from *Pst*I/*Bgl*I digests of the plasmids pDm-72, pDm-60 and pDm-43 to give the corresponding rDNA-CAT plasmids. 4 & 5. In a similar way to that described in 3, the two large fragments from *Pst*I/*Bgl*I digests of pDm-306 and pDmSTN were ligated to the *Pst*I/*Bgl*I fragment containing the CAT gene from p9+CAT to give pDm-306CAT and pDmSTNCAT. 6. The *Taq*I fragment containing the NTS and the transcription start site from pDm238LBH was ligated into the *Acc*I site of p9+CAT to give pDmNTSCAT. 7. The *Alu*I/*Bam*HI fragment containing the transcription initiation site from pDmNTS was cloned into the *Bam*HI site of M13mp8. The insert was subsequently cloned into pEMBL8+ to give pDm-306.

between the *Hin*dIII and *Bam*HI sites of pEMBL8+ to give pDm238LBH. This plasmid was digested with *Eco*RI and *Hin*dIII and the insert was gel purified. After digestion with *Taq*I, the cohesive ends were filled-in using the Klenow fragment of *E.coli* DNA polymerase I and the fragments were inserted into the filled-in *Bam*HI site of pEMBL8+. This process recreates the *Bam*HI sites. Plasmids which contained the large (3.7 kb) *Taq*I fragment, consisting of the

entire NTS and 34bp of the ETS, were identified by hybridisation with nick-translated NTS sequences and by sequencing analysis. Both orientations of insert were obtained and the plasmids were called pDmNTS and pDmSNT (Fig. 1)

The 340 bp fragment in pDm-306 was originally cloned into M13mp8 after gel purification from a *Bam*HI/*Alu*I digest of pDmNTS. The *Bam*HI cohesive end was filled-in using Klenow polymerase and the fragment was cloned into the filled-in *Bam*HI site of mp8. The insert was subsequently excised from the double stranded RF with *Eco*RI and *Hin*dIII and cloned between the *Eco*RI and *Hin*dIII sites of pEMBL8+ to give pDm-306. The orientation of the insert is such that the *Taq*I site is nearest the *Pst*I site of the pEMBL8+ polylinker (Fig. 1).

The plasmid p9+CAT contains the bacterial chloramphenicol acetyl transferase (CAT) gene in pEMBL9+. It was constructed by gel purifying the CAT gene containing the 1.6kb *Bam*HI/*Bgl*II fragment from pSV0CATBglII and inserting it into the *Bam*HI site of pEMBL9+. The orientation of the insert is such that the 5' end of the CAT gene is nearest the *Hin*dIII site of the pEMBL9+ polylinker.

The clone pDm-43CAT.1 was constructed by first digesting pDm-43T with *Pst*I which cuts once just after the ETS end of the insert. Treatment with T4 DNA polymerase in the presence of dCTP had the effect of removing the 3' overhangs created by *Pst*I. The plasmid pSV0CATBglII was digested with *Bam*HI and *Bgl*II to release the fragment containing the CAT gene and the sticky ends were filled using the Klenow fragment of *E.coli* DNA polymerase I. A blunt ended ligation of the gel purified CAT gene fragment and the *Pst*I cut pDm-43T was carried out. Recombinants containing the CAT gene were selected on chloramphenicol and the orientation of the gene was determined by restriction endonuclease digestion and sequencing analysis.

In order to make the plasmids pDm-72CAT, pDm-60CAT and pDm-43CAT, the fragment containing the CAT gene was first purified from a *Pst*I/*Bgl*I digest of p9+CAT (*Bgl*I cuts approximately 150 bp away from the *Eco*RI site in the pEMBL9+ polylinker, in the beta-galactosidase gene sequence present in the vector). The plasmids were then also digested with *Pst*I and *Bgl*I. Since *Bgl*I also cuts the pEMBL plasmids in the beta-lactamase gene, three fragments were produced. The two large fragments were purified and ligated to the CAT gene fragment. Selection of colonies on both ampicillin and chloramphenicol ensured that only plasmids with the CAT gene and the rDNA fragments were recovered. The constructs pDmSTNCAT and pDm-306CAT were made in an analogous manner from pDmSNT and pDm-306. The plasmid pDmNTSCAT was constructed by cloning the purified NTS containing fragment from a *Taq*I digest of pDm238LBH and cloning it into the *Acc*I site of p9+CAT.

The organisation of these constructs was confirmed by restriction endonuclease mapping and sequencing analyses using a conventional single stranded primer (New England Biolabs 15-mer) and a primer complementary to a region of the CAT gene approximately 65 bp from the *Bgl*II site (15).

Transfection of Drosophila Tissue Culture Cells.

Plasmid DNA was introduced into *Drosophila* Kc and SL2 cells using a modification of the calcium phosphate/DNA co-precipitate method described by

Gorman (16). Cells were seeded, 24 h before transfection, at a density of approximately 3 X 10^6 cells per ml in a 75 cm^2 flask. The calcium phosphate/DNA co-precipitate, containing a total of 20ug DNA was added to the cells and left for 5-7 h, after which it was removed and replaced with fresh medium.

Chloramphenicol Acetyltransferase (CAT) Assay.

CAT assays were carried out on protein extracts from tissue culture cells as described by Davies *et al.*, (17).

Primer Directed Reverse Transcription.

Labelling of the synthetic oligonucleotide primer, and the hybridisation and extension reactions were carried out as described by Davies et al., (16) with the following modifications: 30ug of total cellular RNA was ethanol precipitated, centrifuged and resuspended in 6ul of double distilled water containing 10ng of oligonucleotide primer labelled to a specific activity of approximately 5 X 10^7 cpm/ug (see section 2.11.2.). After a 10 minute incubation at 65°C, 2ul of annealing buffer (500mM Tris-HCl pH 8.0, 500mM KCl, 80mM $MgCl_2$) were added. Annealing was carried out at 45°C for 2 h after which 9ul of 2mM dATP, 2mM dCTP, 2mM dGTP, 2mM TTP, 400mM DTT were added along with 13 units of AMV reverse transcriptase (Life Sciences). The reaction was incubated at 42°C for 1 h, phenol extracted and ethanol precipitated. After centrifugation the pellet was taken up in 10ul 60% (v/v) formamide dye mix (99% (v/v) formamide, 0.05% (w/v) bromophenol blue, 0,05% (w/v) xylene cyanol), boiled for 5 minutes and loaded onto a 6% polyacrylamide sequencing gel.

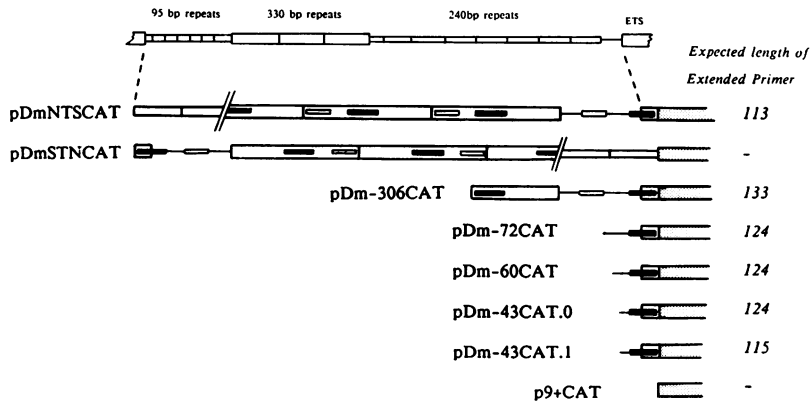

Figure 2. The Upstream Sequences in the rDNACAT constructs and the expected lengths of extended primers synthesised from RNA expressed from the constructs. The CAT gene is indicated by stippled shading. The two blocks of sequences within the 240 base pair repeats that have homology to sequences at the transcription initiation site are indicated by open and filled blocks (see also text and legend to Fig 8.)

RESULTS

Cells Transfected with the rDNA-CAT Plasmids do not Contain Significant Levels of CAT Activity.

We wished to study the cis-acting sequences that regulate *Drosophila* rDNA expression *in vivo* and chose to examine this by transfecting hybrid genes into *Drosophila* tissue culture cells. We chose to construct gene fusions in which the body of the bacterial gene encoding chloramphenicol acetyl transferase (CAT) was fused to the *Drosophila* rDNA promoter and varying amounts of upstream sequences. The construction of these molecules is described in detail in Materials and Methods (see also Fig. 1), and the essential features of the constructs shown in Fig. 2. The plasmids, pDm-43CAT, pDm-60CAT, pDm-72CAT, pDm-306CAT, and pDmNTSCAT contain 43, 60, 72, 306 and 3640 bp of sequences upstream of the rDNA initiation site respectively. The 3640 bp of rDNA in plasmid pDmNTSCAT correspond to the entire "non-transcribed" spacer (NTS) in the orientation expected to drive CAT transcription. pDmSTNCAT contains these same NTS sequences in the opposite orientation with respect to the CAT gene, and p9+CAT contains the CAT gene but with no upstream rDNA sequences whatsoever. CAT was a chosen as a "reporter gene", since there is no equivalent enzyme activity in *Drosophila* cells and it would therefore allow the expression of the transfected plasmid to be distinguished from endogenous rDNA expression. In principle, this could be done either by looking for CAT transcripts or for CAT enzyme activity, and when these experiments were initiated, there were reports that RNA transcripts from protein coding regions under the control of an RNA polymerase I promoter could be translated at low levels (17). We therefore transfected *Drosophila* SL2 cells with each of the rDNA-CAT plasmids and tested protein extracts of the cells for CAT enzyme activity after 48 hours (Fig. 3.). In no case could we detect significant acetylation of chloramphenicol in extracts of cells transfected with these plasmids (Fig. 3). This result is in agreement with the experiments of Grimaldi and DiNocera (9) who found that similar rDNA constructs would not express CAT activity following transfection into *Drosophila* cells. It contrasts with the extremely efficient expression of CAT activity in extracts from cells which had been transfected with the plasmid copia-CAT (18) which contains the CAT gene under the control of the copia promoter (Fig. 3., lane 9). The inability to detect high levels of CAT activity in extracts from cells transfected with the rDNA-CAT plasmids is not the result of a technical problem with transfections, since the experiment was repeated several times and the same results obtained. Moreover, several of the rDNA plasmids directed the synthesis of abundant CAT transcripts in the absence of CAT activity (see below). Furthermore, when the rDNA-CAT constructs were co-transfected into cells with copia-CAT, the high levels of expression characteristic of copia-CAT transfections were seen. In subsequent experiments, we have used the expression of CAT activity from the copia-CAT construct as a control for transfection efficiency. In these cases, the levels of CAT activity produced by the rDNA-CAT constructs is estimated to be less than 1% of that from copia-CAT.

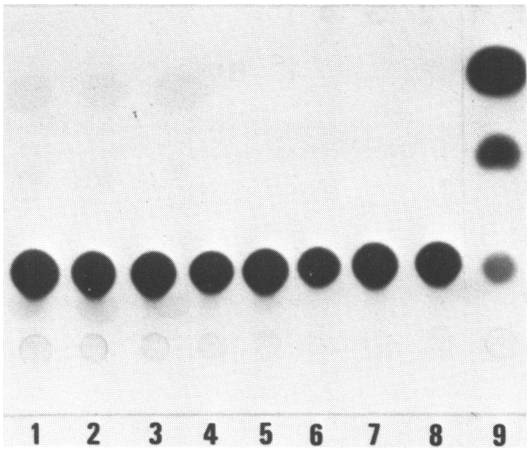

Figure 3. CAT activity in Cells Transfected with the rDNA-CAT Plasmids. Protein extracts were made from dishes of SL2 cells (1.5 x 10^7 cells) transfected with 5ug of rDNACAT plasmid. Equal amounts of protein (200ug) were tested for CAT activity. The lanes correspond to assays of extracts from cells transfected with: 1, pDmNTSCAT; 2, pDm-306CAT; 3, pDm-72CAT; 4, pDm-60CAT; 5, pDmSTNCAT; 6, p9+CAT; 7, pDm-43CAT.0; 8, pDm-43CAT.1. 9, copia-CAT.

pDmNTSCAT is Transcribed Accurately and Efficiently in SL2 Cells.

In order to determine whether our rDNA-CAT constructs were directing transcription *in vivo*, plasmid DNA was transfected into *Drosophila* SL2 cells and levels of CAT RNA determined by primer directed reverse transcription. This requires that RNA is first prepared from the transfected cells and annealed with a radio-labelled synthetic oligonucleotide primer complementary to a sequence close to the 5' end of the CAT gene (14). The primer is then extended by reverse transcriptase, in the presence of deoxyribonucleotides, to give a labelled molecule extending to the 5' end of the RNA. The predicted lengths of the extended primer complementary to transcripts from the various deletion plasmids are shown alongside the maps of these hybrid genes in Fig. 2.

Fig. 4. shows the results from one such experiment in which SL2 cells were co-transfected with copia-CAT and DNA of one of the plasmids, pDmSTNCAT, p9+CAT or pDmNTSCAT. A band can be seen corresponding to extended primer of the expected length complementary to transcripts of pDmNTSCAT initiating at the rDNA start site (Fig.4A lanes 1 and 4, solid triangle). A CAT transcript cannot be detected, however, in cells transfected with the plasmids pDmSTNCAT and p9+CAT (lanes 2 and 3). Panel B of Fig. 3. shows the results of CAT assays carried out on extracts made from these same cells. High levels of CAT activity are present in all of the extracts indicating expression of copia-CAT and thus showing that each of the transfections was successful. Efficient and specific transcription thus requires the rDNA promoter in the correct orientation.

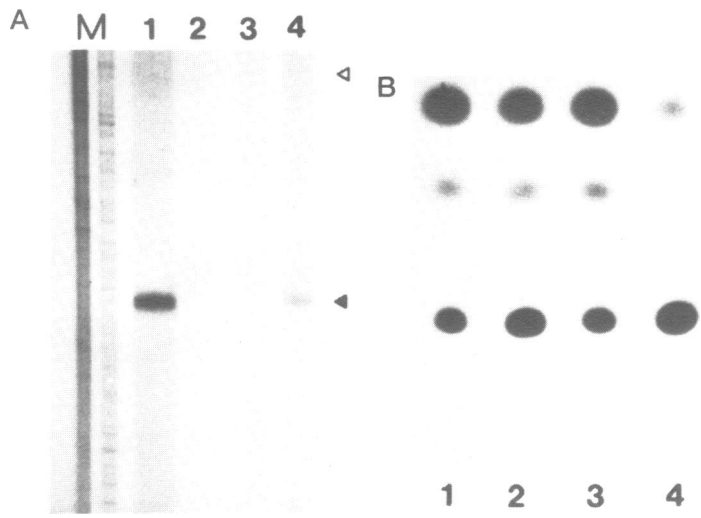

Figure 4. Transcription of pDmNTSCAT. A. Primer extension reactions carried out on 30ug total RNA from SL2 cells transfected with:- lane 1, pDmNTSCAT; lane 2, p9+CAT; lane 3, pDmSTNCAT; lane 4, pDmNTSCAT. In each case 15ug of plasmid was transfected together with 5ug copia-CAT. Molecular length markers were run in lane "M". **B.** CAT assays carried out on 200ug protein from cells transfected with the above plasmids.

Although CAT enzyme is produced in these transfection experiments, we were unable to detect the copia-CAT transcript. This would have given a 215bp extended primer at the position marked with an open triangle in panel A of Fig. 4. Thus either the copia promoter does not direct transcription at sufficient levels to be detected by this technique, or the RNA product is not sufficiently stable to be detected.

Plasmids pDm-306CAT and pDm-43CAT are transcribed more efficiently than pDm-72CAT and pDm-60CAT.1.

SL2 cells were transfected separately with the plasmids pDmNTSCAT, pDm-306CAT, pDm-72CAT, pDm-60CAT or pDm-43CAT.1., and RNA extracted for primer extension reactions as before (Fig. 5.). Specific bands of the predicted size can be seen in tracks 1, 2 and 5, corresponding to transcription from the plasmids pDmNTSCAT, pDm-306CAT and pDm-43CAT.1. However, we were unable to detect transcripts from pDm-72CAT or pDm-60CAT, and yet the cells had been successfully transfected since we could detect CAT activity from the expression of co-transfected copia-CAT (not shown). The experiment was repeated with Kc cells and twice with SL2 cells with the same results. The relative intensities of the extended primer bands produced from pDmNTSCAT, pDm-306CAT and pDm-43CAT.1 transcripts were consistent between these

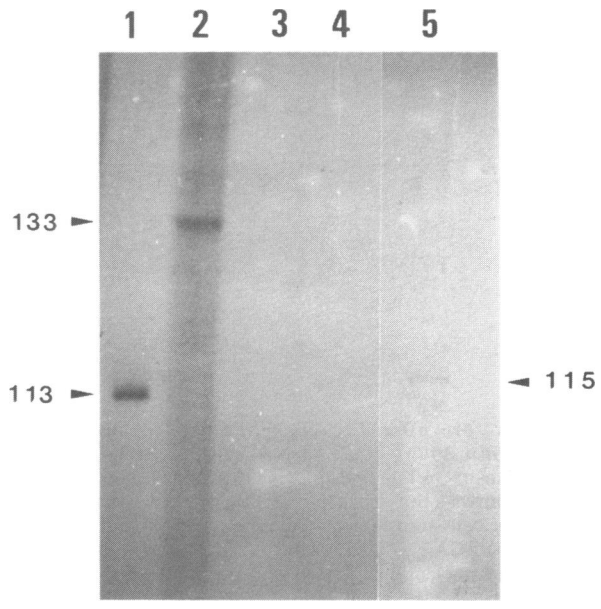

Figure 5. Transcription of the rDNA-CAT Deletion Plasmids. Primer extension reaction carried out on 30ug total RNA from SL2 cells co-transfected with 15ug of rDNA CAT plasmid and 5ug of copiaCAT: lane 1, pDmNTSCAT; lane 2, pDm-306CAT; lane 3, pDm-72CAT; lane 4, pDm-60CAT; lane 5, pDm-43CAT.1. Arrows indicate lengths of reverse transcripts in nucleotides.

experiments when measured using a scanning densitometer, taking into account the transfection efficiency, as judged by copia-CAT activity. This indicated that pDmNTSCAT is transcribed at a level 4-5 times higher than pDm-306CAT which, in turn, is transcribed approximately twice as well as pDm-43CAT.1.

One might have expected the plasmids with 60 bp or 72 bp of upstream sequences to have been transcribed as efficiently as the plasmid pDm-43CAT, which contains the minimum sequence between -43 and +4 expected to have promoter properties (3,4). Furthermore, constructs with 60 bp or 72 bp of upstream sequences were found to be efficiently transcribed in the experiments of Kohorn and Rae (4) with their *in vitro* system. We were concerned that some difference between the plasmid constructs other than the amounts of upstream rDNA sequences might be responsible for this result. The plasmids pDm-72CAT and pDm-60CAT not only differ from pDm-43CAT.1 in having additional upstream rDNA sequences, but also in that they contain an additional 13 bp (consisting mostly of the *Pst*I and *Sal*I sites of the pEMBL9+ polylinker) separating the rDNA and the CAT gene. This is a consequence of the protocols used for plasmid construction (see Materials and Methods). We therefore constructed a second plasmid with 43 bp of upstream rDNA sequences, pDm-43CAT.0, having the

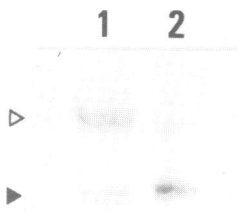

Figure 6. Transcription of pDm-43CAT. Lane 1 shows a primer extension reaction carried out on 30ug total RNA from SL2 cells transfected with equimolar amounts of pDm-43CAT.1 and pDm-43CAT.0 (7.5ug of each added to 4.5 X 10^7 cells). Lane 2 shows the results of a primer extension reaction carried out on 30ug total RNA from cells transfected with equimolar amounts of pDm-43CAT.1 and pDm-72CAT. The open triangle corresponds to 124 nucleotides and the closed triangle to 115 nucleotides.

identical 13 bp separating the rDNA and the CAT gene as pDm-72CAT and pDm-60CAT (see Materials and Methods). The two plasmids, pDm-43CAT.1 and pDm-43CAT.0, were then introduced in equimolar amounts into SL2 cells. Primer extension products synthesised against RNA extracted from these cells is shown in lane 1 of Fig. 6. The lower band, marked with a filled triangle, represents primer extended against transcripts from pDm-43CAT.1 whereas the upper one, marked with an open triangle, runs at the expected position for a reverse transcript made from RNA transcribed from pDm-43CAT.0. This shows that there is nothing inherent in the additional 13 bp present in pDm-43CAT.0, pDm-72CAT, and pDm-60CAT, which may lead either to to inefficient transcription or decreased transcript stability. In fact the presence of these additional sequences appears, if anything, to lead to a slight increase in the level of transcription or of the stability of pDm-43CAT.0 transcripts. The results of an experiment in which SL2 cells were transfected with equimolar amount of pDm-72CAT and pDm-43CAT.1 is also shown in Fig.6 (lane 2). There is a 115 nucleotide band representing pDm-43CAT.1 transcripts, but we cannot detect a band at the position expected for a reverse transcript of pDm-72CAT RNA (124 nucleotides). This demonstrates directly that the plasmid containing 72 bp of sequence upstream from the promoter is inefficiently transcribed under conditions that permit efficient transcription of a plasmid containing 43 bp of upstream sequence.

pDmNTSCAT successfully competes with pDm-306CAT for transcription

In *Xenopus*, competition experiments with rDNA plasmids have showed that transcription occurs preferentially from rDNA plasmids containing longer

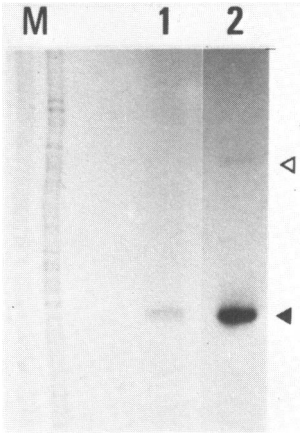

Figure 7. Competition Between pDmNTSCAT and pDm-306CAT. Lane 1 shows the result of a primer extension reaction carried out on 30ug total RNA from SL2 cells transfected with pDmNTSCAT. Lane 2 shows the results of a primer extension reaction carried out on 30ug total RNA from SL2 cells co-transfected in a separate experiment with equimolar amounts of pDmNTSCAT and pDm-306CAT (5.7ug and 3.5ug, respectively, added to 4.5 X 10^7 cells). The dose response of the cells to transfecting DNA indicates amounts of DNA between 0.5 and 15 ug are saturating under these conditions. Molecular length markers were run in lane "M".

upstream sequences. To see if a similar effect could be demonstrated in *Drosophila* tissue culture cells, we introduced equimolar amounts of pDmNTSCAT and pDm-306CAT into SL2 cells and analysed the transcription products in a primer extension reaction. As expected, two bands can be seen corresponding to the products of primer extension on RNA transcribed from pDm-306CAT (Fig. 7, lane 2; band marked with an open triangle) and on the pDmNTSCAT transcript (closed triangle). The ratio of the intensities of these bands was measured using a scanning densitometer and the lower band was found to be approximately 19 more intense than the upper band. This is over three times as high as the ratio of band intensities found for these same plasmids when transfected separately. The plasmid pDm-306CAT is presumably transcribed at a lower relative level in the presence of pDmNTSCAT because the larger plasmid is more successful in competition for some limiting transcription factor.

DISCUSSION.

We have described experiments in which we have used the bacterial gene that encodes chloramphenicol acetyl transferase as a "reporter" to follow the activity of the *Drosophila* rDNA promoter in cultured cells. *Drosophila* cells transfected with the rDNA-CAT plasmids do not contain significant levels of CAT enzyme, in spite of the fact that the CAT gene in several of these plasmids is efficiently transcribed. This confirms previous observations of

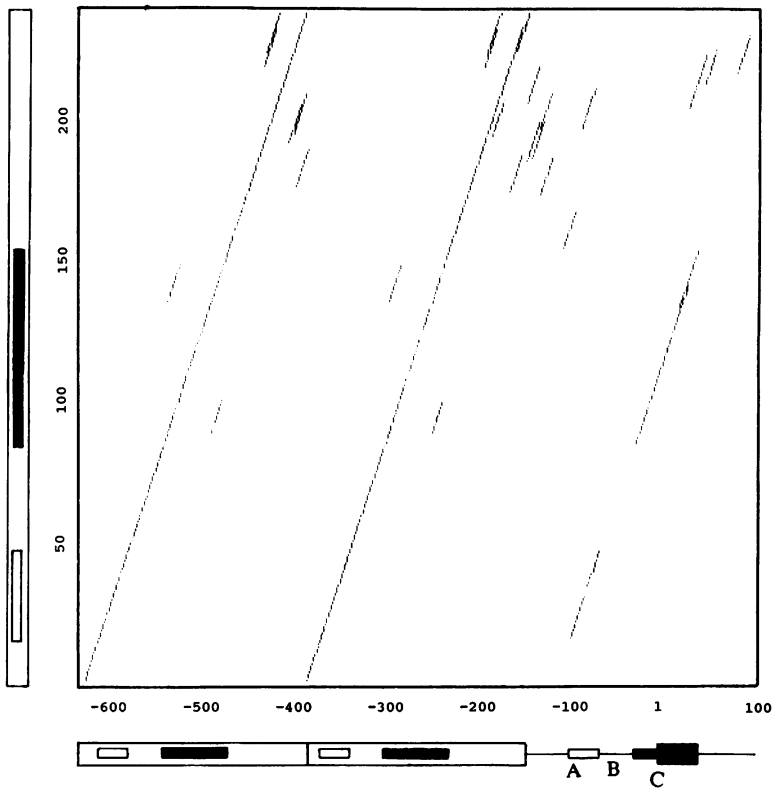

Figure 8. Matrix Homology Comparison between a 240 bp Repeat and Sequences around the major transcription initiation site. The transcription start site is position 1 on the abscissa. The sequence around this point shows two regions of homology, termed "A" and "C", with regions in the 240 base pair repeats, indicated by open and filled blocks respectively.

Grimaldi and DiNocera (9) using similar constructs and is perhaps to be expected since RNA polymerase I transcripts are synthesised in the nucleolus; they undergo different processing events as compared with RNA polymerase II transcripts; they are not capped and are not normally translated *in vivo*. The extremely low levels of CAT enzyme seen in cells transfected with some of the rDNA-CAT plasmids probably results from fortuitous transcription of the CAT gene by RNA polymerase II. This has been shown to be the case with low levels of protein produced by mouse and human rDNA promoter constructs (20,21).

The *Drosophila melanogaster* rDNA NTS contains several tandem reiterations of the sequence at the transcription start site. These reiterations are found within 240 bp repeating units as part of a longer region of interrupted homology with the sequence upstream of the transcription initiation site.

These reiterations are illustrated in the matrix homology comparison shown in Fig. 8 between a 240 bp NTS repeat (the ordinate) and the major transcription start site for the 40S rRNA precursor and its upstream sequences (the abscissa). The regions in the vicinity of the transcription start site with homology to the 240 bp repeat are indicated "A" (-105 to -65) and "C" (-25 to +35). In experiments with cell-free extracts, Kohorn and Rae (3,4) found that the sequence between 43 bp upstream and 4 bp downstream of the transcription start site was the minimal requirement for efficient rDNA transcription. We refer to this sequence as the "minimal rDNA promoter" and it overlaps with region "C" referred to in Fig. 8.

Grimaldi and DiNocera (9) demonstrated that constructs having the rDNA sequences between -180 and +34 of the transcriptional unit were sufficient for expression *in vivo* using a transfection system similar to the one we have employed in this paper. However, the presence of an intact NTS conferred a transcriptional advantage in competition experiments. Recently they have shown that the efficiency of the rDNA promoter is directly proportional to the number of 240 bp repeats (10). In this paper, we have extended their observations using a series of constructs with varying amounts of rDNA sequence immediately upstream of the minimal rDNA promoter region. Attempts to quantitate the levels of transcription of the rDNACAT plasmids were hampered by our inability to detect specific transcription from the control plasmid, copia-CAT. Instead, we used the abundant expression of CAT protein produced from the copia-CAT plasmid to control the efficiency of transfection. Nevertheless, our measurements of the relative level of transcription of the rDNACAT plasmids were remarkably consistent in separate experiments and indicated that a plasmid with the complete NTS, pDmNTSCAT, was expressed at a level 4-5 times higher than a construct with 306 bp of upstream sequence, pDm-306CAT, which, in turn is expressed approximately twice as well as pDm-43CAT.1, having only the minimal promoter region. pDmNTSCAT was expressed at a level approximately 20 times higher than pDm-306CAT when equimolar amounts were simultaneously introduced into cells. This competition effect is consistent with the results of Grimaldi and DiNocera (9) who showed that a plasmid equivalent to pDmNTSCAT would successfully compete for transcription with a plasmid having 180 bp of upstream rDNA sequence. The smallest plasmids to be transcribed pDm-43CAT.1 and pDm-43CAT.0, contain the minimal promoter element, whereas the plasmid, pDm-306CAT, contains a reiteration of the transcription start site (region "C") within part of a 240 bp repeat. pDmNTSCAT, on the other hand, contains seven 240 bp repeats (22), each of which has a transcription start site reiteration. It would appear therefore that these data are consistent with the thesis that transcript levels are in proportion to the numbers of promoter reiterations contained in the plasmids. These findings are in broad agreement with results obtained in competition experiments to study transcription of rDNA in *Xenopus* in which plasmids with the most sequence upstream of the rDNA promoter are preferentially transcribed. The sequences responsible for this effect in *Xenopus* are the

60/81 bp repeats, which contain homology to a region between -114 and -72 with respect to the transcription start site and the reiterations of the transcription start site (11-13).

We believe that our inability to detect transcription *in vivo* from plasmids containing either 72 or 60 bp of upstream sequences, pDm-72CAT and pDm-60CAT probably reflects the transfection efficiencies of our experiments, since in their recent work Grimaldi and DiNocera have been able to see expression with a construct containing 72 bp of upstream sequences although they have not compared this to a -43 bp construct (10). However, it is clear from our observations that pDm-72CAT and pDm-60CAT are transcribed much less efficiently than a construct with less upstream DNA, pDm-43CAT, or one with more upstream DNA, pDm-306CAT. We have confirmed the structure of pDm-72CAT and pDm-60CAT by DNA sequencing. It is unlikely that the physical juxtaposition of vector sequences next to rDNA in both pDm-72CAT and pDm-60CAT acts to create sequence configurations which reduce expression, since this would have to occur in the two constructs even though the vector-rDNA junction is different in each case. Another possibility is that a sequence element essential for efficient transcription *in vivo* has been removed in pDm-72CAT and pDm-60CAT, and that this sequence is artificially reconstructed in pDm-43CAT.1 and pDm-43CAT.0. Such a sequence would have to reside in the NTS between 306 bp and 72 bp upstream of the transcription initiation site, since pDm-306CAT is efficiently transcribed. It is possible to imagine, for example, that a transcription terminator could have such an effect. Sequences directing the termination of *Xenopus* and mouse rDNA transcription have been shown to influence the efficiency of transcription. They probably act by preventing readthrough transcription from interfering with initiation at the rDNA promoter, and by placing RNA polymerase in a favourable position for reinitiation (23-26). A possible candidate for a *Drosophila* transcription termination sequence is the site of 3' end formation of spacer transcripts at a position 125 bp upstream of the transcription initiation site (6). This site falls within a small region which is conserved between species of the melanogaster species subgroup (Hayward and Glover, in preparation). However, there is doubt whether termination of transcription does occur at this site (8).

The DNA between 105 base pairs upstream and 35 base pairs downstream of the transcription start site shows homology to a sequence within the 240 bp repeats. This homology is, however, disrupted between positions -60 and -25 (region "B" in the matrix comparison of Fig 8.). This region contains the 17 nucleotide sequence that distinguishes pDm-60CAT from pDm-43CAT. Transcription not only initiates at the junction of the NTS with the so-called "external transcribed spacer" (ETS), but also at the duplications of promoter sequences within the 240 base pair repeats. This has been demonstrated in the cell free system derived from Kc cell extracts (5) and also *in vivo* (6,7) in which case transcripts of the NTS varying in length by integral values of 240 bp are observed. Furthermore Tautz and Dover (8) have been unable to demonstrate transcription termination at any fixed point from S1 mapping and

nuclear "run-on " assays, suggesting that transcription of one rDNA unit can occur up to the major promoter of the next unit. It is necessary that the sequences at the major promoter site in *Drosophila* differ in some way from the reiterations of these sequences in the NTS in order to direct productive transcription to give the major 40S primary transcript. An alternative explanation of our data therefore would be that by dissecting the major promoter into constituent parts, we have exposed a region of DNA that distinguishes the major promoter from similar sequences in the NTS and which could therefore be of functional significance. It appears that the negative effect of this interaction is not seen if both regions "A" and "C" are available suggesting perhaps that RNA polymerase I or "initiation factors" might interact with these regions to overcome the effect of region "B". It will be of interest to examine such interactions once *Drosophila* RNA polymerase I and its associated transcription factors become available.

ACKNOWLEDGEMENTS

We would like to thank Peter Rae and Bruce Kohorn for providing us with plasmids. We are grateful to the Cancer Research Campaign for supporting all aspects of this work including a Career Development Award to DMG and a Research Studentship to DCH.

*To whom correspondence should be addressed

+Present address: Research School of Biological Sciences, Australian National University, Canberra, ACT, Australia

REFERENCES

1. Coen, E. S. and Dover, G. A. (1982) Nucl. Acids Res. **10**, 7017-7026.
2. Simeone, A., La Volpe, A. and Boncinelli, E. (1985) Nucl. Acids Res. **13**, 1089-1101.
3. Kohorn, B. D. and Rae, P. M. M. (1983) Nature **304**, 179-181.
4. Kohorn, B. D. and Rae, P. M. M. (1983) Proc. Natl. Acad. Sci. USA **80**, 3265-3268.
5. Kohorn, B. D. and Rae, P. M. M. (1982) Nucl. Acids Res. **10**, 6879-6886.
6. Murtif, V. L. and Rae, P. M. M. (1985) Nucl. Acids Res. **13**, 3221-3239.
7. Miller, J. R., Hayward, D. C. and Glover, D. M. (1983) Nucl. Acids Res. **11**, 11-19.
8. Tautz, D. and Dover, G. A. (1986) EMBO J. **5**, 1267-1273.
9. Grimaldi, G. and Di Nocera, P.P. (1986) Nucl. Acids Res. **14**, 6417-6432.
10. Grimaldi, G. and Di Nocera, P.P. (1988) Proc. Natl. Acad. Sci. USA in press
11. Reeder, R. H., Roan, J. G. and Dunaway, M. (1983) Cell **35**, 449-456.
12. Labhart, P. and Reeder, R. H. (1984) Cell **37**, 285-289.
13. De Winter, R. F. J. and Moss, T. (1986) Cell **44**, 313-318.
14. Roiha, H., Miller, J. R., Woods, L. C. and Glover, D. M. (1981) Nature **20**, 749-753.
15. Walker, M.D., Edlund, J., Boulet, A. and Rutter, W. (1983) Nature **306**, 557-561.
16. Gorman, C. (1985) In Glover, D. M. (ed), DNA cloning, IRL Press Ltd., Oxford, Vol. II, pp. 143-190.

17. Davies, J. A., Addison, C. F., Delaney, S. J., Sunkel, C. and Glover, D. M. (1986) J. Mol. Biol. **189**, 13-24.
18. Grummt, I. and Skinner, J. A. (1985) Proc. Natl. Acad. Sci. USA 82, 722-726.
19. Di Nocera, P. P. and Dawid, I. B. (1983) Proc. Natl, Acad. Sci. USA. **80**, 7095-7098.
20. Smale, S. T. and Tjian, R. (1985) Mol. Cell. Biol. **5**, 352-362.
21. Lopata, M. A., Cleveland, D. W. and Sollner-Webb, B. (1986) Proc. Natl. Acad. Sci. USA. **83**, 6677-6681.
22. Tautz, D., Tautz, C., Webb, D. and Dover, G. A. (1987) J. Mol. Biol. **195**, 525-542.
23. Henderson, S. and Sollner-Webb, B. (1986) Cell **47**, 891-900.
24. McStay, B. and Reeder, R. H. (1986) Cell **47**, 913-920.
25. Mitchelson, K. and Moss, T. (1987) Nucl. Acids Res. **15**, 9577-9596.
26. Grummt, I., Kuhn, A., Bartsch, I. and Rosenbauer, H. (1986) Cell **47**, 901-911.