Volume 8 Number 21 1980 Nucleic Acids Research

Two Drosophila melanogaster $tRNA^{Gly}$ genes are contained in a direct duplication at chromosomal locus 56F

N.Davis Hershey and Norman Davidson

Department of Chemistry, California Institute of Technology, Pasadena, CA 91125, USA

Received 3 September 1980

ABSTRACT

One *Drosophila melanogaster* $tRNA^{Gly}$ gene occurs on each 1.1 - 2.0 kb unit of a direct duplication at chromosomal region 56F. The nucleotide sequence of the gene and the 5' flanking region has been determined. The non-transcribed strand sequence of the tRNA gene is: 5' GCATCGGTGGTTCAGTGGTAGAATGCTCGCCTGCCA CGCGGGCGGCCCGGGTTCGATTCCCGGCCGATGCA 3'. This nucleotide sequence is identical to that of the major glycine tRNA in *Bombyx mori* posterior silk gland. Within the 22 kb region mapped, additional tRNA genes are found, an observation consistent with reports that genes for other isoacceptors are present at this locus.

INTRODUCTION

Until recently, the arrangement of the 600-800 *Drosophila melanogaster* (Dm) tRNA genes, comprising approximately 60 different tRNA sequences (1,2,3) has been studied by *in situ* hybridization to polytene chromosomes of either total Dm 4S RNA (4,5) or purified tRNA species (6,7,8,9,10,11,12). These studies have established a number of important facts: a) the tRNA genes are clustered at 54-60 chromosomal loci, which by silver grain counts probably contain quite different numbers of tRNA genes; b) a particular locus may have genes for tRNAs which accept different amino acids; c) genes for a single isoacceptor may be found at more than one locus.

The development of recombinant DNA technology and rapid DNA sequencing methods has allowed the arrangement of eucaryotic tRNA genes to be studied in the greatest possible detail (13,14,15,16,17). The most striking feature of this arrangement by a number of criteria has been the low degree of regularity. Genes within a cluster are irregularly spaced, intermingled with respect to amino acid species accepted and may be encoded on both DNA strands. Flanking sequences show very limited sequence homology and an absence of recurrent secondary structure features. While regularities do exist, the variability in the arrangement of tRNA genes is their more impressive aspect.

The mapping of approximately 22 kb of Dm DNA surrounding two $tRNA^{Gly}$ genes is reported here; these genes exist as part of a duplicated region. A 0.28 kb portion of the 1.1 - 2.0 kb repeat unit, including the gene and a large section of 5' flanking region, has been sequenced.

MATERIALS AND METHODS

Construction of the *Drosophila melanogaster* (Canton S wild type) library in the vector Charon 4 using Dm DNA partially digested with Eco RI has been described (18,19). A similar library with random shear fragments inserted into the vector with synthesized Eco RI linkers was a generous gift of Joyce Lauer (20). Clones were selected from these libraries by plaque filter hybridization (21). Recombinant phage were grown by the PDS method and DNA was isolated after lysis of the phage (22).

Recombinant phage DNA was digested with Eco RI; the resulting fragments were inserted into the vector pKH47 (23), a derivative of pBR322 (24), at the Eco RI site by standard subcloning techniques (15,25). Bacterial colonies containing recombinant plasmids were screened with labeled DNA fragments (26) and plasmid DNA was prepared (13) from the selected clones.

Restriction endonucleases were purchased from New England Biolabs; digests were performed as recommended by the supplier.

Two types of gels were used to analyze DNA fragments and for Southern blots (27). For larger fragments a 0.7% agarose slab gel, 14 x 20 x 0.4 cm, in 40 mM Tris-acetate, 1 mM EDTA pH 7.4 was used at 5 volt/cm. Smaller fragments were separated on a 4% acrylamide gel, 14 x 20 x 0.16 cm, in 50 mM Tris-borate, 1 mM EDTA pH 8.5 at 5 volt/cm. DNA fragments were isolated from agarose gels as described previously (28) and from acrylamide gels according to Maxam and Gilbert (29).

Dm 4S RNA (13) was labeled with ^{32}P either a) at the 5' end with α ^{32}P-ATP and T4 polynucleotide kinase or b) at the 3' end using 5' ^{32}P, 3'-diphosphocytidine and T4 RNA ligase (15). DNA was ^{32}P-labeled by nick translation (30).

Nucleotide sequencing was performed according to Maxam and Gilbert (29). Fragments for sequencing were labeled at the 5' end with α ^{32}P-ATP and T4 polynucleotide kinase.

In situ hybridization to polytene chromosomes was carried out as previously reported (28).

All research involving recombinant DNA was performed in accordance with NIH Guidelines.

RESULTS

Recombinant Phage Clone Selection

Plasmid pCIT5, isolated from a *Drosophila melanogaster* (Oregon R wild type) library in the preliminary stages of a previous study (13) carries a 1.5 kb Dm DNA insert which encodes an undetermined number of tRNA genes, probably one. The phage library having Eco RI partial digest inserts was screened for DNA homologous to labeled pCIT5 DNA. Several clones having homology to pCIT5 were found by this procedure. Clone λDmt56-6, which in addition hybridized to Dm 4S RNA, was chosen for further study.

Since the tRNA genes of λDmt56-6 were located at the extreme end of the 13.5 kb insert (see below), the isolation of λ clones having adjacent, overlapping inserts was undertaken to determine if additional tRNA genes were present in the immediate area. For this purpose, a library having random shear fragments of Dm DNA as inserts was used (20). Plasmid p56-33, the insert of which is the Eco RI G/G' fragment of λDmt56-6 (see below), was used to select two overlapping clones, λDmt56-94 and λDmt56-102. The insert of λDmt56-102 is totally contained within that of λDmt56-94.

Restriction Enzyme Mapping of Recombinant Phage Clones

The Dm inserts of these three clones were characterized by restriction mapping using a battery of standard techniques including gel electrophoresis of restriction enzyme digests, digests with two enzymes, and partial digests. Southern blots of these gels were probed with labeled 4S RNA to map the positions of tRNA genes. Comparison of the three clones in these analyses facilitated construction of the map.

The restriction site map covering ca. 22 kb of Dm DNA is shown in Figure 1A. Since the inserts of all three clones do overlap, they are most likely derived from the same region of the genome. Two Eco RI fragments of identical size (1.1 kb), G and G' of λDmt56-6, carry the tRNA genes of interest. The fact that both G and G' carry tRNA genes was established by hybridization of 4S RNA with Eco RI partial digest fragments of λDmt56-6 in a Southern blotting experiment. Partial fragments comprising either G or G' to the exclusion of the other both hybridized 4S RNA, thus demonstrating that both G and G' encode tRNAs (data not shown). Moreover, the G and G' fragments are identical by restriction analysis (see below); so both must have tRNA genes.

Additional tRNA genes of unknown number and identity were discovered in the course of mapping to be confined to the rightmost ca. 2.0 kb of mapped DNA (Figure 1A). These genes have not yet been investigated and will not be considered further in this manuscript.

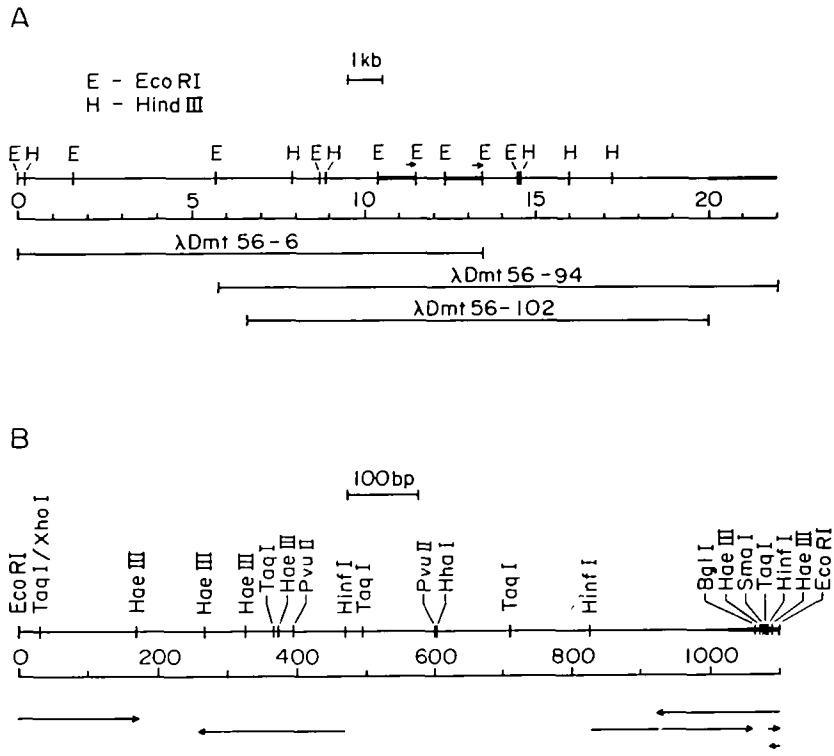

Figure 1. Restriction Site Maps of Cloned *Drosophila* DNA Encoding tRNA Genes at Region 56F

A. Region comprising two $tRNA^{Gly}$ genes. Sites of cleavage by the restriction endonucleases Eco RI (E) and Hind III (H) are indicated. The smallest regions to which Dm 4S RNA hybridization can be confined by this mapping are shown by the heavy lines. The arbitrary orientation of the map is such that the two $tRNA^{Gly}$ genes are transcribed from left to right as indicated by the arrows; the coordinate scale is calibrated in kilobase pairs. The lines below the map show the extent of the inserts in each of the three Charon 4 clones. In all three cases the Charon 4 right arms are to the left of the insert. The Hind III site at coordinate 14.6 kb is abnormal in that a fraction (ca. one half) of the DNA is fully resistant to Hind III cleavage at this site in both clones. The relative position of this Hind III site and the nearby Eco RI site is uncertain; the Hind III site is probably on the right.

B. λDmt56-6/Eco RI G fragment; Sequencing scheme. The coordinate scale is calibrated in base pairs. The position of the tRNA sequence indicated by the heavy line was established by sequencing. The tRNA sequences were known prior to sequencing only to be right of the Hinf I site at coordinate 825 bp. The positions of some sites were determined by sequencing. The only uncertainty in this map is the relative position of the Hha I and Pvu II sites at coordinate 600 bp (the most likely arrangement is depicted).

The direction and extent of sequencing from 5' ^{32}P-labeled termini is shown by the arrows below the map. In several cases sequencing either began or terminated within the vector, pKH47.

The tRNA Genes are Contained Within a Direct Repeat Sequence

To determine if the sequences of the λDmt56-6/Eco RI G and G' fragments on this map were identical, λDmt56-6 DNA was digested with Eco RI, the fragments separated by gel electrophoresis and the DNA of the G/G' band eluted from the gel.

This DNA was labeled by nick translation, digested with the restriction enzymes Hae III or Hha I, and the resulting fragments separated by electrophoresis and detected by autoradiography. If the G and G' fragments are identical, the sum of the fragment lengths generated by this procedure is expected to be 1.1 kb, the length of either one of the fragments. If, however, the G and G' fragments only coincidentaly have the same length, the products of these digests should have lengths totaling 2.2 kb. An autoradiogram of the gels from such digests as described above is shown in Figure 2A. The sum of the fragment lengths are 1112 and 1192 bp for the Hae III and Hha I digests, respectively; whereas the undigested control fragment is 1140 bp in length. These results were fully confirmed by similar digests of isolated G + G' fragments in amounts sufficient for visualization of the products by

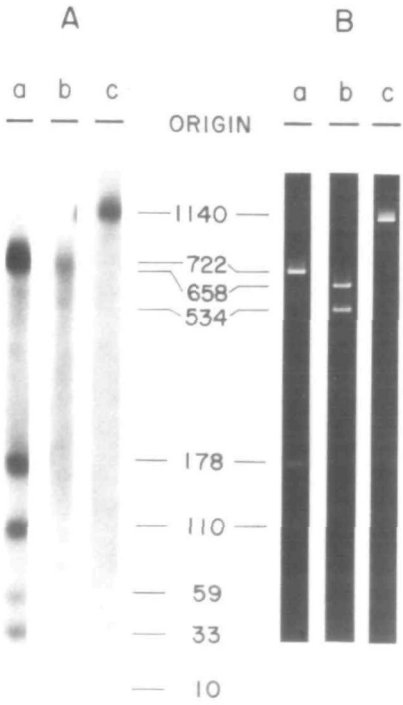

Figure 2. Gel Electrophoresis Pattern of Fragments Produced by Restriction Endonuclease Digestion of Isolated λDmt56-6/Eco RI G + G' Fragments. The G + G' fragments of a λDmt56-6/Eco RI digest were isolated from a 0.7% agarose gel. A fraction of this DNA was labeled by nick translation using α ^{32}P-dCTP. Portions of both labeled and unlabeled DNA were separately digested with restriction endonucleases Hae III or Hha I. The resulting fragments were separated on a 4% polyacrylamide gel. Labeled fragments were detected by autoradiography of the gel for 10 hr at room temperature using Kodak XR-5 film (panel A). Unlabeled fragments were visualized by ethidium bromide staining (panel B). Using the fragments from Hae III and Hinf I digests of pBR322 as length standards, a calibration curve was constructed. Fragment sizes given in this Figure were determined from that curve. Restriction endonucleases used and the sums of fragment sizes are: a) Hae III, 1112; b) Hha I, 1192; c) Control, undigested G + G' fragments, 1140.

ethidium bromide staining (Figure 2B). In this instance, fragments shorter than ca. 0.1 kb could not be seen. Within the limits of resolution of this technique, the G and G' fragments are therefore identical.

Fragments G and G' were found to have a Sma I site very near one end and an Xho I site very near the opposite end (see below). These sites can be positioned on λDmt56-6 with respect to the Hind III site at coordinate 8.9 kb (Figure 1A) for the lefthand repeat unit and with respect to a known Kpn I site in the left arm of Charon 4 (31) for the righthand unit by standard restriction enzyme mapping. The unique interpretation of the results of this analysis (not shown) is that G and G' are parts of a direct repeat which is oriented as indicated in Figure 1A.

The map of Figure 1A reveals that the repeat units, of which G and G' are parts, cannot be longer than the length of G/G' (1.1 kb) plus the distance separating G and G' (0.9). Therefore the unit length must be between 1.1 and 2.0 kb. A further refinement of these limits has not been attempted.

Subcloning and Restriction Mapping of the λDmt56-6/Eco RI G/G' Fragment

The Eco RI fragments of λDmt56-6 were inserted by standard subcloning techniques at the Eco RI site of vector pKH47. Two plasmids, p56-33 and p56-41, having the G/G' fragment inserted in opposite orientation were selected using nick-translated G/G' fragment DNA, which had been eluted from a gel. The opposing orientations of the inserts in these two clones proved to be a convenience for both mapping and sequencing, because this situation offered a wide choice of combinations of vector and insert restriction enzyme sites for any given purpose. Restriction enzyme digests and double digests were analyzed by gel electrophoresis and subjected to Southern blotting and hybridization with a Dm 4S RNA probe to yield the map shown in Figure 1B. (The information on the map has been verified and refined by the sequencing reported below.) This analysis localized the tRNA coding region to the right of the Hinf I site at coordinate 825 (Figure 1B). Accordingly, this region and some others were sequenced by the strategy indicated in Figure 1B. All ambiguities in the 279 ntp region from the Hinf I site to the Eco RI site at the right end of the insert were resolved by sequencing overlapping fragments with the same or reversed polarity.

The resulting sequence (Figure 3) was inspected for the presence of tRNA coding sequences, first by a search for the highly conserved sequence in the loop IV, 5' GTTCPuAXPyC 3', where Pu = purine, Py = pyrimidine and X = any base (32). The familiar tRNA secondary structure could then be deduced from the sequence using the models of tRNA structure as a guide (32). A DNA

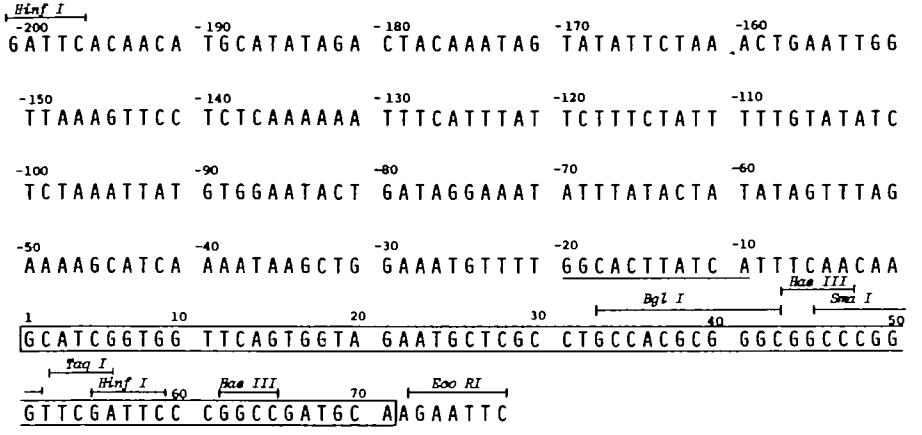

Figure 3. Nucleotide Sequence of the Rightmost 0.28 kb of the λDmt56-6/Eco RI Fragment. Polarity of this non-transcribed DNA strand sequence is 5' to 3' from left to right. All mapped restriction endonuclease sites are indicated above the sequence. Numbering begins at the first nucleotide of the $tRNA^{Gly}$ gene sequence, which is boxed; nucleotides 5' to the gene are numbered negatively beginning with the nucleotide immediately 5' to the gene sequence. The consensus sequence (see text) is underlined.

coding sequence for a single tRNA gene with a glycine anticodon, GCC, was found. This sequence is boxed in Figure 3 and displayed as tRNA with the predicted secondary structure in Figure 4.

In Situ Localization of the λDmt56-6 Insert

In situ hybridization to polytene chromosomes of ^3H-labeled cRNA transcribed from λDmt56-6 led to labeling at a single chromosomal locus, 56F (Figure 5). The insert of λDmt56-6 is 13.5 kb in length, and the duplicated region which includes the tRNA gene is at least 1.1 kb. If the duplicated unit were repeated at another locus, the ratio of the number of grains at that locus to the number at 56F should exceed 1.1/13.5. Exposures calculated to be sufficiently long to give ca. 450 grains at 56F revealed labeling at no other locus. Thus no sequence of the λDmt56-6 insert greater than 1 kb, including the duplicated sequence, is repeated elsewhere in the genome. Furthermore, when the G/G' fragment of λDmt56-6 (subcloned as p56-33) was used to select clones from the Dm genomic library, only overlapping clones originating from the same section of the genome were isolated, as judged by restriction mapping. This result further suggests that the duplicated segments are found only at 56F. Finally, another indication that the λDm56-6 insert is mostly single copy DNA is the following: only fragments present in the insert were hybridized (and to the same relative intensity as their

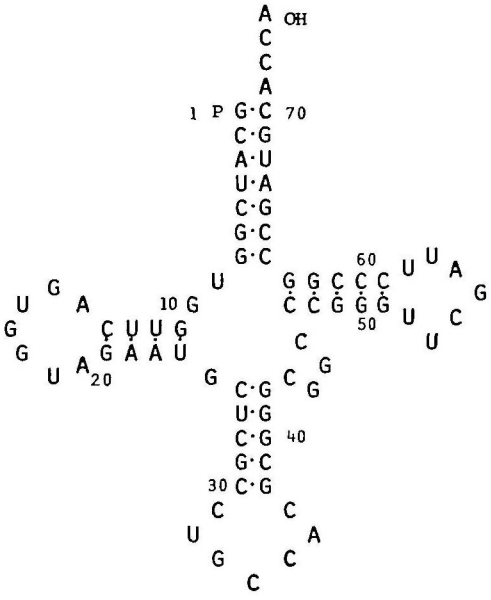

Figure 4. Secondary Structure Predicted for the RNA Transcript of the *Drosophila melanogaster* $tRNA^{Gly}$ Gene. The 3' terminal triplet, CCA, expected in the mature tRNA is shown here, although it is not encoded in the DNA. The nucleotide modifications of the mature tRNA are not known.

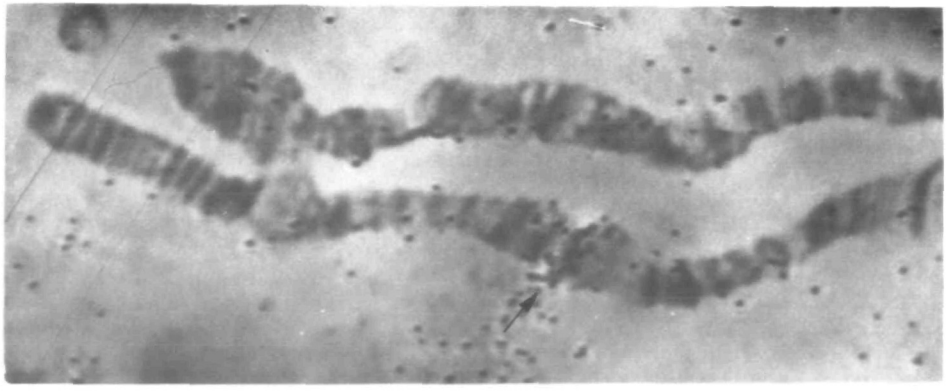

Figure 5. Autoradiogram of *Drosophila melanogaster* Salivary Gland Chromosomes Hybridized *In Situ* to cRNA Transcribed from λDmt56-6. Approximately 0.02 ug cRNA (specific activity 9x 10^7 dpm/ug) was applied to each squash of salivary gland chromosomes from giant mutants of *Drosophila melanogaster*. Labeling occurs only at 56F (arrow); approximately 60 silver grains are present at this locus (9 day exposure).

representation in the insert) when Southern blots of Eco RI digested Dm genomic DNA were probed with ^{32}P-labeled λDmt56-6 DNA (data not shown).

In summary, unless the 13.5 kb of λDmt56-6 insert DNA is included in an even larger repeat in region 56F, the two $tRNA^{Gly}$ genes are part of two 1.1 - 2.0 kb closely spaced direct repeats which are found nowhere else in the genome.

DISCUSSION

Chromosomal Location

In situ hybridization of the major Dm glycine isoacceptor, $tRNA_3^{Gly}$, led to localization at three sites, 22BC, 35BC, and 57BC, all on the second chromosome (10). If the sequences of other glycine tRNA isoacceptors are sufficiently unlike that of $tRNA_3^{Gly}$, additional loci for $tRNA^{Gly}$ genes may have gone undetected. Furthermore, in this experiment some silver grains were seen at 56EF; this hybridization was thought to be due to low levels of 5S rRNA contaminant in the purified tRNA preparation. But the possibility that 56F is a site for $tRNA_3^{Gly}$ genes was also reserved.

Our results show that $tRNA^{Gly}$ genes are indeed present at 56F. Whether or not these genes are the source of the signal in the *in situ* hybridization experiment using $tRNA_3^{Gly}$ is uncertain; the nucleotide sequence of $tRNA_3^{Gly}$ is unknown and consequently the degree of cross-reactivity of $tRNA_3^{Gly}$ with these genes is unknown.

Two other tRNA genes, those for $tRNA_4^{Glu}$ (8) and, tentatively, $tRNA_2^{Phe}$ (12) have been localized to 56F by *in situ* hybridization of purified tRNA species. The presence of additional, unidentified tRNA genes about 7 kb from the $tRNA^{Gly}$ genes (Figure 1A) is consistent with the existence of a larger tRNA gene cluster in this vicinity.

Repeat Structure

Since the duplicated repeat is found nowhere but 56F, other $tRNA^{Gly}$ genes must be arrange differently. Sequence similarities below the level of the minimum repeat size (1.1 kb) among all Dm tRNA genes are not excluded by this observation. The simple duplication described here can be compared to the remarkable case of *Xenopus laevis*, in which a 3.18 kb sequence containing four tRNA genes is repeated 300 times in the haploid genome (17).

Sequence

The primary $tRNA^{Gly}$ sequence deduced from the DNA sequence (Figure 4) is exactly the same as that for *Bombyx mori* $tRNA_1^{Gly}$ (33,34). This coincidence is not surprising in view of the rather close phylogenetic relationship between Diptera and Lepidoptera and the high degree of sequence conservation of $tRNA^{Gly}$ (and other tRNAs) throughout eucaryotic evolution (14,35). As noted previously (36), eucaryotic glycine tRNAs having the same GCC anticodon are highly homologous (94.6% of the nucleotide positions are the same for the $tRNA_1^{Gly}$ from *B. mori* posterior silk gland and from human placenta), whereas two human glycine tRNAs having different anticodons display much less homology (67.6%).

Like most tRNA genes sequenced thus far, the 3' terminal triplet, CCA of the mature tRNA is not encoded in the DNA. Two instances of the "unusual", but allowed, base pairing of G with U occur at nucleotides 10-24 and 27-41.

The DNA sequences for these genes reveals no intervening sequences as occur in certain tRNA genes of yeast (37), *Drosophila* (16) and *Xenopus* (17).

The most striking feature of the extensive sequence 5' to the gene sequence is its high, uniformly distributed, A+T content of 75% (201 ntp). (In contrast the gene is 32% A+T.) Within this region no repeats, direct or inverted, of any significant extent are found, although several short (5-6 bp) inverted repeats separated by rather long intervals (12-16 bp) occur. Repeated sequences have been proposed as possibly involved in regulation of RNA polymerase III transcription initiation (17,38). In no case, however, has an invariant sequence been found precisely located 5' to RNA polymerase III - transcribed genes even within a single species (14,15,16,38). A consensus sequence beginning at ca. nucleotide -24 of three *Drosophila* $tRNA_2^{Lys}$ genes (14) is repeated in the present case with reasonable fidelity beginning at nucleotide -20; these sequences are compared below:

	-24
$tRNA_2^{Lys}$	GGCAGTTTTTA
$tRNA^{Gly}$	GGCACTTATCA
	-20

Only one of the three positions found to be different here is variable in the three $tRNA_2^{Lys}$ gene 5' flanking regions from which the consensus sequence was derived. Recent evidence suggests that little (39) or none (40,41) of the 5' flanking sequence is necessary for transcription by RNA polymerase III.

RNA polymerase III termination occurs within short stretches of thymidylate residues that follow within 25 bp of the 3' end of the gene (16,38,39). For a number of 5S rRNA genes and adenovirus VA RNA_I gene, GC-rich sequences and regions of dyad symmetry preceed these stretches of thymidylate residues (38). But these characteristics are absent in the tRNA genes of *B. mori* (39) and *Drosophila* (14,16).

None of these features is evident in the 11 nucleotides following the 3' end of the $tRNA^{Gly}$ gene. But any or all of them could exist beyond this limited region.

No data on either *in vitro* or *in vivo* transcription of these $tRNA^{Gly}$ genes is available. That the two genes would be transcribed to similar extents in any given situation seems likely in view of their apparently identical flanking sequences and their close proximity on the chromosome.

ACKNOWLEDGMENTS

We thank our collaborators, R. Robinson and P. Yen, for their participation in the construction of the Dm library and isolation of λDmt56-6 and for numerous discussions about all aspects of this project. This work was supported by a grant from the NIH. Department of Chemistry Contribution Number 6300.

REFERENCES

1. Ritossa, F.M., Atwood, K.C. and Spiegelman, S. (1966). Genetics *54*, 663-676.
2. Tartof, K.D. and Perry, R.P. (1970). J. Mol. Biol. *51*, 171-183.
3. Weber, L. and Berger E. (1976). Biochemistry *15*, 5511-5519.
4. Steffenson, D.M. and Wimber, D.E. (1971). Genetics *69*, 163-178.
5. Elder, R., Szabo, P. and Uhlenbeck, O. (1980). In Transfer RNA: Biological Aspects, D. Söll, J.N. Abelson and P.R. Schimmel, eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 317-323.
6. Grigliatti, T.A., White, B.N., Tener, G.M., Kaufman, T.C. and Suzuki, D.T. (1974). Proc. Nat. Acad. Sci. USA *71*, 3527-3531.
7. Delaney, A., Dunn, R., Grigliatti, T.A., Tener, G.M., Kaufman, T.C. and Suzuki, D.T. (1976). Fed. Proc. *35*, 1676.
8. Kubli, E. and Schmidt, T. (1978). Nuc. Acids Res. *5*, 1465-1478.
9. Dunn, R., Hayashi, S., Gillam, I.C., Delaney, A.E., Tener, G.M., Grigliatti, T.A., Kaufman, T.C. and Suzuki, D.T. (1979). J. Mol. Biol. *128*, 277-287.
10. Hayashi, S., Gillam, I.C., Delaney, A.D., Dunn, R., Tener, G.M., Grigliatti, T.A. and Suzuki, D.T. (1980). Chromosoma *76*, 65-84.
11. Kubli, E., Schmidt, T. and Egg, A.H. (1980). In Transfer RNA: Biological Aspects, D. Söll, J.N. Abelson and P.R. Schimmel, eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 309-315
12. Tener, G.M., Hayashi, S., Dunn, R., Delaney, A., Gillam, I.C., Grigliatti, T.A., Kaufman, T.C. and Suzuki, D.T. (1980). In Transfer RNA: Biological Aspects, D. Söll, J.N. Abelson and P.R. Schimmel, eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 295-307.
13. Yen, P.H., Sodja, A., Cohen, M. Jr., Conrad, S.E., Wu, M., Davidson, N. and Ilgen, C. (1977). Cell *11*, 763-777.
14. Hovemann, B., Sharp, S., Yamada, H. and Söll, D. (1980). Cell *19*, 889-895.
15. Yen, P.H. and Davidson, N. (1980). Cell, in press.
16. Robinson, R.R. and Davidson, N. (1980). Cell, submitted.
17. Müller, T. and Clarkson, S.G. (1980). Cell *19*, 345-353.
18. Yen, P., Hershey, N.D., Robinson, R. and Davidson, R. (1979). In ICN-UCLA Symposium on Eucaryotic Gene Regulation, R. Axel and T. Maniatis, eds., Academic Press, New York, pp. 133-142.
19. Davidson, N., Fyrberg, E.A., Hershey, N.D., Kindle, K., Robinson, R.R., Sodja, A. and Yen, P. (1980). In RNA-Polymerase, tRNA and Ribosomes: Their Genetics and Evolution, S. Osawa, H. Ozeki and T. Yura, eds., University of Tokyo Press, Tokyo, in press.
20. Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978). Cell *15*, 687-701.
21. Benton, W.D. and Davis, R.W. (1977). Science *196*, 180-182.
22. Blattner, F.R., Williams, B.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L.-A., Grunwald, D.J., Kiefer, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L. and Smithies, O. (1977). Science *196*, 161-169.
23. Hayashi, K. and Davidson, N. (1980). Gene, in press.

24. Boliver, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L. and Boyer, H.B. (1977). Gene *2*, 95-113.
25. Cohen, S.N., Chang, A.C.Y., Boyer, H.W. and Helling, R.B. (1973). Proc. Nat. Acad. Sci. USA *70*, 3240-3244.
26. Grunstein, M. and Hogness, D.S. (1975). Proc. Nat. Acad. Sci. USA *72*, 3961-3965.
27. Southern, E. (1975). J. Mol. Biol. *98*, 503-517.
28. Fyrberg, E.A., Kindle, K.L., Davidson, N. and Sodja, A. (1980). Cell *19*, 365-378.
29. Maxam, A. and Gilbert, W. (1980). In Methods in Enzymology, Vol. 65, L. Grossman and K. Moldave, eds., Academic Press, New York, pp. 499-560.
30. Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975). Proc. Nat. Acad. Sci. USA *72*, 1184-1188.
31. Blattner, F.R., personal communication.
32. Clark, B.F.C. (1978). In Transfer RNA, S. Altman, ed., M.I.T. Press, Cambridge, pp. 14-47.
33. Garel, J.P. and Keith, G. (1977). Nature *269*, 350-352.
34. Zúñiga, M.C. and Steitz, J.A. (1977). Nuc. Acids Res. *4*, 4175-4196.
35. Sprinzl, M., Grueter, F., Spelzhaus, A. and Gauss, D.H. (1980). Nuc. Acids Res. *8*, r1-r22.
36. Gupta, R.C., Roe, B.A. and Randerath, K. (1979). Nuc. Acids Res. *7*, 959-970.
37. Knapp, G., Ogden, R.C., Peebles, C.L. and Abelson, J. (1979). Cell *18*, 37-45 (references cited herein).
38. Korn, L.J. and Brown, D.D. (1978). Cell *15*, 1145-1156.
39. Garber, R.L. and Gage, L.P. (1979). Cell *18*, 817-828.
40. Sakonju, S., Bogenhagen, D.F. and Brown, D.D. (1980). Cell *19*, 13-25.
41. Bogenhagen, D.F., Sakonju, S. and Brown, D.D. (1980). Cell *19*, 27-35.