The Structures of Genes Hybridizing with tRNA\textsubscript{Val} from Drosophila melanogaster

William R. Addison, Carolee R. Astell, Allen D. Delaney, Ian C. Gillham, Shizu Hayashi, Robert C. Miller, Bhanu Rajput, Michael Smith, Deborah M. Taylor, and Gordon M. Tener

From the Departments of Biochemistry, Medical Genetics, and Microbiology, University of British Columbia, Vancouver, British Columbia V6T 1W5, Canada

(Received for publication, August 17, 1981)

Segments of cloned Drosophila DNA from four recombinant plasmids that hybridize with tRNA\textsubscript{Val} have been sequenced. The segments from pDt55R and pDt120R that hybridize to 90BC on the third polytene chromosome appear to be either repeats or alleles. They contain one structural gene each of identical sequence but differ at eight sites in 506 base pairs. The structural genes differ at four sites from the sequence expected from that of tRNA\textsubscript{Val}. A third plasmid, pDt14, which hybridizes to 89BC on the third chromosome, also contains a structural gene with the same sequence as those in pDt55R and pDt120R. In addition, pDt14 has a gene for tRNA\textsubscript{Phe} 214 base pairs upstream with the same polarity as the tRNA\textsubscript{Val} gene. The tRNA\textsubscript{Phe} gene contains a 23-base pair segment identical with the corresponding segment in the tRNA\textsubscript{Val} genes except for one base pair. The fourth plasmid investigated, pDt55, hybridizes to 70BC. It contains two tRNA\textsubscript{Val} genes 525 base pairs apart with opposite polarity. These genes have identical sequences, which correspond to that expected from the sequence of tRNA\textsubscript{Val}. There is no evidence that the first three tRNA\textsubscript{Val} genes are expressed at any stage during the development of Drosophila.

In Drosophila, the genes for various RNAs can be localized by autoradiography using \textsuperscript{3}H- or \textsuperscript{32}P-labeled RNA annealed with polytene chromosomes of the salivary glands (1, 2). Transfer RNAs are particularly useful probes for these studies because there are between 50 and 90 individual tRNAs, of which a large number can be readily purified (3). Total crude tRNA hybridizes to about 54 sites on the chromosomes (4) while each purified tRNA hybridizes to at least two and in some cases as many as seven sites (2, 5-8). These sites may be on the same or on different chromosomes. It has been estimated that there is an average of about 10 genes for each tRNA in Drosophila (9); thus, the sites of hybridization are expected to contain clusters of genes for one or more tRNAs. This clustering has been found for a number of tRNA genes including those for tRNA\textsubscript{Gin}, tRNA\textsubscript{Ace}, tRNA\textsubscript{His}, tRNA\textsubscript{Val}, tRNA\textsubscript{Asp}, tRNA\textsubscript{Asn}, and tRNA\textsubscript{Glu} (10-16). In addition, genes for other tRNAs have also been found in close proximity to these gene clusters (10-12).

In this paper, we describe the characterization of five tRNA\textsubscript{Val} genes derived from three different chromosomal sites. Three of these genes differ by four nucleotides from the sequence expected from the tRNA\textsubscript{Val} sequence whereas the other two have the expected sequence.

**MATERIALS AND METHODS**

**General**—The Na\textsuperscript{25}I and [\textsuperscript{3}P]dideoxyribonucleotides 5'-triphosphates were purchased from Amersham Corporation. Restriction endonucleases were obtained from New England Biolabs, Inc. and Bethesda Research Laboratories, Inc. and DNA polymerase was obtained from Boehringer Mannheim Canada. Gevaert Curix x-ray film was purchased from local suppliers. Xma I was purified according to the procedure of Endow and Roberts (17).

**DNA**—Plasmids pDt14, pDt55, pDt92R, and pDt120R were prepared by ligating HindIII-cut Drosophila melanogaster (Oregon R) DNA into pBR322 as described earlier (18). The plasmids were grown, and the DNA was isolated by the methods previously described (18).

**In Situ Hybridization**—The plasmid DNA was labeled with \textsuperscript{32}P and hybridized to squashes of acetylated polytene chromosomes as previously described (18).

**Restriction Analysis**—Plasmid DNAs were hydrolyzed with commercially available restriction endonucleases according to the procedures recommended by the suppliers. The resulting fragments were separated by electrophoresis in either polyacrylamide or agarose gels (18). In some cases, the specific fragment(s) carrying the tRNA gene(s) was identified by the method described by Southern (19) using \textsuperscript{32}P-tRNA\textsubscript{Val} as the probe (18).

**Sequence Analysis**—DNA sequencing was performed as described by Maxam and Gilbert (20) using [\textsuperscript{3}P]dideoxynucleotide triphosphates and the Klenow fragment of DNA polymerase to label the restriction enzyme-generated oligonucleotides. Gevaert Curix x-ray film was used for autoradiograms using DuPont Cronex Lightning-Plus intensifying screens when necessary.

**RESULTS**

It is not known if these tRNA\textsubscript{Val} genes are transcribed in vivo nor where the origins or terminations of transcription may be. Throughout the text, the term structural gene refers to a sequence of DNA that could be transcribed into a tRNA\textsubscript{Val}.

**In Situ Hybridization**—Under stringent conditions, a plasmid with a Drosophila tRNA gene insert hybridizes essentially to one site on the polytene chromosomes whereas the tRNA itself typically hybridizes to a number of sites. Thus, pDt55 was previously demonstrated (18) to hybridize to the left arm of the third chromosome, at 70BC while pDt23, which was shown to be identical with pDt14, hybridized to 89BC on the right arm of the chromosome. All subsequent studies were done using pDt14. In the present study, pDt92R and pDt120R were found to hybridize to 90C (Fig. 1); tRNA\textsubscript{Val} hybridizes to all three sites as well as to 50D (2). We have not yet isolated a plasmid that hybridizes to this latter site.

**Restriction Enzyme Analysis**—The limits of the inserted DNAs in the four plasmids are the HindIII sites. All these inserts have a large number of sites which are cleaved by...
other restriction enzymes, and as many as 15 enzymes were tested. The fragments bearing the tRNA genes were detected using the Southern technique (6). The organization of the genes in the inserts is shown diagramatically in Fig. 2. The polarity of the tRNA genes is indicated by arrows with the arrow heads at the 3' ends of the genes.

The tRNA genes contain the sequence CCCGGG, which is cleaved by the endonuclease XmaI. pDt92R and pDt120R each contain one such site in their single tRNA genes; pDt14 has three XmaI sites, one in the tRNA gene, one in a tRNA gene, and the third about 3 kilobase pairs to the left of the tRNA gene. pDt55 has two XmaI sites; one in each of the tRNA genes. The tRNA genes all occur well inside the DNA segments studied. This has allowed us, in most cases, to obtain sequences of at least 100 base pairs on each side of the structural genes.

**Sequences**—The nucleotide sequences of segments of the 

**Drosophila DNA inserts bearing the tRNA genes are shown in** Fig. 3. Only one strand of each segment is shown. It is written in the conventional 5' to 3' direction and except where genes occur with opposite polarity, it is the noncoding strand.

**Nucleotide Sequences of pDt92R and pDt120R**—These two plasmids differ at only eight sites in the 506 base pairs between the HindIII site at the left end of each insert and a HindIII site on the right end of the pDt92R insert. pDt120R lacks this second HindIII site, and its insert is about 1500 base pairs longer. Starting at position 513, it has the sequence AACCTT, which in pDt92R apparently is AAGCTT, the HindIII recognition sequence. The sequence differences between these two inserts lie on both sides of the structural gene, and the differences in pDt92R are noted in Fig. 2 by letters under the sequence of pDt120R.

The structural genes in pDt92R and pDt120R which hybridize with tRNA genes are identical, but their sequence differs from that expected from the tRNA sequence at four sites overlined in the figure: at residue 232, a T instead of a C, position 16 of the structural gene (in the D loop); at residue 245, a C (position 29) and, at 257, a G (position 41) instead of the expected T and A (these code for a base pair in the middle of the anticodon stem); and at residue 273, an A (position 57) instead of a G. It has the expected AAC sequence in the anticodon region, and in common with other eukaryotic tRNA genes it does not code for the 3'-terminal CCA sequence.

**Nucleotide Sequence of pDt14**—The Drosophila DNA in this plasmid also hybridizes with tRNA Val, but the nucleotide sequence of the structural gene is identical to that in pDt92R and pDt120R. Upstream 214 base pairs and with the same polarity as the tRNA Val gene is another tRNA gene. It corresponds to the sequence found for tRNA Val (22). This gene contains 23 base pairs (from 43 to 65) identical with a segment in the tRNA Val gene (from 44 to 66) except for the sixth base pair from the 5' end of the segments. The XmaI sites occur in these segments.

Upstream from the 5' end of the tRNA Val genes there is very little similarity between pDt14 and pDt92R or pDt120R, with the exception that at -28 in pDt14 there is the sequence GAAAA which also occurs in pDt92R at -19. In addition, pDt14 has a 10-base pair sequence TACTTACA at -110 (nucleotide 302). This sequence is found in pDt92R and pDt120R at -116 (nucleotide 101). Immediately beyond the 3' end of pDt92R and pDt120R is a sequence, GTTGGAA, which is also found in the corresponding pDt14
Drosophila tRNA\textsuperscript{Val} Genes

The Nucleotide Sequence of pDt120R with the Differences in pDt92R Under the Sequence

![Nucleotide Sequence](http://www.jbc.org/Downloaded from)

Gene as TTGGA. Beyond these sequences are A-T rich regions which contain blocks of Ts that are thought to be involved in the termination of transcription (23). In pDt14, there is an inverted complementary sequence 11 base pairs long at 71 and 309.

Nucleotide Sequence of pDt55—This segment of Drosophila DNA is significantly different from those just discussed. It contains two identical tRNA\textsuperscript{\textsc{\textalpha}} genes of opposite polarity, 525 base pairs apart. The nucleotide sequence of these genes corresponds to that found in tRNA\textsuperscript{\textsc{\textalpha}} (21). The sequence homology between the two genes extends seven base pairs beyond the 5' ends and six base pairs beyond the 3' ends of the structural genes. In fact, the loop-outs, the homology extends by 24 base pairs beyond the 3' end. An additional feature of the sequence is the possibility of forming hairpin structures upstream from the 5' ends of two structural genes. The stems of these hairpins are 13 and 21 base pairs long, respectively. If loop-outs and G-T pairs are included, these structures extend upstream from base pair -15 and -12, respectively. Upstream at -52 in the first gene there is a sequence GCCAGTT which was also found upstream at -20 in Drosophila tRNA\textsuperscript{\textsc{\textalpha}} genes (24). In the second gene, the corresponding sequence (at -67) is GCCACTT. At present, no function can be assigned to these sequences.

**Discussion**

The transfer RNA genes of Drosophila are scattered across the chromosomes in a number of different organizational patterns. Most individual tRNAs hybridize to more than one site on the chromosomes, and at any one site there may be several copies of the gene. These may be intermixed with genes for other tRNAs. The genes may have similar or opposite polarities. Yen and Davidson (11) have carried out the most extensive analysis to date on a cluster of Drosophila tRNA genes. They found 18 genes occurring in a 46-kilobase pair section of a 94-kilobase pair long segment of DNA which hybridizes to region 42A on the second chromosome. This segment of DNA served as the source of the 9.3-kilobase pair insert of the plasmid pCIT12 (10), which contains eight of the genes. These were sequenced by Silverman et al. (25) and Hovemann et al. (12) and found to be one gene for tRNA\textsuperscript{\textsc{\textalpha}}, three for tRNA\textsuperscript{\textsc{\textbeta}}, one for tRNA\textsuperscript{\textsc{\textgamma}} and three for tRNA\textsuperscript{\textsc{\textdelta}}. A number of these had opposite polarities. The additional 10 genes which lie outside the pCIT12 segment hybridize with tRNAs for asparagine, lysine, or arginine. The genes occur in small clusters randomly distributed along the DNA (11). The 42A region appears to be particularly rich in tRNA genes.

Genes studied here are not as tightly clustered. For example, the plasmid pDt14, which contains a 16.8-kilobase pair insert arising from 89BC, appears to have only one gene for tRNA\textsuperscript{\textsc{\textgamma}} and one for tRNA\textsuperscript{\textsc{\textalpha}} located 214 base pairs apart near the center of the insert. Hybridization of fragments of this plasmid produced by HindIII, DdeI, RsaI, TaqI or BglII with 13\textsuperscript{\texttextsc{\textalpha}}-RNA (total 4S-RNA) gave no evidence for the presence of other genes. However, since the 89BC region was found to contain four or more genes for 4S-RNA (4), additional genes may lie outside this cloned DNA segment.

The occurrence of a 23-base pair sequence common to the tRNA\textsuperscript{\textsc{\textalpha}} and tRNA\textsuperscript{\textsc{\textgamma}} genes of pDt14 suggests that at least this segment of each gene arose from a common ancestral gene. A striking example of the maintenance of sequence fidelity in genes with the tRNA\textsuperscript{\textsc{\textalpha}} genes of pDt92R, pDt120R and pDt14. On the chromosome, these genes occur many tens of thousands of base pairs apart and yet have identical sequences. The mechanism by which the organism maintains duplicate sequences is unknown.

The inserted sequences of pDt92R and pDt120R are very similar but not identical over 500 base pairs. It may be supposed that these plasmids were generated because the
DNA used was obtained from a nonisogenic stock of Drosophila melanogaster (Oregon R). The variations in nucleotide sequence outside the structural genes may represent differences in the DNA of homologous chromosomes or differences between repeated sequences in the DNA of the same chromosome. At present, we do not know which alternative is correct.

Of the four plasmids containing coding sequences for tRNAVal, only pDt55 appears to be transcribed in vivo. The reasoning behind this statement is as follows. On RPC-5 columns at pH 4.5, valine isoaccepting activities can be resolved into at least seven peaks. However, the major activities fall in only two peaks, labeled tRNAVal and tRNAVal'. The tRNAVal fraction has been subdivided into tRNAVal' and tRNAVal'' (26). The three major tRNAsVal have different nucleotide sequences and the genes in pDt55 correspond to the tRNAVal'' sequence (21). The three tRNAsVal also have different codon preferences and together they recognize all four valine codons (26). Therefore, these three tRNAs should be sufficient for growth and development. The minor species may represent small changes in modification of the bases. However, none of these peaks has been characterized, because of the paucity of material for such investigations. The pattern of RPC-5 chromatography does not appear to change for the valine tRNAs as a function of development at the level of the whole organism (3). In determining the sequence of tRNAVal, no evidence was seen that it contained any heterogeneity (21) nor was there any indication that appreciable amounts of tRNAs of closely related sequence had been separated from it. Therefore, it appears that the structural genes coding for tRNAVal are represented by pDt55 which hybridizes to one of the known major sites of tRNA hybridization, 70BC. The tRNAVal genes in plasmids pDt14, pDt92R, and pDt120R hybridizing at 89BC and 90C appear not to be transcribed, as the tRNA sequences encoded in them have not been identified by RPC-5 chromatography. One could also argue that the genes on plasmids pDt14, pDt92R, and pDt120R are expressed at very low levels and consequently are difficult to detect by RPC-5 chromatography. Either case is interesting. It should be noted that the genes in pDt14, pDt92R, and pDt120R have been cloned in several independent “shot-gun” recombinant DNA experiments almost as frequently as those in pDt55 (18). Thus, the same sequence which is encoded in at least three different genes, represented by pDt14, pDt92R, and pDt120R, is readily isolated by recombinant DNA techniques yet is expressed at low levels or not at all during several stages of Drosophila development. It is probably significant that two of the differences between the structure of the tRNA and of these genes involve the substitution of a complementary base pair in what would be the anticodon stem of a tRNA product of transcription. This suggests that a functional tRNA may be produced under some conditions. The apparently paradoxical nature and function of these sequences remain to be explored.

REFERENCES

1. Gall, J. G., and Pardue, M. L. (1969) Proc. Natl. Acad. Sci. U. S. A. 63, 378–383
2. 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
3. White, B. N., Tener, G. M., Holden, J., and Suzuki, D. T. (1973) Dev. Biol. 33, 185–195
4. Elder, R. T., Uhlenbeck, O. C., and Szabo, P. (1980) in Transfer RNA: Biological Aspects (Soll, D., Abelson, J. N., and Schimmel, P. R., eds) pp. 317–323, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
5. Kubli, E., and Schmidt, T. (1978) Nucleic Acids Res. 5, 1465–1478
6. Dunn, R., Hayashi, S., Gillam, I. C., Delaney, A. D., Tener, G. M., Grigliatti, T., Kaufman, T. C., and Suzuki, D. T. (1979) J. Mol. Biol. 128, 277–287
7. Elder, R. T., Szabo, P., and Uhlenbeck, O. C. (1980) J. Mol. Biol. 142, 1–17
8. Schmidt, T., and Kubli, E. (1980) Chromosoma 80, 277–287
9. Weber, L., and Berger, E. (1976) Biochemistry 15, 5511–5519
10. Yen, P. H., Sodja, A., Cohen, M., Jr., Conrad, S. E., Wu, M., Davidson, N., and Igen, C. (1977) Cell 11, 763–777
11. Yen, P. H., and Davidson, N. (1980) Cell 22, 137–148
12. Hovemann, B., Sharp, S., Yamada, H., and Soll, D. (1980) Cell 19, 889–896
13. Hosbach, H. A., Silberklang, M., and McCarthy, B. J. (1980) Cell 21, 169–178
14. Hershey, N. D., and Davidson, N. (1980) Nucleic Acids Res. 8, 4899–4910
15. Robinson, R. R., and Davidson, N. (1981) Cell 23, 251–259
16. Gergen, J. P., Loewenberg, J. Y., and Wensink, P. (1981) J. Mol. Biol. 147, 475–499
17. Endow, S. A., and Roberts, R. C., Jr. (1977) J. Mol. Biol. 112, 521–529
18. Dunn, R., Delaney, A. D., Gillam, I. C., Hayashi, S., Tener, G. M., Grigliatti, T., Misra, V., Spurt, M. G., Taylor, D. M., and Miller, R. C. Jr. (1979) Gene 7, 197–215
19. Southern, E. M. (1975) J. Mol. Biol. 98, 503–517
20. Maxam, A. M., and Gilbert, W. (1979) Proc. Natl. Acad. Sci. U. S. A. 74, 560–564
21. Addison, W. R., Gillam, I. C., and Tener, G. M. (1982) J. Biol. Chem. 257, 674–677
22. Altwegg, M., and Kubli, E. (1979) Nucleic Acids Res. 7, 93–105
23. Korn, L. J., and Brown, D. D. (1978) Cell 15, 1145–1156
24. Hovemann, B., Schmidt, T., Yamada, H., Silverman, S., Mao, J., DeFranco, D., and Soll, D. (1980) in Transfer RNA: Biological Aspects (Soll, D., Abelson, J. N., and Schimmel, P. R., eds) pp. 325–328, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
25. Silverman, S., Schmidt, T., Soll, D., and Hovemann, B. (1979) J. Biol. Chem. 254, 10290–10294
26. Dunn, R., Addison, W. R., Gillam, I. C., and Tener, G. M. (1978) Can. J. Biochem. 56, 618–623
The structures of genes hybridizing with tRNA4Val from Drosophila melanogaster.

W R Addison, C R Astell, A D Delaney, I C Gillam, S Hayashi, R C Miller, B Rajput, M Smith, D M Taylor and G M Tener

J. Biol. Chem. 1982, 257:670-673.