Homology maps of the *Drosophila* α -tubulin gene family: one of the four genes is different

Howard J. Baum, Yehudah Livneh* and Pieter C. Wensink

Department of Biochemistry and Rosenstiel Center, Brandeis University, Waltham, MA 02254, USA

Received 18 March 1983; Revised and Accepted 22 July 1983

ABSTRACT

We describe the intron-exon structure of and the homology among the four α -tubulin genes of *Drosophila melanogaster*. Three of the genes share a highly conserved 1.3 kb sequence which corresponds to most of the RNA complementary portion of the genes. The fourth gene is different. Its 5' half has weak homology and its 3' half has moderate homology to the other three genes. The homology maps were first determined by electron microscopy of heteroduplexes between pairs of genes. Higher resolution maps were then obtained by gel analysis of heteroduplexes that had been digested with S1 nuclease.

INTRODUCTION

Tubulin, a dimeric protein composed of one α and one β subunit, is a major structural component of eukaryotic cells. Although the amino acid sequence of each subunit has been highly conserved through evolution (1-5), there is often minor sequence heterogeneity among α or β subunits from an individual organism (4-10). In all cases, this heterogeneity involves only one or a few amino acids.

Tubulin subunit heterogeneity, coupled with the wide variety of microtubular structures within individual organisms, has led to speculation that different microtubular structures may contain different α or β subunits (11). Structure-specific tubulins found in *Naegleria* flagella (6) and *Drosophila* sperm (7-8) support this hypothesis. Thus specific microtubular structures may be due to cellular control of the relative concentrations of different α - and β -tubulins. Such control may occur by post-translational modification of tubulins (12-13). It also appears that such control may occur by transcription of alternative tubulin genes (6,9,10,14,15).

The α -tubulin gene family of *Drosophila* is a likely candidate for transcriptional control of different tubulin structures. There are four *Drosophila* α -tubulin genes (15-18), each with a different chromosomal location (α 1 at 84B3-C8; α 2 at 85E6-15; α 3 at 84D5-8; and α 4 at 67C4-6; (17-18)). We have established that each gene is complementary to α -tubulin mRNA and is

transcribed to produce an RNA that occurs on polysomes (15,17). Therefore, we presume that each encodes α -tubulin. We have also shown that the relative concentrations of transcripts from three of the four genes vary in a tissue and/or developmentally specific manner (15). This indicates a potential for transcriptional control of alternative α -tubulins.

In this paper we describe electron microscopic and S1 nuclease mapping of heteroduplexes among these four genes and between the genes and their transcripts. The results give the first maps of homologies among all members of an α -tubulin gene family. These maps indicate that three genes are very similar to each other. A fourth gene is weakly homologous to the other three in its 5' half and is moderately homologous to the other three in its 3' half.

MATERIALS AND METHODS

Plasmid DNA and *D. melanogaster* RNAs were isolated as described by Kalfayan and Wensink (15). Poly A⁺ RNA from adult flies was a gift from Paul Mitsis.

Agarose gel electrophoresis was carried out as described by Hung *et al.* (19). Polyacrylamide gel electrophoresis was performed according to the method of Sanger and Coulson (20).

DNA restriction fragments used for the mapping experiments were radiolabelled at their termini as described by Hung *et al.* (19). DNA molecular weight standards were radiolabelled using T4 DNA polymerase (21).

Heteroduplexes were formed and prepared for electron microscopy essentially by the method of Davis *et al.* (22). The hybridized DNA was spread under 50% formamide hyperphase and 10% formamide hypophase conditions. R-loops of α_1 , α_2 , and α_3 were formed and prepared for electron microscopy according to Hung *et al.* (19). Where indicated, the R-loops were treated with 0.1 mg/ml RNase A for 5 minutes at 25°C. In reporting lengths, " $\pm$ " indicates the sample standard deviation.

The nuclease protection mapping of RNA on α_4 was done essentially as described by Hung *et al.* (19). During hybridization, the molar ratio of DNA to its RNA complement was approximately 3:1, assuming that tubulin mRNA is 0.1% of total mRNA. Hybridization occurred at 50°C for 20-26 hr.

S1 heteroduplex mapping was carried out as follows. A restriction digest of the unlabelled gene was mixed with a ³²P-end-labelled gene restriction fragment to give a final concentration of 10 μ g/ml unlabelled DNA and a 1:25 to 1:70 molar ratio of labelled:unlabelled DNA. The DNA mixture was denatured

and heteroduplexes formed according to the method of Davis *et al.* (22). It was necessary to adjust the DNA concentrations and the length of hybridization time (60 min) in order to increase heteroduplex formation and decrease heteroduplex strand displacement by the unlabelled DNA. Also, it was often useful to reduce the strand displacement reaction by digesting the unlabelled DNA with restriction enzymes that reduced the length of single strands in the heteroduplex. The hybridization was stopped and S1 nuclease digestion (45°C, 30 min) performed according to Berk and Sharp (23). The digested heteroduplex samples were split in half and each half was either electrophoresed in a neutral agarose gel or heat-denatured and then electrophoresed in an alkaline agarose gel.

RESULTS

General strategy for mapping α -tubulin gene homologies

DNA sequence homologies among all four genes were mapped by a two step procedure. First, homologous regions were located by electron microscopy of heteroduplexes between each pair of genes. Then, a higher resolution method was used to map S1 nuclease resistant heteroduplexes relative to restriction endonuclease sites.

Figures 1 and 2 show representative electron micrographs of the heteroduplex molecules observed. Each heteroduplex had a single continuous region of homology. Those among $\alpha 1$, $\alpha 2$, and $\alpha 3$ were approximately 1.3 kb in length. Their lengths and locations are summarized in Figure 3. Our earlier Southern blot experiments (17) demonstrated that homology between these three genes and $\alpha 4$ is weak. In agreement with this observation, heteroduplexes between $\alpha 4$ and each of the other genes were rare and the homologous regions had considerable variance in length (the $\alpha 1:\alpha 4$ heteroduplex was 1.1 ± 0.1 kb (Fig. 2B); $\alpha 2:\alpha 4$, 1.4 ± 0.4 kb; $\alpha 3:\alpha 4$, 1.8 ± 0.5 kb) and in location (the average standard deviation for localizations was ± 0.5 kb).

More precise lengths and localizations were determined by a modification of the Berk and Sharp (23) method for mapping DNA:RNA duplexes. This modification gives high resolution maps of DNA:DNA heteroduplex regions that are resistant to the single-strand specific nuclease, S1. In this method a restriction enzyme site in a region of homology is end labelled with ^{32}P and then heteroduplexed by annealing to a large excess of unlabelled DNA from another gene. As diagrammed in Figure 4, the resulting heteroduplexes are digested with S1 and then the homologous DNA on the labelled strand is mapped by gel electrophoresis and autoradiography. S1 sensitive interruptions in the

Figure 1. Heteroduplexes between $\alpha 1$ and either $\alpha 2$ or $\alpha 3$.

In each case the heteroduplex to the left is between the 3' ends of the genes and the one to the right is between the 5' ends. Schematic drawings of the heteroduplexes are shown beneath each micrograph. The thick solid lines represent vector sequences and the thin lines represent *Drosophila* segments. The single-stranded regions are depicted as non-horizontal lines. All lengths are in kb. (A) Electron micrographs of the heteroduplexes formed between $\alpha 1$ and $\alpha 2$. The measurements are from 23 and 16 molecules respectively for the heteroduplexes on the left and right. (B) Electron micrographs of the heteroduplexes formed between $\alpha 1$ and $\alpha 3$. The measurements for the heteroduplexes on the left are from 27, 26, and 27 molecules for the long and short single stranded regions and the duplexed regions respectively. The measurement for the heteroduplex on the right are from 20 molecules.

Figure 2. Heteroduplexes between $\alpha 2$ and $\alpha 3$ and between $\alpha 1$ and $\alpha 4$. The symbols in this figure are the same as Figure 1. (A) Electron micrographs of the heteroduplexes formed between $\alpha 2$ and $\alpha 3$. The measurements for the heteroduplexes on the left are from 22 and 12 molecules for the single and double-stranded regions, respectively. The measurements for the heteroduplexes on the right are from 23, 28, and 28 molecules for the long and short double-stranded regions and single-stranded regions respectively. (B) Electron micrographs formed between $\alpha 1$ and $\alpha 4$. The micrograph on the right shows the end of the $\alpha 4$ single-stranded region pairing with the homologous vector region of $\alpha 1$. This is due to the fact that the region between the Bam HI and Hind III sites of the vector pBR322 is located at the end of the single stranded region of $\alpha 4$. The schematic drawing shows one of the two possible interpretations of the heteroduplex between $\alpha 1$ and $\alpha 4$. The measurements are from 18 molecules.

Figure 3. Summary of the homology between $\alpha 1$, $\alpha 2$, and $\alpha 3$ as determined in the electron microscope.

The restriction endonuclease sites are ($\blacklozenge$) Hind III; ($\circ$) Xba I; ($\blacksquare$) Xho I; ($\blacktriangle$) Eco RI; ($\blacksquare$) Bam HI; ($\blacklozenge$) Ava I; ($\blacktriangledown$) Kpn I. Only the *Drosophila* DNA inserts are shown. The vectors are pBR322 for $\alpha 1$ and $\alpha 4$, and pBR325 for $\alpha 2$ and $\alpha 3$. If a restriction enzyme site appears on a map, all the sites recognized by that enzyme are shown. The region of homology detected in the electron microscope is represented by the open boxes and the standard error is represented by horizontal lines.

labelled strand homology are detected by a combination of denaturing and non-denaturing gels (Fig. 4b). Since this method only maps homology in the labelled strand, separate experiments are necessary to map homology of each heteroduplex strand. However, the location of homology on the unlabelled strand can be determined by measuring changes in the heteroduplex length that are caused by different restriction digestions of the unlabelled DNA. Labelled DNA fragments used in our experiments and a summary of the results are shown in Figure 5. Steps in determining these results are described below.

Homology among the 3' ends of α -tubulin genes

Restriction fragment F (Fig. 5) from the 3' end of $\alpha 3$ was used to map homology to the other three genes. An alkaline gel (Fig. 6a, lane 2) shows that, after S1 digestion, a heteroduplex between fragment F and the $\alpha 1$ gene contains an $\alpha 3$ strand that extends 560 nucleotides downstream (in the 3' direction) from the end-labelled Stu I site. A less intense band, 600 nucleotides long, also occurs and is due either to secondary structure in the

Figure 4. Scheme for the high resolution heteroduplex mapping. Heteroduplexes are formed between end-labelled DNA and unlabelled DNA. * represents the end-label (^{32}P). In part a, heteroduplexes are formed between two genes that have a continuous region of homology while in part b, one of the genes has an interruption in its homology. The straight horizontal lines represent homologous DNA and the wavy lines represent nonhomologous DNA. Homoduplexes are also formed but they are not shown because either they are unlabelled or are the same as the original, labelled molecule. The heteroduplexes are digested with S1 nuclease and electrophoresed on either non-denaturing (N) or denaturing (D) gels. The expected sizes of the heteroduplexes on both gels are shown.

DNA or to partial homology extending 50 nucleotides further downstream. Examination of the same products on a neutral gel gave the same results, indicating that there is no S1 sensitive interruption in the labelled $\alpha 3$ strand of the heteroduplex. The location of this homology within the unlabelled $\alpha 1$ strand of the heteroduplex was determined by digesting the $\alpha 1$ gene before forming heteroduplexes. Xho I digestion reduced the heteroduplex length to 400 nucleotides (Fig. 6a, lane 3) and Sal I digestion reduced it to 230 nucleotides (Fig. 6a, lane 4). These results are diagrammed in Figure 5.

Several other bands that are characteristic of this technique also appear in the alkaline gel of Figure 6a. A band of 1150 nucleotides is present in all lanes. This is a homoduplex of fragment F because it is the same length as fragment F (Fig. 6a, lane 15) and is formed during annealing in the absence

Figure 5. Summary of the homology among all four α -tubulin genes and the location of their homologous RNA's.

The subclones and restriction endonuclease sites shown are the same as in Figure 3 with the following additional restriction endonuclease sites: (∇) Pvu II; (Φ) Sal I; ($\circ$) Pst I; ($\top$) Stu I. If a restriction enzyme site appears on a clone map, all the sites recognized by that enzyme are shown except for the Eco RI sites on $\alpha 4$. The regions of strong homology among the tubulin genes, as determined by the high resolution technique described in Figure 4, are represented by the filled boxes. $\alpha 1$, $\alpha 2$, and $\alpha 3$ share the same homologous region. The $\alpha 4$ homology is only mapped against $\alpha 1$. Its homology is indicated by the cross hatched box. The fragments used to determine the homologies are indicated by lines A-J. The fragments are end-labelled at the tail of the arrow with the vertical line representing a 5' end-label and the open box and standard deviations are represented by the horizontal lines. The $\alpha 1$, $\alpha 2$, and $\alpha 3$ RNA maps were determined using the electron microscope. The $\alpha 4$ RNA map error was ± 5 nucleotides, so no error bars are shown. Fragments H, I, and J were used to construct the $\alpha 4$ RNA map using nuclease protection methodology.

Figure 6. High resolution mapping of the homology among the 3' ends of the α -tubulin genes.

(a) Alkaline gel of heteroduplexes between 5' end-labelled fragment F (Stu I+Eco RI α 3, see Fig. 5) and various restriction digests of the other tubulin genes. The unlabelled DNA restriction fragment used for the heteroduplex is indicated above each lane using the following abbreviations: Hind III (H), Xho I (X), Sal I (S), Eco RI (R), Xba I (Xb), Bam HI (B), Pst I (P), and Stu I (St). Lanes 11 and 12 are homoduplexes with unlabelled α 3. Lane 13 is an end-labelled molecular length standard consisting of four different digests of pBR322 producing bands at 3614, 2936, 1854, 1631, 1426, 1060, 928, 756, 748, 517, 506, 396, 383, 344, 298, 221, 220, 154, 121, 75, and 13 bases. Lane 14 is blank. Lanes 15 and 16 are the native and the denatured probe, respectively. Lane 17 is a homoduplex of the probe with no unlabelled DNA present, subject to S1 nuclease digestion. Lane 18 is a mixture of the denatured probe and denatured, unlabelled Eco RI α 3 DNA subject to S1 nuclease digestion. (b) Alkaline gel of heteroduplexes between 3' end labelled fragment H (Xho I+Pvu II α 4) and α 1. The bands at 265 and 165 bases are indicated.

of unlabelled α 1 DNA (Fig. 6a, lane 17). Another characteristic band (Fig. 6a, lanes 3 and 4) is produced when unlabelled α 1 DNA is cleaved in the homology region and then heteroduplexed to labelled fragment F. All S1 digested heteroduplexes should be shorter than the entire 3' homology region, yet a band of the entire length does occur. This band appears to be from fragment F that has been heteroduplexed to two different α 1 restriction fragments.

The same F fragment was used to map homology between the 3' region of $\alpha 3$ and the $\alpha 2$ and $\alpha 4$ genes. As in the heteroduplexes with $\alpha 1$, fragment F heteroduplexed to $\alpha 2$ or $\alpha 4$ produces a band of 560 nucleotides on alkaline (Fig. 6a, lanes 6 and 10) and neutral gels (Table I). The map of $\alpha 3$ homology on $\alpha 2$ extends 470 bases from the $\alpha 2$ Xho I site and 240 bases from the Sal I site (Fig. 6a, lanes 8 and 7). Homology between fragment F and $\alpha 4$ was considerably weaker as revealed by a very faint band 560 nucleotides in length (Fig. 6a, lane 10; Table I). These results indicate that the 3' regions of $\alpha 1$, $\alpha 2$, and $\alpha 4$ have the same maps on $\alpha 3$.

Other experiments mapped the $\alpha 1$, $\alpha 2$ and $\alpha 4$ strands of heteroduplexes among the 3' regions of the four genes (Table I). For example, a fragment from $\alpha 1$ (fragment A) when annealed to unlabelled $\alpha 2$ or $\alpha 3$ DNA (Table I) confirmed that a similar sequence occurs at or near the 3' end of the $\alpha 1$, $\alpha 2$ and $\alpha 3$ gene. Thus we detected no 3' homology that is peculiar to any pair of these three genes. Because of the common 3' homology among $\alpha 1$, $\alpha 2$, and $\alpha 3$, the $\alpha 4$ homology was examined further only with $\alpha 1$. Heteroduplexes between end-labelled $\alpha 4$ fragment H and $\alpha 1$ DNA revealed homology extending 265 nucleotides downstream from the $\alpha 4$ Xho I site (Fig. 6b, lane 2) and 100 nucleotides (265 minus 165) downstream from the $\alpha 1$ Xho I site (Fig. 6b, lane 4). As in all other experiments with α -tubulin gene 3' homology, the alkaline and neutral gels gave identical results indicating that there is no interruption in the homology. These mapping experiments yield a self-consistent map of 3' homologies among the four genes. This map is diagrammed in Figure 5.

Homology among the remainder of the gene regions

Fragments containing the remainder of $\alpha 1$ and $\alpha 3$ (B and G, Fig. 5) were used to demonstrate 1300 nucleotides of uninterrupted homology between $\alpha 1$ and $\alpha 3$. These fragments gave identical results in alkaline and neutral gels (Table I) indicating the homology diagrammed in Figure 5.

Similar experiments revealed that $\alpha 2$ also has 1300 nucleotides of homology to $\alpha 1$ and $\alpha 3$ (fragments A, B, F, G; Table I). However, this homology is interrupted by a short region of non-homology because when end-labelled $\alpha 2$ fragment C was heteroduplexed with either $\alpha 1$ or $\alpha 3$, and digested with S1, it had a shorter length on an alkaline gel (Fig. 7a, lanes 7, 10, 12, and 14) than on a neutral gel (Fig. 7b, lanes 7, 10, 12, and 14). This interruption was further mapped with 5' end-labelled fragments D1 and D2 which are opposite strands of an $\alpha 2$ restriction fragment (Fig. 5). Both fragments were simultaneously annealed to $\alpha 1$ or $\alpha 3$ producing two S1 resistant hetero-

Table I - Summary of Heteroduplex Results

End-Labelled DNA Fragment	Unlabelled DNA	Alkaline Gel	Neutral Gel	End-Labelled DNA Fragment	Unlabelled DNA	Alkaline Gel	Neutral Gel
A	Eco RI $\alpha 2$	1000	960,860	E	Hind III-Xba I $\alpha 1$	310	-
	Eco RI-Xho I $\alpha 2$	920	880,780		Eco RI $\alpha 3$	310	-
	Eco RI-Sal I $\alpha 2$	670	630,540	F	Hind III $\alpha 1$	600,560	560
	Eco RI $\alpha 3$	1000	960		Hind III-Xho I $\alpha 1$	460	420
	Pvu II $\alpha 3$	500	460		Hind III-Sal I $\alpha 1$	240	220
	Eco RI-Stu I $\alpha 3$	480	445		Eco RI $\alpha 2$	560	560
B	Eco RI $\alpha 2$	580	580		Eco RI-Xho I $\alpha 2$	500	440
	Eco RI-Bam HI $\alpha 2$	580	580		Eco RI-Sal I $\alpha 2$	250	220
	Eco RI-Sal I $\alpha 2$	250	280		Eco RI $\alpha 4$	560	560
	Eco RI-Pst I $\alpha 2$	220	-	G	Hind III $\alpha 1$	790	820
	Eco RI $\alpha 3$	580	580		Hind III-Pvu II $\alpha 1$	550	520
	Eco RI-Pvu II $\alpha 3$	360	350		Hind III-Bam HI $\alpha 1$	450	480
C	Hind III $\alpha 1$	740	1300		Eco RI $\alpha 2$	810	850
	Hind III-Bam HI $\alpha 1$	740	820		Eco RI-Sal I $\alpha 2$	460	520
	Hind III-Sal I $\alpha 1$	650,250	650,250		Eco RI-Pst I $\alpha 2$	430	490
	Eco RI $\alpha 3$	740	1300	H	Hind III-Bam HI $\alpha 1$	265	290
	Eco RI-Pst I $\alpha 3$	740	880		Hind III-Sal I $\alpha 1$	265	290
	Eco RI-Stu I $\alpha 3$	420	450		Hind III-Xho I-Xba I $\alpha 1$	165	190
D1, D2	Hind III $\alpha 1$	490,180	670,590	I	Hind III $\alpha 1$	440	470
	Hind III-Xba $\alpha 1$	490,180	670,590		Hind III-Pvu II $\alpha 1$	230	280
	Hind III-Sal $\alpha 1$	410,180	400,180		Hind III-Sal I $\alpha 1$	92	-
	Eco RI $\alpha 3$	490,180	670,590				

The lengths are in nucleotides. The locations of the end-labelled fragments are shown in Figure 5. The error is approximately ± 50 bases. For fragments A and D, two bands were seen on native gels. The upper band corresponds to the band on the denaturing gel, while the lower band probably results from secondary structure. All the doublets occur when the end-labelled fragment maps close to the interruption on $\alpha 2$ and only involve $\alpha 2$ heteroduplexes near the interruption.

Figure 7. High resolution mapping of the interruption on $\alpha 2$. The symbols are the same as in Figure 6. (a) Alkaline gel of heteroduplexes between fragment C (Xho I+Eco RI $\alpha 2$, see Fig. 5) and $\alpha 1$ or $\alpha 3$. Lane 1 is the same molecular length standard described in Figure 6. Lane 2 is blank. Lanes 3 and 4 are the native and the denatured probe, respectively. Lane 5 is a homoduplex of the probe with no unlabelled DNA present, subjected to SI

nuclease digestion. Lane 6 is a mixture of the denatured probe and denatured unlabelled Eco RI $\alpha 2$ DNA subjected to S1 nuclease digestion. Lanes 15 and 16 are homoduplexes. (b) Neutral gel of the same heteroduplexes and same lane pattern as a. (c) Alkaline gel of heteroduplexes between fragments D1 and D2 (Sal I+Sal I $\alpha 2$) and $\alpha 1$ or $\alpha 3$. Lanes 7 and 8 represent homoduplexes. The short 180 nucleotide product gives a less intense signal as might be expected from the lower thermostability of its shorter heteroduplex and from the preferential loss of short fragments that occurs during vacuum drying of gels prior to autoradiography.

duplexes of 180 and 490 nucleotides as measured on an alkaline gel (Fig. 7c, lanes 2, 6). The 490 heteroduplex contains D2 because $\alpha 1$ DNA digestion with Hind III and Sal I shortens this heteroduplex (Fig. 7c, lane 3). The size of the insertion is deduced from the 740 nucleotide length of the D fragments (Fig. 7c, lane 8) minus the heteroduplex lengths (180 and 490 nucleotides), resulting in 70 nucleotides.

A region of S1 resistant homology shared by $\alpha 4$ and the other 3 genes is 700 nucleotides in length in contrast to the 1300 nucleotide region shared among $\alpha 1$, $\alpha 2$, and $\alpha 3$. This homology extends a short distance upstream of and is continuous with the 3' homology determined in the previous section. This was demonstrated by mapping the $\alpha 4$ I fragment on $\alpha 1$ DNA (Table I). Experiments with other labelled fragments gave some indication of very weak homology further upstream. These results are summarized in Table I and diagrammed in Figure 5.

Location of RNA complementary regions

The poly A⁺ RNA from late embryos was mapped on the $\alpha 1$ gene using the electron microscopic R-looping method of Thomas *et al.* (24). One $\alpha 1$ R-looped molecule is shown in Figure 8a. It has a 1.4 kb single-stranded R loop and an adjacent 0.5 kb double-stranded loop. This double-stranded loop was present in 63% of the R-looped molecules. It indicates (25) that a 0.5 kb intron separates the 1.4 kb exon from an exon that is too small to form a visible R loop. This interpretation is shown in Figure 8b and is supported by our observation that RNase treatment reduced the frequency of the double-stranded loop from 63% to 12%. The $\alpha 1$ transcript map resulting from these R loops is diagrammed in Figure 5.

Since our mapping results demonstrate that the $\alpha 1$, $\alpha 2$ and $\alpha 3$ genes form well base-paired duplexes with each other, we presume that each of the three can form R loops with transcripts from any of the others. The poly A⁺ RNA used for all our R-looping experiments was from late embryos in which $\alpha 1$ mRNA is the major α -tubulin transcript (15). For this reason we presume that all R loops contain $\alpha 1$ mRNA. The $\alpha 2$ and $\alpha 3$ genes gave R loops that were similar to

Figure 8. R loops formed on $\alpha 1$.

(a) Electron micrograph formed when hybridizing poly (A)⁺ RNA from late embryos to $\alpha 1$ DNA. (b) Interpretive drawing of a.

those of $\alpha 1$ with two exceptions. First, no short double-stranded loop was formed at the end of the $\alpha 2$ and $\alpha 3$ R loops. This indicates that if $\alpha 2$ and $\alpha 3$ have upstream exons, either the exons are too short to form stable R loops under these conditions or the exons are not present within the upstream DNA included in the cloned genes. The second exception is that $\alpha 2$ formed two large R loops separated by a short double-stranded region. This region maps at the site within $\alpha 2$ that is not homologous to $\alpha 1$ or $\alpha 3$. These results are diagrammed in Figure 5.

As might be expected from the low homology between $\alpha 1$ and $\alpha 4$ and from the lack of $\alpha 4$ transcripts in late embryos (15), no R loops between late embryonic RNA and $\alpha 4$ were observed. Therefore the location of RNA complementarity in $\alpha 4$ was determined by nuclease protection methodology using RNA from adult flies which have abundant $\alpha 4$ transcripts (15). The 3' end of the transcript was mapped with fragment H (Fig. 5). Digesting the RNA:DNA hybrid with S1 or Exo VII nuclease produced bands of approximately the same length (Fig. 9a), indicating that there are no additional exons within the downstream 1.7 kb. The rest of the transcript was mapped with fragment I demonstrating a total uninterrupted transcript length of 1.5 kb. Fragment J was used to map the 5' end of this transcript with greater precision. Both S1 and Exo VII treated RNA hybrids with fragment J gave bands migrating at the same length (Fig. 9b) indicating that there are no additional upstream exons within 1.9 kb. A tenfold molar excess of unlabelled $\alpha 1$, $\alpha 2$, $\alpha 3$, or $\alpha 4$ DNA during hybridization

Figure 9. Map of the RNA transcript on $\alpha 4$.

The S1 or Exo VII (X7) resistant material, as indicated above each lane, was electrophoresed on a 4% polyacrylamide sequencing gel. Lane M is the same molecular length standard described in Figure 6. Note that the S1 and Exo VII digested hybrids migrate at slightly different molecular lengths because extensive Exo VII digestion leaves 2 to 3 nucleotide single-stranded tails whereas S1 degrades ends (19,26). (a) Location of the 3' end of the $\alpha 4$ transcript. Fragment H (3' end-labelled Xho I+Pvu II, see Fig. 5) was used to map the 3' end. (b) Location of the 5' end of the $\alpha 4$ transcript. Fragment J (5' end-labelled Ava I+Bam HI $\alpha 4$) was used to map the 5' end. Four of the DNA-RNA duplexes, as indicated, were formed in the presence of a 10 fold molar excess of unlabelled DNA. The symbols for the restriction digests of the unlabelled DNA are the same as in Figure 6.

indicated that this hybrid was between $\alpha 4$ DNA and $\alpha 4$ transcript (Fig. 9b). Figure 5 summarizes this RNA map.

DISCUSSION

Drosophila has one different and three similar α -tubulin genes

Sequence homology among the four *Drosophila* α -tubulin genes was first mapped by electron microscopy. Although the electron microscopic maps (Fig. 3) determined double-stranded heteroduplex lengths to within 100 nucleotides, the large variance in single-strand lengths led to homology localizations that were insufficient to determine whether regions homologous between two genes

were also homologous to a third gene. Furthermore, as our subsequent results demonstrated, electron microscopy had insufficient resolution to reveal a short interruption that occurs in the homology between genes.

Because of these resolution limits to the electron microscopic method, we also mapped homologies by an S1 nuclease protection method. The S1 method has its resolution limited by the specificity of S1 nuclease and by the resolution of the gel system used to examine protected fragments. Therefore, unlike the electron microscopic method, end points of a region of homology could be determined to within one or a few nucleotides (27). Also, short, poorly base-paired or completely unpaired regions within a longer region of homology could be detected. The limits of this detection have been studied by Dodgson and Wells (28) who have shown that although sensitivity to S1 digestion of mismatched portions of a duplex is influenced by local (A + T) composition, the nuclease efficiently makes double-stranded breaks only when a mismatch is 3 or more nucleotides in length. Because our digestion conditions were similar to those of Dodgson and Wells, we presume that homologies detected by this method differ by no more than isolated nucleotides or pairs of nucleotides.

Both electron microscopy and S1 nuclease mapping revealed 1.3 kb of homology among the $\alpha 1$, $\alpha 2$, and $\alpha 3$ genes. The homology regions of $\alpha 1$ and $\alpha 3$ are continuous, but that of $\alpha 2$ has a 70 nucleotide interruption 500 nucleotides from its 5' end (Fig. 5). As will be discussed below, these 1.3 kb regions appear to correspond to the three α -tubulin structural genes. Earlier restriction endonuclease site mapping of these genes (17) demonstrated that there are sequence differences among these homology regions. Given the mispairing sensitivity of S1 mapping and the general sequence conservation among α -tubulins (1-5), it is likely that most differences between these genes are at wobble positions in codons and that the protein products have few differences.

Although three of the *Drosophila* α -tubulin genes are very similar, $\alpha 4$ is different. It has less homology to the other three genes. The S1 mapping experiments revealed only one 700 nucleotide homology region at the 3' end of the 1.3 kb regions of $\alpha 1$, $\alpha 2$ and $\alpha 3$ (Fig. 5). Also, the heteroduplexes formed were less abundant than the 1.3 kb heteroduplexes among the other genes (Fig. 6). These observations indicate that $\alpha 4$ homology is not only shorter but also less homologous to the other genes. Electron microscopy (Fig. 2) indicates that the $\alpha 4$ homology extends many hundreds of nucleotides upstream and probably includes the remainder of the 1.3 kb region (Fig. 5). The

inability to detect this additional homology with the S1 technique using fragments end-labelled at several sites within this region (unpublished observation) suggests that the 5' half of $\alpha 4$ contains many clustered mismatches to the other genes. It is likely that this upstream homology is detected by electron microscopy only because the microscope allows visualization of duplexes with mismatches as long as 50 nucleotides. We conclude that the 5' half of the $\alpha 4$ gene is weakly homologous to the other three genes and the 3' half is more homologous to these genes.

Mapping RNA complementary regions of the α -tubulin genes

The RNA complementary regions were mapped and found to coincide with the DNA homology regions described above. The RNA complementarity of $\alpha 1$, $\alpha 2$, and $\alpha 3$ were mapped by an electron microscopic R loop method. In the 5' to 3' direction, $\alpha 1$ has a small (<50 nucleotide) exon, a 500 nucleotide intron and a 1.4 kb exon. Because of the relative abundance and shared homologies of the RNAs, the $\alpha 2$ and $\alpha 3$ RNA complementary regions are most likely to have been mapped with $\alpha 1$ transcripts and therefore reflect their heteroduplexes with $\alpha 1$ (Fig. 5). The RNA complementary region of the less related $\alpha 4$ gene was mapped with its own transcript by a nuclease protection and DNA competition method. This revealed only a single 1.5 kb exon within a 5.1 kb region (Fig. 5). As diagrammed in Figure 5, the RNA complementary regions coincide with the sequences that are shared between the genes.

The 1.3 kb homology region is likely to encode most if not all of the α -tubulin protein. This region in $\alpha 1$ is complementary to poly A⁺ RNA which can be translated into an α -tubulin protein (Fig. 5) (15). Furthermore, our partial sequencing data (unpublished) indicate that of the two termini of the $\alpha 1$ homology region, one probably begins with the codon for the second amino acid of α -tubulin and the other probably ends with the final amino acid. For this reason we hypothesize that these three regions in $\alpha 1$, $\alpha 2$, and $\alpha 3$ as well as the complementary and partially complementary regions in $\alpha 4$ all code for almost the entire α -tubulin or α -tubulin-like protein.

α -tubulin gene and protein heterogeneity

The *Drosophila* α -tubulin genes are heterogenous. The slight heterogeneity among the $\alpha 1$, $\alpha 2$ and $\alpha 3$ genes is probably similar to the heterogeneity reported for genes encoding the most abundant porcine α -tubulins (4). The $\alpha 4$ gene, however, is substantially different from the other three *Drosophila* genes. In some organisms, heterogeneity within tubulin gene families is due to pseudogenes that do not encode proteins (29). It does not appear likely that any of the four α -tubulin genes of *Drosophila* is a

pseudogene because each produces an RNA that occurs on polysomes and therefore is likely to be translated (15).

The protein encoded by $\alpha 4$ is likely to have an amino-terminal half that is different from the other α tubulins and a carboxy-terminal half that is similar to the other α tubulins. This places the $\alpha 4$ protein in a distinct category. No other study has indicated substantial heterogeneity among α tubulins or among β tubulins. This raises several possibilities. Perhaps other quite different tubulins have escaped detection because, unlike $\alpha 4$ tubulin, most unusual tubulins are of low abundance. It is also possible that $\alpha 4$ is not an α tubulin, but instead is a related protein, perhaps a microtubule associated protein (30) that has some features similar to α -tubulin which allow it to interact with tubulin proteins.

Acknowledgements

We thank Mike Wormington, Chan Fulton, Michael Garabedian, Peggy Harris, Mien-Chie Hung, and Ellen Z. Baum for useful criticism of this manuscript. This work was supported by research grant GM21626 and training grant 5T32-GM07596 from the NIH.

*Current address: Biotechnology General Ltd., Kiryat Weitzmann, Rehovot, Israel

REFERENCES

1. Kirschner, M.W. (1978) *Int. Rev. Cytol.* 54, 1-71.
2. Cleveland, D.W., Lopata, M.A., MacDonald, R.J., Cowan, N.J., Rutter, W.J., and Kirschner, M.W. (1980) *Cell* 20, 95-105.
3. Lemischka, I.R., Farmer, S., Racaniello, V.R., and Sharp, P.A. (1981) *J. Mol. Biol.* 151, 101-120.
4. Ponstingl, H., Krauhs, E., Little, M., and Kempf, T. (1981) *Proc. Nat. Acad. Sci. USA* 78, 2757-2761.
5. Krauhs, E., Little, M., Kempf, T., Hofer-Warbinek, R., Ade, W., and Ponstingl, H. (1981) *Proc. Nat. Acad. Sci. USA* 78, 4156-4160.
6. Lai, E.Y., Walsh, C., Wardell, D., and Fulton, C. (1979) *Cell* 17, 867-878.
7. Kempf, K.J., Raff, R.A., Kaufman, T.C., and Raff, E.C. (1979) *Proc. Nat. Acad. Sci. USA* 76, 3991-3995.
8. Kempf, K.J., Raff, E.C., Raff, R.A., and Kaufman, T.C. (1980) *Cell* 21, 445-451.
9. Silflow, C.D. and Rosenbaum, J.L. (1981) *Cell* 24, 81-88.
10. Raff, E.C., Fuller, M.T., Kaufman, T.C., Kempf, K.J., Rudolph, J.E., and Raff, R.A. (1982) *Cell* 28, 33-40.
11. Fulton, C. and Simpson, P.A. (1976) in *Cell Motility*, Goldman, R., Pollard, T., and Rosenbaum, J. Eds., pp. 987-1005, Cold Spring Harbor Laboratory, New York.
12. Raybin, D. and Flavin, M. (1977) *Biochem.* 16, 2189-2194.
13. Brunke, K.J., Collis, P.S., and Weeks, D.P. (1982) *Nature* 297, 516-518.
14. Alexandraki, D. and Ruderman, J.V. (1981) *Mol. Cell. Biol.* 1, 1125-1137.
15. Kalfayan, L. and Wensink, P.C. (1982) *Cell* 29, 91-98.

16. Sánchez, F., Natzle, J.E., Cleveland, D.W., Kirschner, M.W., and McCarthy, B.J. (1980) *Cell* 22, 845-854.
17. Kalfayan, L. and Wensink, P.C. (1981) *Cell* 24, 97-106.
18. Mischke, D. and Pardue, M.L. (1982) *J. Mol. Biol.* 156, 449-466.
19. Hung, M.C., Barnett, T., Woolford, C., and Wensink, P.C. (1982) *J. Mol. Biol.* 154, 581-602.
20. Sanger, F. and Coulson, A.R. (1978) *FEBS Lett.* 87, 107-110.
21. O'Farrell, P.H., Kutter, E., and Nakanishi, M. (1980) *Molec. Gen. Genet.* 179, 421-435.
22. Davis, R.W., Simon, M., and Davidson, N. (1971) in *Methods in Enzymology*, Grossman, L. and Moldave, K. Eds., Vol 21, pp. 413-428, Academic Press, New York.
23. Berk, A.J. and Sharp, P.A. (1977) *Cell* 12, 721-732.
24. Thomas, M., White, R.L., and Davis, R.W. (1976) *Proc. Nat. Acad. Sci. USA* 73, 2294-2298.
25. Hirsh, J. and Davidson, N. (1981) *Mol. Cell Biol.* 1, 475-485.
26. Changas, G.S. and Wu, R. (1975) *J. Biol. Chem.* 250, 4601-4606.
27. Hung, M.C. and Wensink, P.C. (1981) *Nucl. Acid Res.* 9, 6407-6419.
28. Dodgson, J.B. and Wells, R.D. (1977) *Biochem.* 16, 2374-2379.
29. Lemischka, I. and Sharp, P.A. (1982) *Nature* 300, 330-335.
30. Solomon, F. (1981) *Cold Spring Harbor Symp. Quant. Biol.* XLVI, 17-22.