Cell, Vol. 14, 901-919, August 1978, Copyright © 1978 by MIT

A Novel Arrangement of Tandemly Repeated Genes at a Major Heat Shock Site in D. melanogaster

John T. Lis,* Louise Prestidge and David S. Hogness†
Department of Biochemistry
Stanford University School of Medicine
Stanford, California 94305

Summary

Three cloned segments of Drosophila melanogaster DNA have been isolated that derive from the major heat shock site at 87C in chromosome 3. Each of these segments contains sequences homologous to a class of polysomal poly(A)$^+$ RNAs whose synthesis is induced by heat shock of cultured cells. A combination of R loop, heteroduplex and restriction fragment maps of these segments reveals that their RNA-homologous sequences are arranged in tandemly repeated units, each unit consisting of an α element (0.49 kb) joined to a β element (1.10 kb). The polysomal RNAs homologous to these $\alpha\beta$ units (1.59 kb) are distributed into three size classes exhibiting approximate lengths of 1, 2 and 3 kb. R loop mapping demonstrated that the sequence of the 2 kb RNA is $\alpha\beta\alpha$, indicating that it, and presumably the 3 kb RNA, derive from transcripts covering more than one repeated unit.

One of the cloned segments contains a variant repeat unit, $\alpha\gamma$, located between two $\alpha\beta$ units. This unit has the same α element, but the β element has been replaced by a nonhomologous γ element (0.87 kb). Analysis of the total D. melanogaster DNA indicates that the 87C locus contains at least 21 tandemly repeated units, twelve of which were identified as $\alpha\beta$ units and six as $\alpha\gamma$ units. The 21 or more units are distributed among at least three different tandem arrays separated by different spacer regions, one of which is within a cloned segment. Sequences in the γ element, but not those in the α or β elements, are also found at 87A, which contains another heat shock site. The possible roles of the $\alpha\beta$-type RNAs and of the γ sequences are discussed in the light of these results.

Introduction

Characteristics of the Heat Shock Response

We are interested in hierarchies of genetic regulation in Drosophila melanogaster. The response triggered by heat shock of Drosophila cells appears to represent such a hierarchy. Heat shock is usually achieved in the laboratory by rapidly increasing the cell temperature from 25 to 37°C. The response to this temperature jump appears to include regulatory processes operative at three levels of genetic expression: transcription, translation and RNA processing. Regulation at the transcriptional level was first suggested by site-specific changes in the puff topography of polytene chromosomes in larval salivary glands (Ritossa, 1962; Ashburner, 1970). Within minutes after the temperature jump, new puffs appear at nine sites, while most preexisting puffs regress. This transition in puff topography was subsequently correlated with changes in the sites of ^{3}H-uridine incorporation (Tissières, Mitchell and Tracy, 1974; Bonner and Pardue, 1976) and of RNA polymerase II accumulation (Jamrich, Greenleaf and Bautz, 1977), in a manner indicating that heat shock induces an increased rate of transcription at the new puff sites and decreases the rate at many preexisting puffs.

A parallel transition in the pattern of protein synthesis is observed in a variety of cell types (Tissières et al., 1974; Lewis, Helmsing and Ashburner, 1975; McKenzie, Henikoff and Meselson, 1975; Mirault et al., 1978; Moran et al., 1978). Heat shock induces the synthesis of at least six and perhaps nine distinct polypeptides, and markedly reduces the rate of synthesis of most proteins synthesized at 25°C. This transition results from a very rapid disappearance of the highly heterogeneous population of preexisting polysomes and a more gradual appearance of a new, considerably less heterogeneous population (McKenzie et al., 1975; Moran et al., 1978). Most of the mRNAs in the new population derive from the nine induced chromosomal sites (McKenzie et al., 1975; Spradling, Penman and Pardue, 1975; Bonner and Pardue, 1976; Spradling, Pardue and Penman, 1977) and code for the induced polypeptides (Mirault et al., 1978; Moran et al., 1978). Mirault et al. (1978) also found that an appreciable fraction of the mRNA formed at 25°C is retained in the heat-shocked cells in a form which is efficiently translated in vitro by a rabbit reticulocyte system, although it is inefficiently translated in the heat-shocked cells and virtually absent from their polysomes. It is thus improbable that the sudden disappearance of the preexisting polysomes results from the rapid degradation of their mRNAs. Furthermore, the kinetics of the polysomal transition are not compatible with a simple concentration-dependent competition for the two sorts of mRNA. It therefore appears that heat shock causes a change in translational specificity characterized by the preferential initiation of the new over the old mRNAs.

Finally, heat shock produces an arrest of the

* Present address: Department of Biochemistry, Molecular and Cell Biology, Wing Hall, Cornell University, Ithaca, New York 14853.
† To whom requests for reprints should be addressed.

processing of some transcripts, such as the 135 nucleotide precursor of 5S RNA (Rubin and Hogness, 1975; Jacq, Jourdan and Jordan, 1977) and the large molecular weight precursor of the 18S and 28S rRNAs (Jacq et al., 1977; D. M. Glover, personal communication).

The overall effect of heat shock thus appears to be the induction of a select set of genes, coupled with transcriptional, translational and processing arrests in the expression of most previously active genes. This response, or at least the transitions in puff topography and protein synthesis, can also be triggered at constant temperature (25°C) by an anaerobic pulse (Ashburner, 1970; Lewis et al., 1975), indicating that it does not result simply from the inactivation of an arbitrary collection of heat-sensitive molecules. Because the response is reversible for short-lived perturbations of temperature and oxygen supply, can be triggered by perturbations of amplitudes encountered in nature and appears to be universal with respect to cell type, it has been postulated to be homeostatic (Lewis et al., 1975). Perhaps a temporary and simultaneous arrest of different developmental pathways via the response allows the organism to cope with environmental insults that would otherwise disturb the synchrony among them.

Experimental Plan

These observations and arguments suggest that the response elicited by both temperature and oxygen perturbations is triggered by the same molecular signals which act at the top of a regulatory hierarchy that dictates and coordinates its many facets. This, at least, is the concept which has guided our experimental approach. We have focused our attention on the chromosomal loci where transcription is rapidly induced by heat shock. The general plan is to isolate cloned segments of D. melanogaster DNA (Dm segments) that comprise these loci, and to map the sequences in segments that are homologous to the primary and processed transcripts present in heat-shocked cells. The ultimate goals of these mapping experiments are to define the number and arrangement of the units of transcription at each locus, to determine the routes by which their primary transcripts are processed and, using probes defined by the maps, to analyze the kinetics of the induced transcription and of processing. We are particularly interested in determining whether the units of transcription fall into one or more classes of coordinate induction, and whether the units in such a class share common sequences, particularly at or near their transcription initiation sites.

We have initiated this plan by screening our library of independently cloned Dm segments for those which contain the heat shock-induced genes. The library consists of E. coli K12 clones carrying Col E1 hybrid plasmids (cDm plasmids) which contain shear-generated Dm segments inserted at the Eco R1 site of the Col E1 vector by poly(dA)-poly(dT) connectors (Wensink et al., 1974; D. J. Finnegan, G. M. Rubin, J. Bower and D. S. Hogness, manuscript in preparation). A two-phase screening procedure similar to that described by Young and Hogness (1977) was used to isolate the heat shock genes. In the first phase, clones containing sequences homologous to ^{32}P cytoplasmic poly(A)$^+$ RNA labeled in heat-shocked cells were identified by colony hybridization (Grunstein and Hogness, 1975) with this probe. Separate secondary hybridizations with probes consisting of ^{32}P cytoplasmic poly(A)$^+$ RNA labeled at 25°C and ^{32}P-cRNA transcribed from a cloned rDNA repeat unit (Young and Hogness, 1977) were then used to eliminate the false positives that result from contaminating RNAs equivalent to those in the secondary probes.

In this paper, we characterize three of the cDm clones that survived this screen. Each contains a Dm segment derived from 87C, the locus of a major heat shock-induced puff in the right arm of chromosome 3, and each segment contains sequences homologous to the same kind of heat shock-induced polysomal poly(A)$^+$ RNA. We have mapped these sequences by hybridization of the RNA to an ordered set of restriction fragments from each segment, and by electron microscopy of R loops (White and Hogness, 1977) and of heteroduplexes formed between the segments. These maps reveal an arrangement of the RNA-homologous sequences that is unusual in three respects. First, these sequences are contained in tandemly repeated units that form several clusters, or tandem arrays, within the 87C locus. Second, certain of the polysomal poly(A)$^+$ RNAs span more than one of the tandem repeats, indicating that the transcriptional unit from which they derive includes multiple repeats. Finally, a variant of the standard repeat unit, but not the standard unit itself, contains sequences that are also present at 87A, the site of another major heat shock-induced puff. This provides an interesting correlation with the heat shock gene described in the accompanying paper (Schedl et al., 1978); this gene codes for an mRNA quite different from any of the RNAs described here, yet it also is found at 87C and, in addition, at 87A.

Results

Restriction Fragments Provide Low Resolution Maps of Sequences Homologous to Heat Shock RNA

The three hybrid plasmids selected from among those that survive the screen described in the

Introduction are called cDm702, cDm703 and cDm704, where the numbers designate the Dm segment contained in each plasmid. Quantitative filter hybridization of increasing amounts of each plasmid DNA with a constant amount of total ^{32}P cytoplasmic poly(A)$^+$ RNA labeled in Kc_0 cells during a 70 min period following a 25-37°C temperature jump resulted in the hybridization of 3.1, 3.2 and 4.4% of this RNA at saturating amounts of cDm702, cDm703 and cDm704 DNAs, respectively. By contrast, each DNA hybridized <0.2% of ^{32}P cytoplasmic poly(A)$^+$ RNA labeled in Kc_0 cells for 70 min at 25°C.

The sequences in each Dm segment that are responsible for this specific hybridization to heat shock-induced RNAs were mapped at low resolution by cutting the cDm plasmids with a set of restriction endonucleases, ordering the resulting fragments and hybridizing each fragment to heat shock poly(A)$^+$ RNA or to cDNA copies of this RNA. The maps in Figure 1 show the positions of the Eco RI, Hind III, Sal I and Sma I cleavage sites on each cDm plasmid, and were constructed as indicated in the legend to this figure. The fragments used for the hybridization tests are those generated by double digestion with Sma I and Hind III. Since the Col E1 vector contributes the only Sma I site in the hybrids and contains no Hind III sites, the two Sma I-Hind III fragments in each digest contain both Col E1 and Dm DNA, and are designated S and L in Figure 1, according to whether they contain the short or long segment of Col E1 DNA. By contrast, the Hind III fragments contain only Dm DNA, and in Figure 1 the major fragments of this class are labeled according to their lengths in kb units (kb is a unit of length equal to 1000 bases or base pairs in single- or double-stranded nucleic acids, respectively).

These fragments were separated by agarose gel electrophoresis and transferred to nitrocellulose strips by the "blotting" technique of Southern (1975). The DNA on such strips was then hybridized with three heat shock probes: the total ^{32}P cytoplasmic poly(A)$^+$ RNA described above, ^{32}P polysomal poly(A)$^+$ RNA prepared from the polysomes of heat-shocked cells labeled as for the first probe, and ^{32}P-cDNA synthesized in vitro from the polysomal poly(A)$^+$ RNA of heat-shocked cells with the aid of reverse transcriptase (Experimental Procedures). Each probe produced essentially the same result, and Figure 2 shows that obtained with the ^{32}P-cDNA probe.

The results are also indicated by the black bars on the maps in Figure 1. Considering the cDm703 plasmid first, we note that all three of its fragments hybridize with the heat shock probes. For reasons that will become apparent, we know that all sequences in its 1.5 kb Hind III fragment are homologous to those in the heat shock RNA. Comparison of the hybridization levels for the S, L and 1.5 kb fragments from densitometric tracings of the autoradiographs indicates that the amounts of DNA in S and L which are homologous to the probe are equivalent to the lengths of the Dm regions in these fragments. Hence the black bars encompass these regions as well as the 1.5 kb fragment.

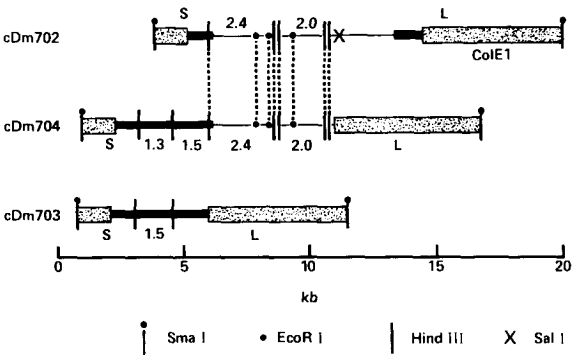


Figure 1. Restriction Maps of the cDm Plasmids

The normally circular plasmids are shown in the linear form produced by cleavage at the single Sma I site located in the Col E1 DNA (stippled bars) at a distance of 1.2 and 5.35 kb from its Eco RI site, where the Dm segments have been inserted by poly(dA)-poly(dT) connectors that destroy this site (Wensink et al., 1974). The other restriction sites were located with respect to the Sma I termini from the lengths of fragments generated by single, double and partial digests produced by the indicated enzymes. Fragment lengths were determined by agarose gel electrophoresis (Figure 2) using the length standards given in Experimental Procedures. The lengths of the S and L Sma I-Hind III fragments in kb units are, respectively: cDm702–2.15 and 9.0; cDm703–2.18 and 6.5; cDm704–2.22 and 5.5. The Hind III fragments labeled 1.3, 2.0 and 2.4 exhibit lengths of 1.27, 2.04 and 2.40 kb, while those labeled 1.5 in cDm703 and cDm704 are 1.48 and 1.50 kb long, respectively. Given that the two minor Hind III fragments in Dm702 and 704 are 0.05 and 0.2 kb long, the total lengths of cDm702, cDm703 and cDm704 obtained from fragment lengths are 15.8, 10.2 and 15.2 kb, respectively, while the comparable lengths determined by electron microscopy of the Sma I-cut DNAs are 16.2 ± 0.5 (N = 32), 10.8 ± 0.5 (N = 42) and 15.9 ± 0.5 kb (N = 42). The black bars within the Dm segments indicate regions homologous to the heat shock RNAs (see text). None of the plasmids contains Bam HI or Sst I sites.

Similarly, the Dm regions in the S fragments of Dm702 and Dm704 consist of RNA-homologous sequences. The level of hybridization to the other fragments from cDm702 indicates the presence of a small amount of RNA-homologous sequences in the 2.4 kb Hind III fragment (0.1-0.2 kb) and 1.1 kb in L. We have placed these sequences adjacent to those in S and at the right end of the Dm region in L, respectively, for reasons that will become apparent below. We have located the RNA-homologous sequences on the map of cDm704 in a similar manner, with one exception: the level of hybridization to the 1.3 kb Hind III fragment was 60% of that expected if all its sequences were homologous to the probes, and we have not attempted to indicate

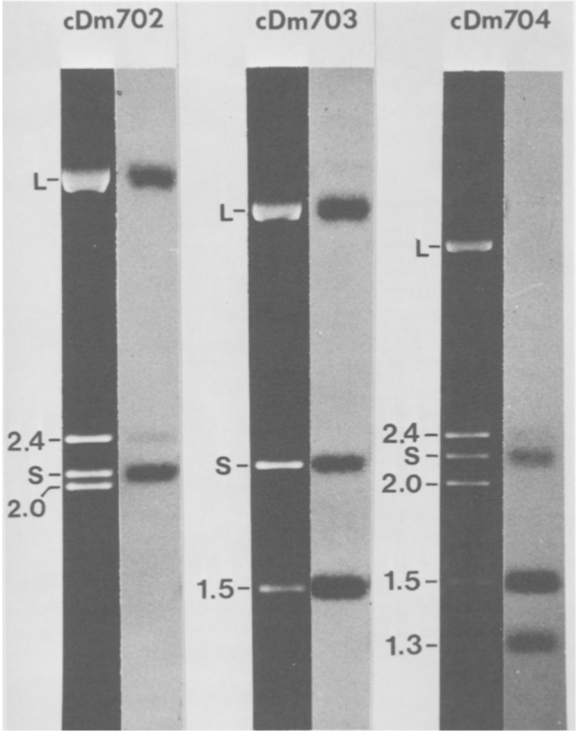


Figure 2. Restriction Fragments Containing Sequences Homologous to Heat Shock RNAs

Fragments present in Sma I, Hind III double digests of the cDm plasmids were separated by electrophoresis at 10 mA in a 1.0% agarose horizontal slab gel (21 × 14.5 × 0.5 cm) and are labeled as in Figure 1. The left-hand track of each pair shows the ethidium fluorescence pattern of the fragments in the gel when irradiated with short wave ultraviolet light with a peak intensity at 254 nm (Mineralight Transilluminator, Ultraviolet Products), which also introduces nicks into the DNA that facilitate its transfer to nitrocellulose strips after denaturation (Southern, 1975). The right track shows the autoradiographic response obtained after ^{32}P-cDNA (6×10^7 cpm/μg, 6×10^4 cpm/ml) complementary to heat shock polysomal poly(A)$^+$ RNA was hybridized to the fragments on the strips (Experimental Procedures). Exposure was 16 hr at −70°C, using a Lightning Plus intensifier screen (Dupont).

the nonhomologous sequences suggested by this observation.

Comparison of the restriction maps of cDm702 and cDm704 indicates an overlap between the respective Dm segments, and suggests that they were generated by overlapping breakpoints in the chromosomal DNA. Results consistent with this postulated overlap were obtained when ^{32}P-cRNA transcribed in vitro from cDm702 DNA was hybridized to the Hind III fragments of cDm704 by the same techniques used to generate the data in Figure 2. The 1.5, 2.0 and 2.4 kb fragments all exhibit strong hybridization with the cRNA probe (data not shown), as expected from such an overlap. The 1.3 kb Hind III fragment, however, also hybridized with this cRNA, and this is not directly predictable from the overlap. The Dm702 segment evidently contains sequences that are also present just to its left within the chromosomal region represented by the left end of Dm702 and Dm704. Homologies involving sequences in the S and L fragments cannot be examined by this cross-hybridization test, since these fragments possess common Col E1 sequences.

The 1.3 and 1.5 kb Hind III fragments of cDm704 also hybridize with cRNA from cDm703, as does the 2.4 kb fragment at a low level. The correspondence between this hybridization pattern and that obtained with the heat shock RNA sequences (Figure 2) is striking, and indicates that the homologous sequences in Dm703 and Dm704 are those sequences which hybridize with the heat shock-induced RNA. Furthermore, the restriction maps of cDm703 and cDm704 (Figure 1) demonstrate that, at most, only a small fraction of Dm703 can overlap Dm704 in the chromosome; hence their sequence homologies must result from sequence repetition. This repetition is also indicated by the observation that cRNA from cDm702 hybridizes to the 1.5 kb fragment of Dm703.

Heteroduplexes Provide High Resolution Maps of Sequence Homologies among the DNA Segments and Demonstrate That the RNA-Homologous Regions Consist of Tandemly Repeated Units

We have determined the homologies among the three Dm segments in more detail by electron microscopic examination of heteroduplex molecules consisting of one DNA strand from each of two hybrid plasmids that have been cut with Sma I. We consider first the homologies between Dm703 and Dm704. Three different heteroduplex configurations are formed from this pair, which are shown in Figures 3A–3C. The A and B configurations appear in approximately equal frequencies (54 and 44%, respectively; N = 135), while C represents a minority class (2%). Examination of A and B reveals that a 2 kb region in Dm704, located between the single-stranded regions in both configurations, is capable of pairing with two overlapping regions in Dm703. This 2 kb region is designated $\alpha\beta\alpha$ in the figure, where α represents the region of overlap. Thus the region defined by $\alpha\beta\alpha\beta\alpha$ in Dm703 pairs with the $\alpha\beta\alpha$ region in Dm704 in two ways: in A, Dm703 contributes the right-hand $\alpha\beta\alpha$ triplet, while in B, it contributes the left-hand $\alpha\beta\alpha$ triplet, each configuration using the central α of Dm703.

The substitution loop in the A configuration demonstrates that the 0.9 kb γ element, located immediately to the left of the $\alpha\beta\alpha$ region in Dm704, exhibits no detectable homology with the 1.1 kb β element. The γ element must be preceded on its left by an α element, since this region pairs with the left-hand α of the $\alpha\beta\alpha\beta\alpha$ sequences of Dm703

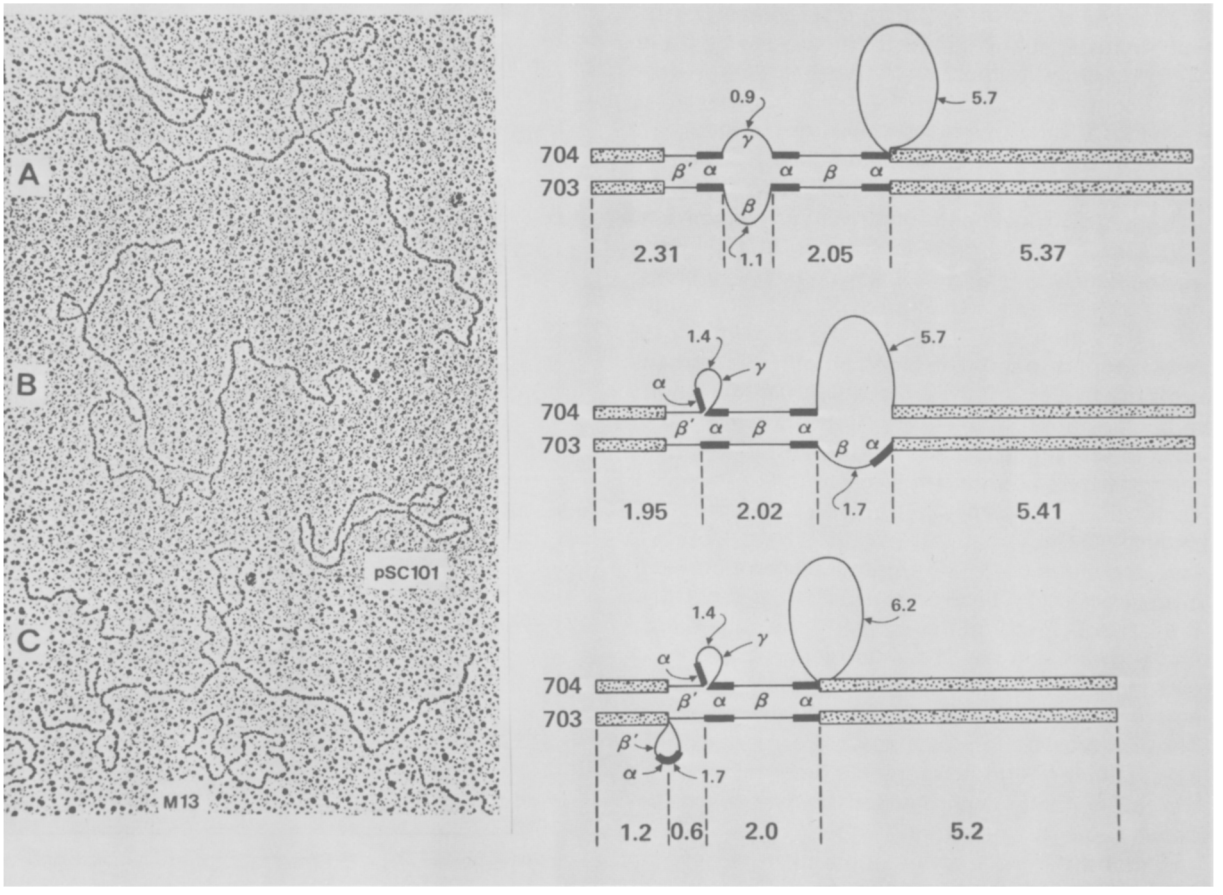


Figure 3. cDm703-cDm704 Heteroduplexes

Heteroduplexes between Sma I-cut cDm703 and cDm704 were formed and measured as described in Experimental Procedures. A montage of three electron micrographs shows examples of the A, B and C configurations (see text), as well as the length standards, pSC101 and M13 (Experimental Procedures). Immediately to the right of each example is the heteroduplex map for the configuration; the mean lengths in kb units were obtained from 33 ± 1 and 33 molecules for A and B, respectively, and from one molecule for C. Standard deviations for critical lengths are given in the text; the average fractional standard deviation for all duplex lengths shown here and in Figure 4 is 0.036 (±0.027), where the values in parentheses give the range, and that for single-stranded lengths is 0.069 (±0.029). The Col E1 segments and the α elements in each strand are represented by stippled and black bars, respectively; β and γ elements are indicated by the labeled thin lines. Straight horizontal pairs of strands indicate duplex regions; curved strands represent unpaired regions.

(configuration A). Finally, the existence of the rare C form requires that the left-terminal regions of both Dm703 and Dm704 consist of a part of the β segment, designated β' in Figure 3. (Additional evidence for this conclusion is provided by the cDm702-cDm703 heteroduplexes described below). The Dm703 segment is thus defined by the tandem array $\beta'\alpha\beta\alpha\beta\alpha$, while the Dm704 segment consists of a similar array that differs by the substitution of a γ for a β element ($\beta'\alpha\gamma\alpha\beta\alpha$) and is followed by a long region containing no sequences homologous to Dm703.

Comparison of these heteroduplex maps for cDm703 and cDm704 with their restriction maps (Figure 1) indicates that the α element contains a Hind III cleavage site near its right end. For example, the distances from the left-hand Sma I terminus of cDm704 to the first, second and third Hind III sites are 2.2, 3.5 and 5.0 kb on the restriction map, while the mean distance from this Sma I terminus to the right-hand ends of the first, second and third α elements on the heteroduplex map are 2.31 ± 0.06 (N = 33), 3.74 ± 0.15 (N = 67) and 5.33 ± 0.15 kb (N = 67), where the ± values are standard deviations. The distance to the right end of an α element is always slightly greater than the distance to the corresponding Hind III site, not only in cDm704, but also in cDm703 and, as described below, in cDm702. The weighted mean of six such differences obtained from all three plasmids is 0.14 kb, and the standard deviation in this mean (that is, the standard error) is ± 0.04 kb. We therefore estimate that the Hind III site in α is located 0.1–0.2 kb from its right end. This small part of α must therefore extend into the 2.4 kb Hind III site of cDm704. It accounts for the small amount of RNA-

homologous sequences in this fragment (0.1–0.2 kb) since all sequences in both the α and β elements are homologous to heat shock-induced RNA.

The localization of the Hind III sites in the tandem arrays of Dm703 and Dm704 demonstrates that the 1.5 kb Hind III fragment present in both segments represents an $\alpha\beta$ repeat unit. Similarly, the 1.3 kb fragment in Dm704 represents an $\alpha\gamma$ unit, which is also repeated in the genome. We have independently determined the lengths of these two units from the lengths of the α, β and γ elements in the heteroduplex maps. The mean lengths of the β and γ elements determined from the substitution loop in configuration A are 1.10 ± 0.06 and 0.87 ± 0.09 kb, respectively (N = 33). The weighted mean for the length of α obtained from both the cDm703-cDm704 and cDm702-cDm703 heteroduplexes (Figures 3 and 4) is 0.49 ± 0.10 kb (N = 214; N is large because there are three α elements in each cDm703-cDm704 heteroduplex and the lengths of most can be determined by using the values for β and γ given above). The mean lengths of the $\alpha\beta$ and $\alpha\gamma$ repeat units obtained by electron microscopy are therefore 1.36 ± 0.13 and 1.59 ± 0.12 kb. These values are in reasonable agreement with the lengths of the Hind III fragments determined by gel electrophoresis, given that for the most part, they derive from single-strand length measurements, and that different length standards were used for the two determinations (Experimental Procedures).

Heteroduplexes formed from Sma I-cut cDm702 and cDm703 plasmids are also found in three different configurations. These are shown in Figures 4A–4C and account, respectively, for 48, 19 and 33% of the population (N = 27). It is evident from these heteroduplexes that Dm702 contains a $\beta'\alpha$ sequence at its left end and a complete β element at its right end, with a "spacer" of 7.36 ± 0.44 kb (N = 27) separating them.

Comparison of the B with the A and C configurations confirms the conclusion that the β' region is part of the β element. The lengths of β' obtained from the cDm703-cDm704 heteroduplexes (0.62 ± 0.13 kb; N = 67) and the cDm702-cDm703 heteroduplexes (0.54 + 0.13 kb; N = 27) indicate little if any significant difference among the β' regions in all three Dm segments, particularly since differences in the poly(dA)-poly(dT) connector lengths will be reflected in the β' lengths. (The connector lengths in these plasmids have not been determined, but they should center at ~0.08 kb, the mean length of the polynucleotide tails used in their construction.) The fact that three of the six shear-induced breakpoints required to generate the three Dm segments are located in approximately the same position in different β elements

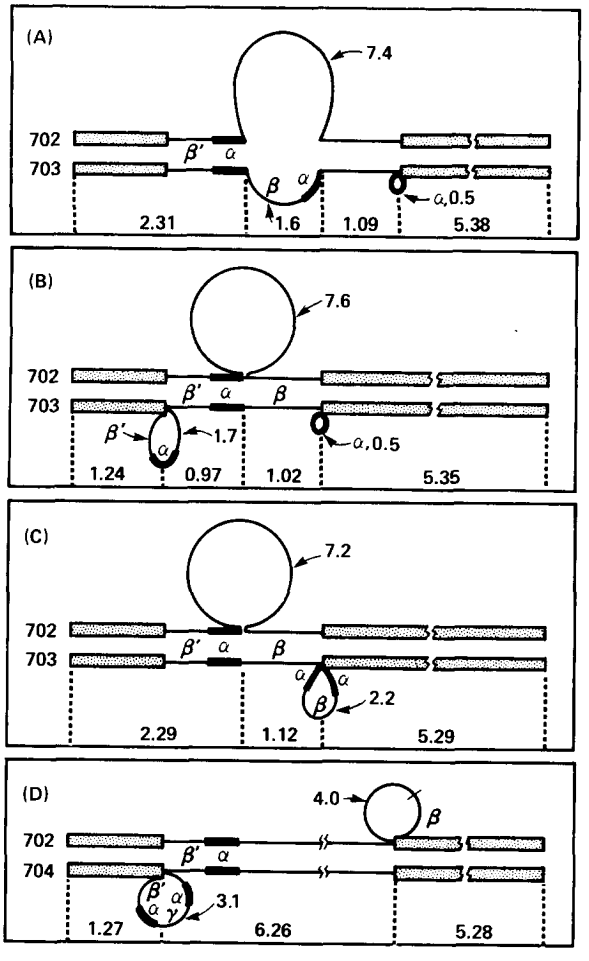


Figure 4. cDm702-cDm703 and cDm702-cDm704 Heteroduplexes

The comments and symbols given for the heteroduplex maps in Figure 3 apply here. The number of heteroduplexes measured to obtain the mean lengths given in the A, B, C and D maps was 13, 5, 9 and 11, respectively. See the text and the legend to Figure 3 for standard deviations of these lengths.

suggests a sequence-specific weak point to shear breakage. One can imagine several causes for such a weak point, including an increased probability of shear-sensitive single-stranded states arising from a pause in transcription (Gilbert, 1976) or in replication, or from a high density of dA:dT base pairs.

The β element at the right end of Dm702 accounts for the 1.1 kb of RNA-homologous sequences previously observed in the L fragment of cDm702 (Figure 1). Similarly, the length of the $\beta'\alpha$ region (1.03 ± 0.10 kb; N = 27) at the other end of Dm702 is in reasonable agreement with the amount of RNA-homologous sequences in the S and 2.4 kb fragments (1.0 and 0.1–0.2 kb, respectively). Indeed, virtually all of the RNA-homologous regions in each Dm segment can be accounted for by sequences in their α and β elements. The only apparent exception is the 0.8 kb of RNA-homolo-

gous sequences observed for the 1.3 kb Hind III fragment of Dm704. This is somewhat greater than can be accounted for by the α element in this $\alpha\gamma$ unit and may indicate the presence of some γ sequences in the heat shock RNA.

The last heteroduplex which we examined demonstrates that Dm704 contains part of the spacer region in Dm702, and confirms the overlap between these segments that was apparent from their restriction maps. The predominant configuration of the cDm702-cDm704 heteroduplexes is illustrated in Figure 4D. We did not make an extensive search for the other configurations of these heteroduplexes, because they are expected to be rare and because the structural information which they could provide had already been obtained from the cDm702-cDm703 and cDm703-cDm704 heteroduplexes.

Heat Shock RNAs Homologous to $\alpha\beta$ Come in Different Sizes

The length distribution of the heat shock cytoplasmic and polysomal poly(A)$^+$ RNAs that hybridize to the three Dm segments was determined by electrophoresis of these RNAs in two kinds of denaturing gels followed by hybridization of the separated RNAs to either cDm703 or cDm704 DNA fixed to filters. Figure 5 shows the results obtained when ^{3}H polysomal poly(A)$^+$ RNA labeled in heat-shocked cells was fractionated in a formamide-acrylamide gel, and slices from the gel were assayed for total labeled RNA and for the amount of labeled RNA that hybridized to cDm704 DNA. The length distribution of the total labeled RNA is similar to that observed by Spradling et al. (1977) for cytoplasmic poly(A)$^+$ RNA transcribed in heat-shocked cells. The hybridizable RNA falls into length classes I and II, which exhibit modes at 2.8 and 1.9 kb, respectively.

Essentially the same result was obtained when this experiment was repeated with ^{32}P cytoplasmic poly(A)$^+$ RNA labeled in heat-shocked cells, except that the slight hybridization exhibited by the RNAs in slices 23–25 of Figure 5 was accentuated, indicating the existence of a distinct third length class (III) with a mode at 0.8 kb.

Length classes I, II and III were also observed when the heat shock polysomal poly(A)$^+$ RNA was fractionated on methyl mercuric hydroxide-agarose gels, and the RNAs in the gel were then assayed by hybridization with cDm703 DNA according to the technique of Alwine, Kemp and Stark (1977). In this case, the RNA was directly transferred and covalently coupled to diazobenzyloxymethyl paper where, after appropriate treatment to destroy remaining diazo groups, it was hybridized with a ^{32}P-cDm703 DNA probe labeled by nick translation (Rigby et al., 1977) and the hybridization was monitored by autoradiography (see Experimental Procedures for conditions). The three zones of hybridization evident on the autoradiographs (not shown) exhibit midpoints corresponding to RNA lengths of 2.8, 1.8 and 0.8 kb, in good agreement with the previous results. Evidently all three length classes contain sequences present in the $\alpha\beta$ units which comprise Dm703. This was also demonstrated for classes I and II by hybridizing the RNA in each class (obtained from the peak fractions shown in Figure 5) to each of the Hind III fragments of Dm704, and the relative level of hybridization exhibited by the 1.3, 1.5 and 2.4 kb fragments was approximately the same as that previously observed with unfractionated polysomal poly(A)$^+$ RNA from heat-shocked cells.

It should be emphasized that the relative amounts of RNA in the three classes varied for unknown reasons (see Discussion). Whereas class II always represented a major or the major fraction

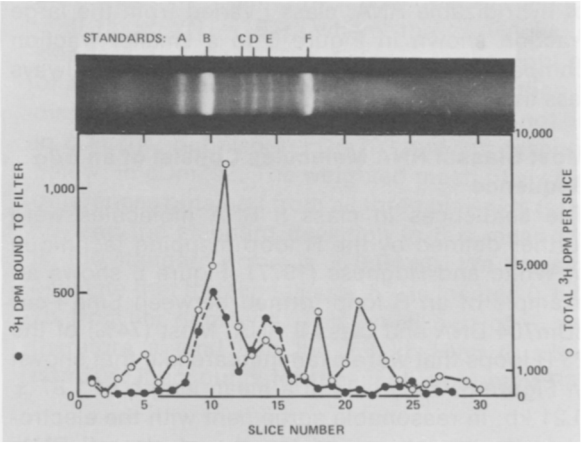


Figure 5. Electrophoretic Analysis of Polysomal Heat Shock RNAs that Hybridize with cDm704 DNA

^{3}H-poly(A)$^+$ RNA labeled in heat-shocked cells and isolated from the polysomal fraction obtained by zone sedimentation in a sucrose gradient (Experimental Procedures) was fractionated by electrophoresis in a 96% formamide, 3.5% acrylamide gel (Duesberg and Vogt, 1973). Slices from the gel were sonicated (Young and Hogness, 1977), and portions of each slice were assayed for total ^{3}H-RNA (open circles) and for that which hybridizes with 2.5 μg of cDm704 DNA bound to nitrocellulose filters (closed circles; Experimental Procedures). A background determined for each slice with blank filters and averaging 25 dpm was subtracted. The ethidium fluorescence pattern of the length standards is shown above the graph. (A) Qβ RNA (3.6 kb, taken as equivalent to MS2 RNA; Fiers et al., 1976); (B) E. coli 23S rRNA (3.05 kb; Van Holde and Hill, 1974); (C) (2.3 kb) and (E) (1.9 kb) are the processed fragments of D. melanogaster 28S RNA (Jordan, Jourdan and Jacq, 1976); and (D) (2.2 kb) is D. melanogaster 18S RNA (lengths from D. Kemp, personal communication); (F) E. coli 16S RNA (1.52 kb; Fellner, 1974). The lengths of class I (slices 9–11) and class II (slices 13–15) RNAs were interpolated from the straight line obtained when the logarithm of the lengths of the standards was plotted against their mobilities (see text for class III RNAs).

of hybridizable RNA, class I varied from the large fraction shown in Figure 5 to a smaller fraction comparable to that for class III, which was always less than that for class II.

Most Class II RNA Molecules Consist of an $\alpha\beta\alpha$ Sequence

The sequences in class II RNA molecules were further defined by the R loop mapping technique of White and Hogness (1977). Figure 6 shows an example of an R loop formed between Sma I-cut cDm704 DNA and class II RNA. Most (74%) of the 23 R loops that were examined are like that shown in Figure 6. They exhibit a mean length of 1.87 ± 0.21 kb, in reasonable agreement with the electrophoretically determined lengths of class II RNA, and are separated by 3.29 ± 0.24 kb of duplex DNA from the nearest Sma I terminus of cDm704. This terminus is the left terminus on the restriction and

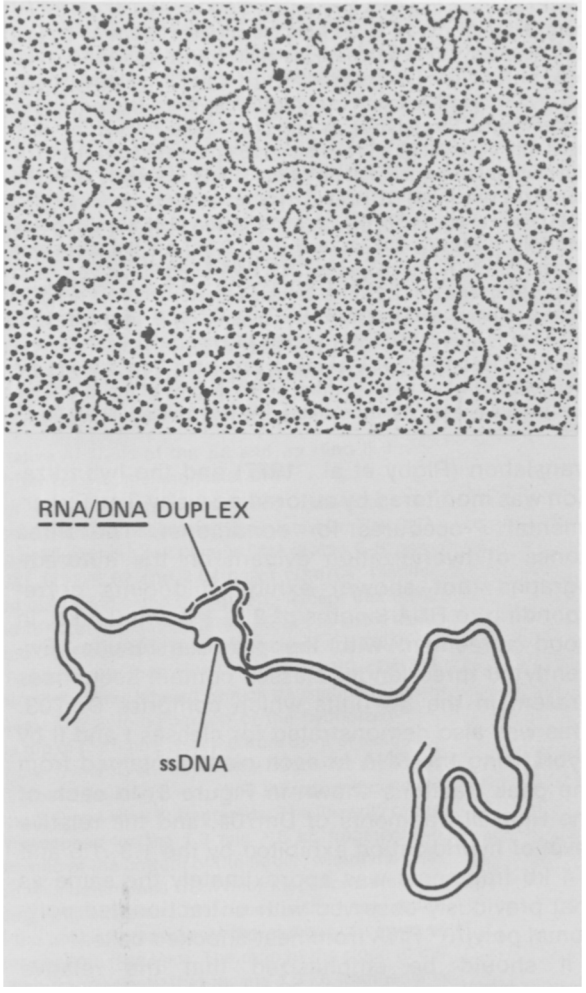


Figure 6. Electron Micrograph and Explanatory Diagram of an R Loop Formed between a Class II RNA Molecule and the $\alpha\beta\alpha$ Sequence in Sma I-Cut cDm704 DNA

heteroduplex maps of cDm704 shown in Figures 1 and 3, since the alternative assignment would place the R loop within the Col E1 DNA.

The length of the R loop may underestimate the total length of class II RNA molecules, since it should not include the poly(A) at their 3′ ends. Indeed, we have observed a very short single-stranded tail (0.1–0.2 kb) extending from the right end of a majority of the R loops, which we presume to be the poly(A) of the RNA because the 5′ to 3′ orientation of the RNA is from left to right (see next section).

Comparison of the position of these R loops with the heteroduplex map of cDm704 (Figure 3) indicates that the R loops encompass the $\alpha\beta\alpha$ region. The distances from the Sma I terminus to the left and right ends of this region are 3.30 ± 0.18 and 5.33 + 0.15 kb (N = 67)—distances which are not significantly different from those to the left and right ends of the R loops (3.29 ± 0.24 and 5.16 ± 0.32 kb; N = 17). We therefore conclude that most class II RNA molecules consist of an $\alpha\beta\alpha$ sequence and hence contain a terminal redundancy consisting of the α element.

Such an RNA should also form R loops with the terminal $\beta'\alpha$ region, which is separated from the above $\alpha\beta\alpha$ region by the γ element. It is clear that the remaining 26% of the R loops which we examined occupy this region. Thus the mean distances from the ends of these R loops to the closest Sma I terminus are 1.33 ± 0.20 and 2.32 ± 0.15 kb (N = 6)—values which are to be compared with 1.26 ± 0.07 (N = 17) and 2.31 ± 0.06 kb (N = 33), which are the distances to the left and right ends of the $\beta'\alpha$ region. We do not know, however, whether these R loops are made by $\alpha\beta\alpha$ RNA molecules, since we did not observe the single-stranded tails which would be expected to extend from the left ends of the R loops and which represent the unpaired 5′ halves of these RNAs. It may be that such a tail is too highly folded to be apparent under our spreading conditions (Experimental Procedures), or that these R loops result from shear breakage of more complex structures in which the tail has formed an R loop in the $\alpha\beta\alpha$ region of another cDm704 molecule. [See White and Hogness (1977) for examples of such complex configurations; we cannot comment on their existence in our preparation, since our analysis was restricted to simple forms.] Alternatively, it may be that the R loops in this class are formed by shorter $\beta'\alpha$ RNAs which contaminate the class II preparation—for example, class III molecules.

Transcription Proceeds from Left to Right

Figure 7 outlines the procedure by which we determined the 5′ to 3′ orientation of the $\alpha\beta$-type RNAs

relative to the cDm703 map. This procedure depends upon the way in which the poly(dA)-poly(dT) connectors in cDm703 were constructed; specifically, it depends upon the fact that the poly(dA) elements of these connectors were added to the 3′ ends of the Dm segment, while the poly(dT) elements were added to the 3′ ends of the Eco R1-cut Col E1. This is indicated in Figure 7 by the extensions of the thick and thin lines into the two connectors in the Sma I-cut cDm703 molecule represented at the top of the diagram. The S and L fragments, obtained by cleavage of this DNA with Hind III and fractionation in agarose gels (Figure 2), each contain a connector, and the strands of these fragments differ according to whether they contain a poly(dA) or poly(dT) element. Thus if we designate the strands by l and r according to their leftward and rightward 5′ to 3′ orientations, we see that the L-l and S-r strands contain poly(dT), whereas the L-r and S-l strands contain poly(dA).

We have used this difference to separate the strands of both fragments by chromatography on oligo(dA)- and oligo(dT)-cellulose (Figure 7, legend). Increasing amounts of the isolated strands were fixed to nitrocellulose filters and hybridized to a fixed amount of ^{32}P polysomal poly(A)$^+$ RNA from heat-shocked cells in the presence of a large excess of unlabeled poly(A) to prevent irrelevant hybridization of the RNA via its poly(A) tail. The ratio of labeled RNA which hybridized to saturating amounts of the L-l strands to that which hybridized to comparable amounts of L-r strands is 26. (234 ± 3 cpm of ^{32}P-RNA hybridized to saturating amounts of L-l DNA, and 9 ± 6 cpm hybridized to the L-r DNA, these values representing the means and ranges of two and four determinations, respectively, after subtraction of 51 cpm obtained for control filters containing no DNA.) When the same experiment was carried out with the strands of the S fragment, an S-l to S-r hybridization ratio of 4.3 was observed, the mean for two determinations at S-l DNA saturation being 137 ± 10 cpm above the same control value. (This somewhat lower ratio can be attributed to l strand contaminants in the S-r preparation, since hybridization to S-r as a function of the amount of DNA per filter did not exhibit a saturation plateau, but increased continuously with a slope, indicating a 10–15% contamination; thus the S-l to S-r hybridization ratio increases to 7 at amounts of DNA per filter which half-saturate the S-l hybridization.)

Since the heat shock RNA exhibits a strong preferential hybridization to the l strands of both fragments, we conclude that the 5′ to 3′ orientation of the RNA is from left to right on the cDm703 map. Furthermore, we can conclude that this orientation of transcription also applies to the α and β elements in cDm702 and cDm704, since the heteroduplex analysis demonstrates a common orientation of these elements in all three plasmids. In the case of cDm704, this places the γ element adjacent to the 5′ end of the $\alpha\beta\alpha$ RNA-homologous segment.

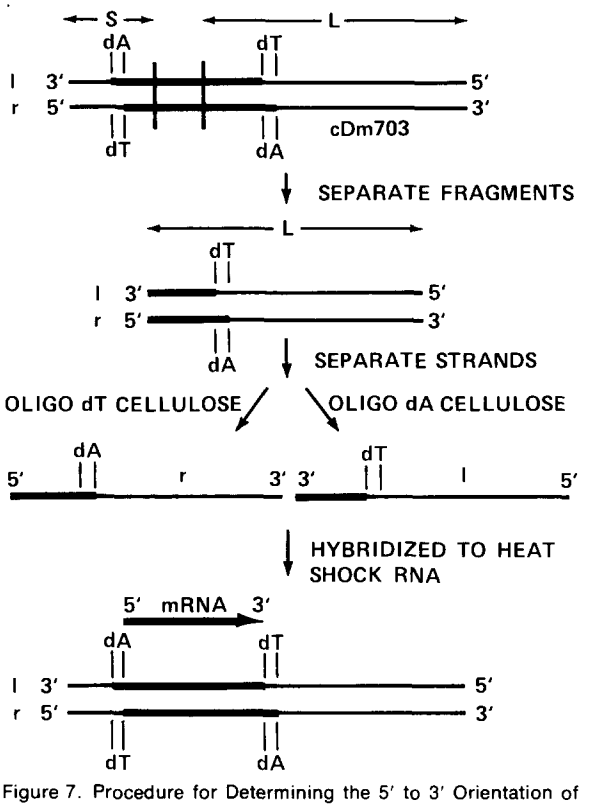
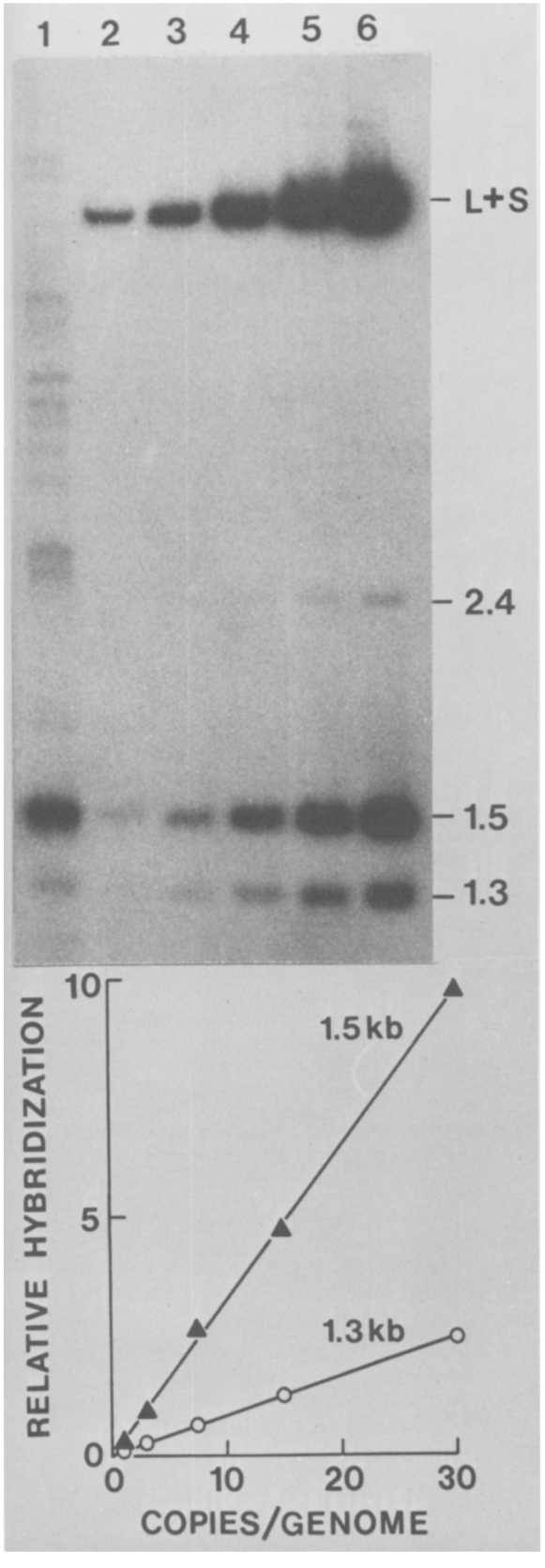


Figure 7. Procedure for Determining the 5′ to 3′ Orientation of the Heat Shock RNA Relative to cDm703

Sma I-cut ^3H-cDm703 DNA (top line; see text for symbols) was cleaved with Hind III, and the S and L fragments were isolated by gel electrophoresis as indicated in Figure 2. After denaturation with alkali and neutralization, the strands of each fragment (≪1 μg/ml) were immediately passed over oligo(dT)-cellulose columns as described for the fractionation of poly(A)$^+$ RNA (Young and Hogness, 1977). The bound strands were eluted from the column by the low salt wash and passed through the column a second time without prior denaturation, and the twice bound strands from the L and S fragments were labeled L-r and S-l, respectively (see text). The strands which failed to bind to the oligo(dT)-cellulose column on the first pass were applied to an oligo(dA)-cellulose column; those that bound and were subsequently eluted by the low salt wash are called L-l and S-r. The four ^3H-labeled strand preparations were fixed to nitrocellulose filters in increasing amounts (0.004–0.033 μg, determined from the ^3H radioactivity retained on the filter after hybridization), and each filter was hybridized with ^{32}P polysomal poly(A)$^+$ RNA labeled in heat-shocked cells (2.0×10^4 cpm per filter; see Experimental Procedures for conditions of this hybridization) with the results given in the text.

There Are at Least Twelve $\alpha\beta$ and Six $\alpha\gamma$ Units Arranged in Tandem Arrays, and These Account for Half the $\alpha\beta$ Homologous Sequences in the Haploid Genome

The three Dm segments which we have analyzed contain all or part of at least five, and possibly six, different $\alpha\beta$ units and one $\alpha\gamma$ unit arranged in



tandem arrays. What is the total number of tandemly repeated $\alpha\beta$ and $\alpha\gamma$ units in the haploid genome? A minimum value was obtained from the number of 1.5 and 1.3 kb Hind III fragments that derive from the tandem arrays within the genome. Total embryonic D. melanogaster DNA was digested to completion with Hind III, and the resulting fragments were fractionated by agarose gel electrophoresis. After transfer to nitrocellulose strips, the fragments were assayed by quantitative hybridization with saturating amounts of ^{32}P-labeled $\alpha\beta$ and $\alpha\gamma$ sequences obtained, respectively, from cDm703 and mkDm750, a plasmid consisting of the 1.3 kb Hind III fragments of cDm704 inserted into the pML21 vector (Experimental Procedures).

Figure 8 shows the autoradiographic results of such an experiment in which ^{32}P cDm703 DNA was used as the probe. Track 1 indicates the Hind III fragments in total D. melanogaster DNA that contain $\alpha\beta$ sequences. Clearly the 1.5 kb fragments exhibit the strongest autoradiographic response. Quantitative analysis of this response was achieved by comparing it to that obtained with Hind III-cut cDm704 DNA (tracks 2-6). Track 1 contains 0.70 μg of total D. melanogaster DNA, and one copy of the 1.5 kb fragment per haploid genome (165,000 kb; Rasch, Barr and Rasch, 1971; Rudkin 1972) represents 6.4 pg of this DNA. Track 2 contains 0.70 μg of Hind III-cut calf thymus DNA plus an amount of Hind III-cut cDm704 equivalent to 6.4 pg of the 1.5 kb fragment (that is, 68 pg of cDm704), and tracks 3-6 contain 3, 7.5, 15 and 30 times this amount of cDm704 DNA mixed with 0.7 μg of calf thymus DNA. Thus the autoradiographic response

Figure 8. Quantitative Analysis of the $\alpha\beta$ and $\alpha\gamma$ Hind III Fragments Obtained from Total D. melanogaster DNA

Electrophoresis of Hind III digests of embryonic D. melanogaster DNA (0.7 μg; track 1) and of increasing amounts cDm704 DNA mixed with 0.7 μg of Hind III-cut calf thymus DNA (tracks 2-6, containing 0.068, 0.20, 0.51, 1.02 and 2.04 ng of cDm704, respectively) was carried out as indicated in Figure 2, except that 0.5% agarose was used to form the slab. After exposure to short wave ultraviolet light and transfer to a nitrocellulose sheet, the DNAs on the sheet were hybridized with 5 ml of denatured ^{32}P-cDm703 DNA 7 × 10^6 cpm/ml) that was labeled by nick translation (Experimental Procedures). The autoradiograph of the hybridized ^{32}P-DNA shown at the top of the figure resulted from 39 hr of exposure at room temperature. Symbols identifying the fragments are those used in Figure 1, where the long S + L Hind III fragment consists of the linked L and S Sma I-Hind III fragments. The linearity of the autoradiography response for the 1.3 and 1.5 kb fragments in tracks 2-6 is indicated at the bottom of the figure, where the relative hybridization was determined from microdensitometric tracings of autoradiographs exhibiting responses in the linear range of the film, and the abscissa represents the number of copies of the fragments per genome (see text). The ratio of the slope of the curve for the 1.3 kb fragments to that for the 1.5 kb fragments is 0.26, a value in reasonable agreement with the ratio of α sequences in the $\alpha\gamma$ unit (0.49 kb) to $\alpha + \beta$ sequences in the $\alpha\beta$ unit (1.59 kb), as determined from the heteroduplex maps.

in tracks 2–6 is equal to that expected for total D. melanogaster DNA, were it to contain 1, 3, 7.5, 15 or 30 copies of this $\alpha\beta$ fragment per genome. The graph below the autoradiograph shows that the autoradiographic response is a linear function of this copy number, allowing one to determine easily the number of 1.5 kb $\alpha\beta$ fragments per genome from the autoradiographic response in track 1. A mean of 12.3 ± 2.6 copies per haploid genome was obtained from four experiments in which the amount of D. melanogaster DNA varied from 0.7 to 5.0 μg and ^{32}P-cRNA was also used as the probe.

Figure 8 also shows that the 1.3 kb $\alpha\gamma$ fragment can be assayed in a similar manner, even with the $\alpha\beta$ sequences in the cDm703 probe. A stronger response, however, is obtained with the $\alpha\gamma$ probe provided by mkDm750. Using both probes, we have obtained an average of 5.7 and 1.6 copies of the $\alpha\gamma$ fragment per haploid genome.

A 1.5 or 1.3 kb Hind III fragment will be obtained whenever a β or γ element is flanked at both ends by an α element. Hence an individual $\alpha\beta$ or $\alpha\gamma$ unit that is not part of a tandem array should not generate such a fragment, nor should one of the two terminal units in an array. For example, a $\beta\alpha\gamma\alpha\beta\alpha$ array will produce a 1.3 and a 1.5 kb fragment consisting, for the most part, of sequences from the central and right-hand units, respectively, whereas virtually all sequences in the left-hand unit will be in a Hind III fragment whose length is determined by the position of the first Hind III site to the left of the array. Hence the number of units in the tandem arrays should equal the sum of the numbers of 1.5 and 1.3 kb fragments plus the number of tandem arrays per genome, or 18 + T, where T equals the number of tandem arrays. In the Discussion, we show that T ≥ 3 and, therefore, that the number of units in such arrays is ≥ 21.

Track 1 of Figure 8 reveals that other Hind III fragments contain $\alpha\beta$ sequences. More than 20 of these extra hybridization bands are reproducibly observed in autoradiographs obtained when greater amounts of D. melanogaster DNA or longer exposures are used than those indicated for this figure. Presumably they arise from sequences contained in terminal units of the arrays and from sequences located outside of these arrays, an assumption that has recently been confirmed by analysis of the DNA from mutants containing deletions which include all of the tandem arrays (J. T. Lis, D. Ish-Horowicz and D. S. Hogness, manuscript in preparation; see Discussion).

Approximately 40% of the autoradiographic response is due to the 1.3 and 1.5 kb fragments, with the remaining 60% resulting from the extra bands. From this distribution, we estimate that all of the α and β sequences, whether in tandem arrays or elsewhere, occupy approximately 50 kb of the DNA in the haploid genome. A better determination of the amount of DNA occupied by all α and β sequences was obtained from their reassociation kinetics. Figure 9 shows the reassociation of a minute amount of short (0.4 kb) single-stranded ^{32}P-cDM703 DNA in the presence of a large amount of equally short, single-stranded ^{3}H-DNA from D. melanogaster embryos.

The curve labeled Dm703 is the computer-generated best fit of the function, $C^{Dm}/Co^{DM} = [1 + (K/n)Cot]^{-n}$, to the data points, where C^{Dm} and Co^{Dm} are the concentrations of renaturable sequences in the labeled Dm703 segment at t and t = 0, and Co is the initial concentration of total D. melanogaster

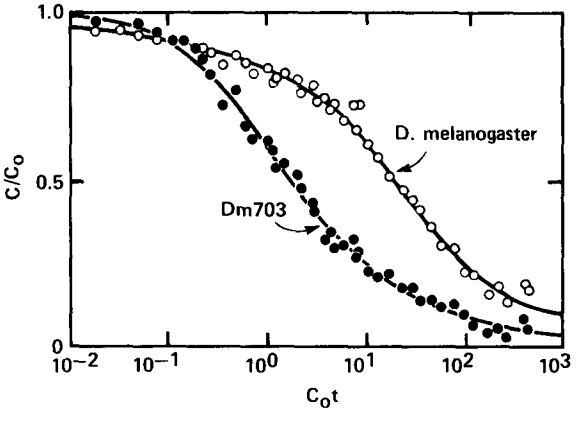


Figure 9. Reassociation of Dm703 DNA Driven by Total D. melanogaster DNA

The Cot of the abscissa is given in M × sec, where Co is the initial concentration of ^{3}H-labeled embryonic D. melanogaster DNA, which was 2.7 × 10^{-4} M and 1.4 × 10^{-2} M in the two reaction mixtures used here. Both contained the ^{32}P-cDm703 DNA probe at 4.5 × 10^{-9} M. Each mixture was sheared and denatured, and the reassociation of the labeled DNAs was carried out and assayed as indicated by Wensink et al. (1974), except that assays with S1 nuclease were performed in the presence of 500 μg/ml of calf thymus duplex DNA. The ordinate for the reassociation of D. melanogaster DNA (open circles) represents the ratio of the single-strand concentration of that DNA at time t to that at t = 0, whereas it represents the corresponding ratio of labeled Dm703 sequences in the ^{32}P-cDm703 probe for the Dm703 reassociation curve (closed circles)—that is, the C^{Dm}/Co^{Dm} ratio given in the text. Values of C^{Dm}/Co^{Dm} were obtained by subtracting out the contribution of the Col E1 sequences in the probe using a computer program devised by Wensink et al. (1974) which was modified to take account of the small portion of label in some probes that never enters the renatured fraction—that is, never becomes S1-resistant. Among the ^{32}P-cDm703 probes which we tested, the fraction of label that does not become S1-resistant, even in the rapid reassociation reactions driven by the homologous cDm703 DNA, varied within the range 15 ± 5%. The probes are labeled by nick translation (Rigby et al., 1977), and we believe that the residual S1-sensitive label derives from contaminants which are variably introduced in this reaction. For the Dm703 curve given above, the computed correction for this putative contamination results in a normalization to 100% of an 86% observed "maximum" reassociation of the Dm703 sequences.

DNA (Wensink et al., 1974). It represents the reassociation of the α and β sequences in Dm703 which is driven by the vast excess of these sequences in the D. melanogaster DNA. This best fit is obtained for n = 0.44 and K = 0.94 M^{-1} sec^{-1}, where n is an arbitrary constant introduced by Morrow (1974) to account for the effects of the S1 nuclease assay, and K is a kinetic coefficient proportional to the repetition frequency per haploid genome of the reassociating sequences (Wensink et al., 1974).

The ratio of K to the equivalent kinetic coefficient for the nonrepetitive sequences in D. melanogaster DNA is 27, and would represent the repetition frequencies of both the α and β sequences if they reassociate as the single kinetic component, as assumed in the above treatment of the data. The ratio, however, of the number of α to β elements in the twelve $\alpha\beta$ and six $\alpha\gamma$ Hind III fragments derived from the tandem arrays is 1.5, and curves generated by α and β components in this ratio (or its reciprocal) are not significantly different from the single-component curve in their fit to the data. It is therefore more appropriate to regard the above ratio as the weighted mean, $\bar{N}$, of the repetition frequencies for the α and β sequences—that is, $\bar{N} = f_\alpha N_\alpha + f_\beta N_\beta$, where N_α and N_β are the repetition frequencies of the α and β sequences, and f_α = 0.31 and f_β = 0.69, which are the fractions of an $\alpha\beta$ unit occupied by the α and β elements. Indeed, we have shown that $\bar{N}$ does not differ from 27 by more than 3% for values of N_α between 20 and 40 copies per genome, the corresponding values for N_β being 31 and 22. For the condition in which the ratio of the repetition frequencies of the α to β sequences is 1.5, N_α = 35.3 and N_β = 23.5.

The amount of DNA occupied by all α and β sequences in the genome is $\overline{NL}_{\alpha\beta}$, where $L_{\alpha\beta}$ is the length of an $\alpha\beta$ unit, or 1.59 kb. Since $\bar{N}$ is virtually invariant in the range of N_α given above, this $\alpha\beta$ DNA equals (27)(1.59) = 43 kb. The α and β elements contained in the $\alpha\beta$ and $\alpha\gamma$ Hind III fragments generated from the tandem arrays account for 22 kb or 51% of this total amount. Since these Hind III fragments do not include one terminal element from each array, we conclude that these arrays account for significantly more than half of the $\alpha\beta$ DNA. This fraction must, however, be appreciably <100%, because most of the extra Hind III fragments that contain $\alpha\beta$ sequences (Figure 8, track 1) are not eliminated by deletions that include all arrays (J. T. Lis, D. Ish-Horowicz and D. S. Hogness, manuscript in preparation.)

One can imagine that all arrays are identical, and that a higher order tandemly repeated unit exists which consists of the spacer in Dm702 plus one of these arrays. We have tested this possibility by an experiment similar to that shown in Figure 8, except that the total D. melanogaster DNA was digested with Eco RI instead of Hind III. Since the Dm702 spacer contains Eco RI sites and the $\alpha\beta$ and $\alpha\gamma$ units do not, such a higher order repetition should generate Eco RI fragments of a particular length class which contains all of the ≥21 tandemly repeated $\alpha\beta$ and $\alpha\gamma$ units. This is not the case. A variety of Eco RI fragments containing $\alpha\beta$ sequences were detected, similar in number to the product of Hind III digestion, but no single length class contained more than the equivalent of five copies of $\alpha\beta$ units per genome.

Chromosomal Mapping

Autoradiography of polytene chromosomes that have been hybridized in situ with ^{3}H-cRNA to cDm703 reveals a strong hybridization at the site of a major heat shock puff in 87C on the right arm of chromosome 3. Figures 10A and 10B show this autoradiographic response after short and long exposures, respectively. The short exposure shows that the grains are largely confined to bands 87C1-3. The long exposure demonstrates that there is no hybridization within 87A, a region that contains the site of another heat shock puff; even when there are more than 500 grains at 87C, no hybridization is detected in 87A.

In addition to the strong hybridization at 87C, a weaker hybridization with the α and β sequences in Dm703 is observed at the chromocenter, to which Spradling et al. (1975, 1977) had previously observed hybridization with heat shock RNAs. The number of grains over the entire chromocenter is approximately 30% of that at 87C (this includes grains that were sometimes observed over the neighboring 42B region). After very long exposures, a minute level of hybridization was observed at 10B.

Figure 10C shows the autoradiographic response in the 87 region after hybridization with the ^{3}H-labeled α and γ sequences in Dm750, the cloned $\alpha\gamma$ unit derived from Dm704. The major site of hybridization is again in 87C. In contrast to the results obtained with the $\alpha\beta$ sequences, the $\alpha\gamma$ probe also exhibits a significant reproducible hybridization in 87A, which must be due to the γ sequences. The ratio of grains in 87C to those in 87A is 25:1, which reflects the ratio of the sum of the α and γ sequences in 87C to the γ sequences in 87A, rather than the ratio of the γ sequences in the two regions (see Discussion). The Dm750 probe also hybridizes to the chromocenter, and more weakly to several other sites in the chromosome arms, none of which corresponds to any of the known heat shock sites.

cDm702 contains the 7.4 kb spacer element in

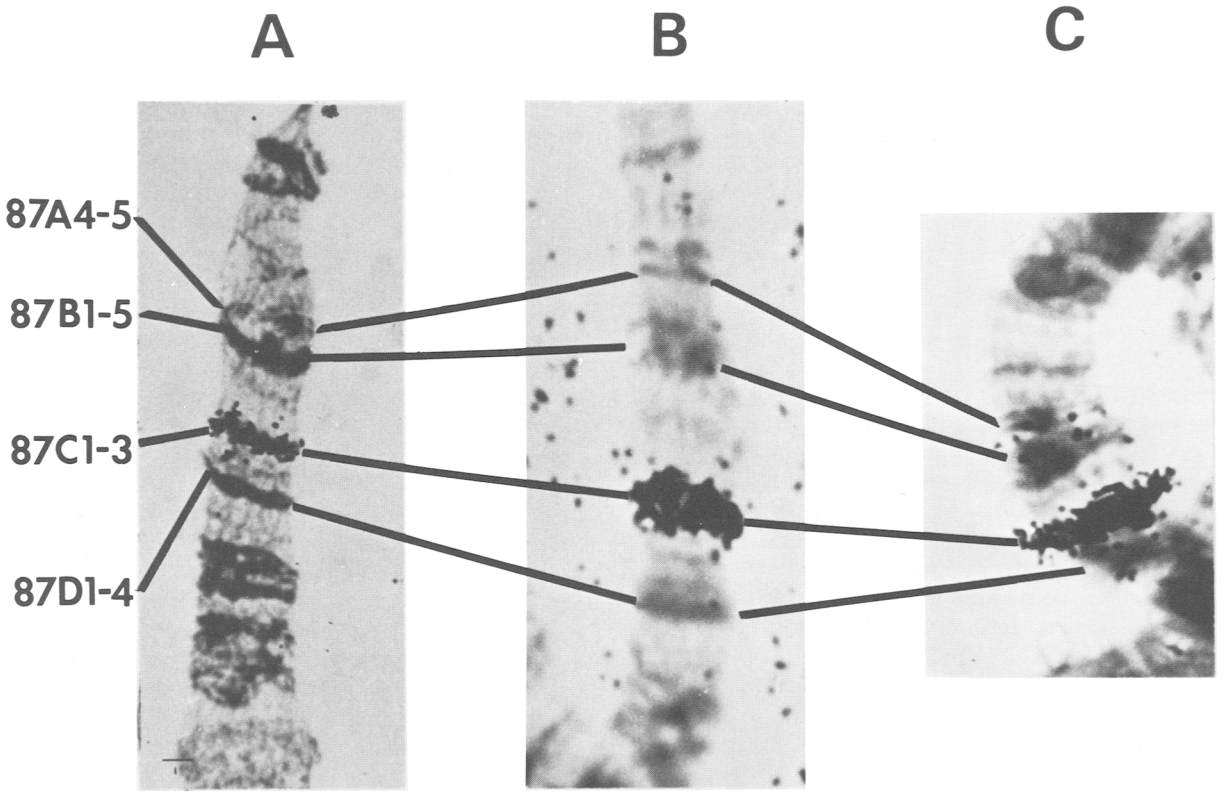


Figure 10. In Situ Hybridization of $\alpha\beta$ and $\alpha\gamma$ Units to Polytene Chromosomes

In situ hybridization was carried out as described by Wensink et al. (1974), except that the 70°C incubation introduced by Bonner and Pardue (1976) was added, and in some cases, ^{3}H-cDm DNA prepared by nick translation (Rigby et al., 1977) was used instead of ^{3}H-cRNA (Glover et al., 1975). The scale is indicated by the fact that the width of each of the chromosome arms shown is approximately 7 μm.

(A) Labeling of 87C1-3 by ^{3}H-cRNA to cDm703; 3.9 × 10^5 cpm per slide, 3.5 day exposure.

(B) More extensive labeling of the same region as in (A), with greater amounts of the same probe (11.3 × 10^5 cpm per slide) and longer exposure (67 days). No grains were observed over 87A, and the same result was obtained with a ^{3}H-cDm703 DNA probe prepared by nick translation.

(C) Labeling of both 87C and 87A by ^{3}H-mkDm750 DNA prepared by nick translation; 3.5 × 10^5 cpm per slide, 34 day exposure.

addition to α and β sequences. Hence any sites of hybridization generated by ^{3}H-cRNA to cDm702 which do not overlap those obtained with the cDm703 probe must result from sequences in the spacer. The hybridization pattern obtained with the Dm702 sequences is given in Table 1. The strongest hybridization response is again in 87C, the number of grains here being twice that at the next most heavily labeled locus. This pattern also overlaps that for Dm703 at the chromocenter. The remaining 31 hybridization loci must then contain sequences in the spacer—including 87A, because Dm702 contains no γ sequences, and 10B, because the autoradiographic response here is much stronger than the very weak response obtained with Dm703. Evidently the spacer contains dispersed repetitive sequences.

Dm704 contains all but the right-hand third of this spacer in addition to α, β and γ sequences. Table 1 shows that its hybridization pattern overlaps that for Dm702 not only at 87C (again the strongest hybridization site by a factor of two) and the chromocenter, but also at 27 of the 31 loci containing spacer sequences. The four loci that respond to Dm702 but not to Dm704 (27B, 65EF, 84DE, and 85DE) should therefore contain dispersed repetitive sequences located in the right third of the spacer. Similarly, the other 27 loci should contain sequences located in the remaining two thirds, 87A being included in this second group because its hybridization with Dm704 is greater than can be accounted for by the γ element. Dm704 also exhibits hybridization at three loci (18B, 79A and 79C) not present in the Dm702 hybridization pattern, and we presume that these result from the γ sequences in Dm704. With the exception of 87A and 87C, none of the bands that hybridize with the Dm702 and Dm704 probes corresponds to a known heat shock locus.

These in situ hybridization maps indicate that Dm702, Dm703 and Dm704 must derive either from 87C1-3 or from the chromocenter, with preference

given to 87C1-3. Conclusive proof that all three segments derive from 87C1-3 has recently been obtained by J. T. Lis, D. Ish-Horowicz and D. S. Hogness (manuscript in preparation), who showed that the 1.5 kb $\alpha\beta$ and 1.3 kb $\alpha\gamma$ Hind III fragments generated by tandem arrays, such as are present in the Dm702-Dm704 and Dm703 chromosomal segments, are missing from Hind III digests of total DNA from D. melanogaster possessing a homozygous deficiency for 87C.

Discussion

Tandem Topographies of the $\alpha\beta$ and $\alpha\gamma$ Units

The mapping data for the $\alpha\beta$ and $\alpha\gamma$ units in the three cloned Dm segments are summarized in Figure 11. The overlapping Dm702 and Dm704 define a chromosomal segment that extends from a tandem array of at least three units on the left througha spacer to a β element, which we presume marks the beginning of another array on the right. The Dm703 segment may be part of either of these two arrays. If Dm703 overlaps Dm702, the right-hand array would consist of at least three $\alpha\beta$ units in tandem, whereas if it overlaps Dm704, the left-hand array would contain at least five units in the order $\beta\alpha\beta\alpha\beta\alpha\gamma\alpha\beta\alpha$, each arrangement accounting for six units. Alternatively, Dm703 may not exhibit a chromosomal overlap with either segment, in which case the three segments would account for seven units, whether Dm703 derives from one of the above arrays or represents a third.

The six or seven units represented in these cloned segments account for no more than one third of all tandemly repeated units in the genome, since we have shown that the total number of these units is 18 + T, where T is the number of arrays per genome. All these arrays are contained in a single chromosomal segment located within the 87C1-3 bands. This is so because each array must generate at least one of the 18 Hind III fragments that represent single tandemly repeated units; deletion of the 87C region eliminates all 18 of these fragments (J. T. Lis, D. Ish-Horowicz and D. S. Hogness, manuscript in preparation); and in situ hybridization of $\alpha\beta$ sequences within this region is restricted to the 87C1-3 bands (Figure 10A).

How many arrays does this chromosomal segment contain? This question cannot yet be answered with much precision, but a lower limit of $T \geq 3$ can be determined from the present data. The finding that none of the Eco RI fragments generated from total D. melanogaster DNA contains more $\alpha\beta$ sequences than the equivalent of five $\alpha\beta$ units is not compatible with only two arrays. Given that each array is entirely contained within some Eco RI fragment because the units contain no Eco RI sites, then a division of 20 units between two arrays must yield an Eco RI fragment that does not satisfy the above constraint. A division of 21 units among three arrays, however, can satisfy this constraint, and hence $T \geq 3$. The smallest upper limit of general applicability that is determinable from the present data is $T \leq 16$, although one can show that $T \leq 6$ for the special case where all arrays terminate in $\alpha\beta$ units, a condition suggested by the arrangement of the units in the cloned segments.

Whatever the upper limit, our results indicate that 21 or more tandemly repeated $\alpha\beta$ and $\alpha\gamma$ units are arranged in multiple arrays separated by a heterogeneous group of spacers that together occupy a chromosomal segment of at least 40 kb which is located within the 87C1-3 bands. This topography is like that recently discovered for the

Table 1. Chromosomal Loci Hybridized by Dm702 and Dm704 Sequences

Chromosome Arm:				
X	2L	2R	3L	3R
(1EF)	21F-22B	47DE	62EF	(83BC)
2B8-14)	25D	49F	65EF	84DE
5EF	(27B)	51CD	68B	(84A)
10B		57B	(72F-73B)	(85DE)
16B		58C	75C	87A
(18B)		(59AB)	76F	87C
19-20		(60A)	(79A)	(87D)
			79C	(88B)
				98A
				98CD

In situ hybridization was carried out with ^{3}H-cRNA to cDm702 and cDm704, as indicated in Figure 10. The parentheses indicate minor sites that produce 0.5-2 grains per week of exposure; all other sites produce >2 grains per week. Sites labeled only by the Dm702 probe are indicated with a solid line, whereas those labeled only by the Dm704 probe are indicated with a dashed line. All other sites, including the chromocenter, are labeled by both.

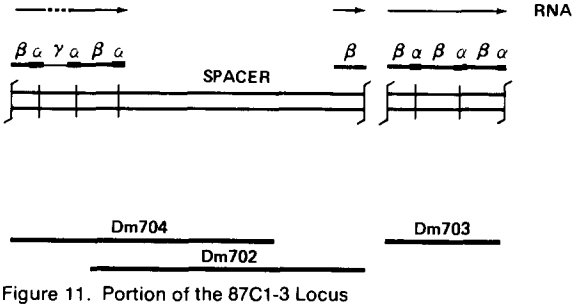


Figure 11. Portion of the 87C1-3 Locus

This map represents a composite of the restriction and heteroduplex maps for the three cloned segments (Figures 1, 3 and 4) which occupy the regions indicated below the map. Only the Hind III restriction sites within the $\alpha\beta$ and $\alpha\gamma$ units are shown (short vertical lines). The regions that are homologous to heat shock RNA and the 5′ to 3′ orientation of that RNA are indicated by the horizontal arrows, where the region corresponding to the γ element is represented by a dashed line to indicate that some γ sequence may also be transcribed.

tandemly repeated units which contain the D. melanogaster histone genes (Karp, 1978; Lifton et al., 1978; R. Lifton, personal communication). In this case, approximately 100 units, each of which contains the full complement of five histone genes, are similarly divided into multiple tandem arrays separated by heterogeneous spacers which together occupy a chromosomal segment of over 500 kb that is contained within several contiguous bands. The dimensions are larger than for the 87C1-3 locus, but the basic principles of chromosomal organization may be the same.

RNAs Transcribed from the Tandem Arrays

Most class II RNAs consist of an $\alpha\beta\alpha$ sequence. We presume that these RNAs are transcribed at the 87C1-3 locus, and not from the hetero-chromatin of the chromocenter, because any chromosomal segment containing an $\alpha\beta\alpha$ sequence is capable of generating an $\alpha\beta$ Hind III fragment, and this capability is restricted to the DNA within 87C1-3. Clearly it cannot be transcribed from any of the other heat shock sites, including 87A, since these contain neither α nor β sequences.

The class II RNA could be formed by either of two general mechanisms. It could represent a primary transcript, where, for example, transcription is initiated at or near the γ/α boundary at the right end of the γ element in Dm704, proceeds to the right and terminates at or near the α/spacer boundary. Alternatively, the $\alpha\beta\alpha$ RNA could arise from the processing of longer primary transcripts which include sequences from multiple units, perhaps including all units in an array, or even in more than one array. Such a transcript could represent a common precursor of all three cytoplasmic RNA length classes. The parallel results obtained when class I, II and III RNAs were hybridized to the Hind III fragments of cDm704 and to cDm703 indicate that class I and III molecules also consist, for the most part, of $\alpha\beta$ sequences, presumably in the order found in the tandem arrays. One can therefore imagine that these three RNAs are formed from a common precursor by either serial or parallel processing routes. Clearly the acquisition of maps for the class I and class III RNAs comparable to that obtained for class II molecules, and the determination of the length and sequence content of homologous nuclear RNAs, are the first steps to be taken in evaluating these various possibilities.

What is the function of these RNAs? The fact that they co-sediment with the polysomal fraction suggests that they are mRNAs. The sequence repetition present in the class II RNA, and probably in class I (although not in class III) molecules, does not seriously inhibit this suggestion. One or both of the terminally redundant α elements in class II molecules may not be translated; or they may be in different reading frames so that they yield different amino acid sequences; or translation of these RNAs may yield a polypeptide containing an amino acid sequence repetition that is subsequently removed by polypeptide cleavage. Furthermore, redundancies of approximately the same length have been observed in other repeated gene families of D. melanogaster (Finnegan et al., 1978), although in these cases, they do not result from tandem repetition.

This suggestion, however, is appreciably perturbed by the observation of Ish-Horowicz, Holden and Gehring (1977) that the deletion of 87C has no apparent effect on the synthesis of the heat shock proteins identified thus far. Since the deletions which they used are the same as those which eliminate the tandem arrays, the RNAs transcribed from these arrays evidently do not code for these proteins. While this result does not affect the possibility that these RNAs code for a less conspicuous heat shock protein, it does encourage speculation that they may contain regulatory rather than coding sequences. Because of their apparent association with polysomes, our speculation has centered on the possibility that these RNAs effect an alteration in translational specificity that appears to be part of the heat shock response (see Introduction).

The 87C site, however, does contain a structural gene for the major or 70,000 dalton heat shock protein. Deletion of both the 87A and 87C heat shock sites results in the loss of this protein (Ish-Horowicz et al., 1977), and the mRNA from which it can be translated in vitro (Mirault et al., 1978; Moran et al., 1978) exhibits in situ hybridization to 87A and 87C (Henikoff and Meselson, 1977; Spradling et al., 1977). These investigators have therefore postulated that copies of this structural gene are located at both sites, Henikoff and Meselson (1977) estimating 1–2 copies at 87A and 2–3 copies at 87C. Cloned Dm segments containing this structural gene are described in the accompanying paper (Schedl et al., 1978), and their characteristics confirm this postulate.

Several properties of the $\alpha\beta$ RNAs clearly distinguish them from this mRNA. First, they contain sequences not present at 87A; second, the sum of their abundancies is an order of magnitude less than that for the abundant mRNA, at least at 37°C; and finally, the electrophoretic mobilities of class I and II RNAs are, respectively, smaller and greater than that of the abundant mRNA, whose electrophoretically determined length is 2.6 kb (Henikoff and Meselson, 1977; Moran et al., 1978; in Figure 5, this RNA would migrate with the peak of total ^{3}H-labeled RNA). The 87C heat shock site must therefore contain members of two quite different re-

peated gene families, one of which is also represented at 87A.

Spradling et al. (1977) and Henikoff and Meselson (1977) have proposed an arrangement of this sort on the basis of the differential hybridization of two RNA fractions to 87A and 87C. In each case, one of these was the mRNA coding for the 70,000 dalton protein. The second RNA analyzed by Spradling and his co-workers was another polysomal RNA called A4, which exhibits preferential hybridization to 87C and, from its mobility in gels, is a probable candidate for our class II RNA. In addition, the concentration of A4 is highly dependent upon the shock temperature, reaching a maximum at 35°C and dropping off rapidly at higher temperatures. J. Lengyel (personal communication) has observed a similar temperature dependence for the cytoplasmic $\alpha\beta$ RNA, assayed by hybridization to cDm703. At 37°C, the amount of this RNA should be quite sensitive to minor temperature variations, and this may account for the variability which we observed in the amount and size distribution of the $\alpha\beta$ RNAs.

The second of the RNA fractions examined by Henikoff and Meselson (1977) consisted of the nuclear RNA which could not be competed away by the mRNA for the 70,000 dalton protein when hybridized to 87A and 87C. It seems probable that this RNA is the nuclear counterpart of the cytoplasmic $\alpha\beta$ RNAs defined here, since it hybridizes to 87C at levels indicating the presence of approximately 20 kb of homologous DNA, but exhibits little hybridization to 87A. Furthermore, this RNA appears to accumulate with increasing time at 37°C, and such an accumulation has also been observed for the nuclear $\alpha\beta$ RNAs (J. Lengyel, personal communication). This accumulation, and the fact that heat shock puffs regress after 30 min at 37°C, led Henikoff and Meselson to propose that the accumulation of this nuclear RNA has a regulatory role in reducing the rate of synthesis of mRNAs that derive from these puff sites. Perhaps the $\alpha\beta$ RNAs have multiple regulatory functions, operating at the translational level in the cytoplasm, and at the transcriptional or processing levels in the nucleus.

The γ Connection

At least six γ elements are present at 87C in the tandem arrays that comprise the repeated gene family from which the $\alpha\beta$ RNAs derive. Yet the γ sequences, like the genes that code for the 70,000 dalton protein, are found in 87A as well as 87C. Indeed, we estimate that the amount of DNA occupied by the γ sequences in 87A is equivalent to 1–2 γ elements, which equals the estimated number of structural genes for the 70,000 dalton protein at this locus (Henikoff and Meselson, 1977). [A lower limit of 0.6 kb for the γ sequences at 87A is obtained by dividing by 25 the amount of $\alpha\gamma$ DNA (that is, that occupied by α and γ sequences) in the six $\alpha\gamma$ and twelve $\alpha\beta$ units known to be present in 87C, the value for the divisor coming from the ratio of $\alpha\gamma$ DNA in 87C to that in 87A, as determined by in situ hybridization with the $\alpha\gamma$ probe (Figure 10C). An upper limit of 1.8 kb was obtained by assuming that all $\alpha\beta$ DNA (that is, that occupied by α and β sequences) at 87C which is not accounted for by the twelve $\alpha\beta$ units is occupied by α elements in $\alpha\gamma$ units, and dividing by 25 the total $\alpha\gamma$ DNA in 87C that results from such a distribution. The amount of $\alpha\beta$ DNA at 87C is $\leq (1 + r)^{-1}$ (43 kb) = 33 kb, where 43 kb is the total amount of $\alpha\beta$ DNA in the genome, and where r = 0.3, the ratio of $\alpha\beta$ DNA in the chromocenter to that at 87C, as determined by in situ hybridization with Dm703; 33 kb is a maximum value because of the potential for underreplication at the chromocenter (Spear and Gall, 1973). The upper limit is then equal to $\{(33\text{–}12\ L_{\alpha\beta})\ (L_{\alpha\gamma}/L_{\alpha}) + 12\ L_{\alpha}\}/25$, where L_{α}, $L_{\alpha\beta}$ and $L_{\alpha\gamma}$ are the lengths of the α element (0.49 kb), the $\alpha\beta$ unit (1.59 kb) and the $\alpha\gamma$ unit (1.36 kb).]

These simple correlations prompt the speculation that the γ sequences may provide a connection between the two repeated gene families that coordinates their expression. Consequently, it will be of considerable interest to define the topographic relationship between the genes for the 70,000 dalton protein and the γ sequences at the molecular level, and to determine whether the transcripts obtained from either of the two gene families include γ sequences. Our present data on the hybridization of heat shock RNA to the $\alpha\gamma$ unit in Dm704 suggest that at least some γ sequences are transcribed, but indicate nothing about the source of such transcription.

Experimental Procedures

Labeling of Cell Cultures

The Kc_0 line of D. melanogaster cells obtained from W. Gehring is a subline of the Kc line established by Eschalier and Ohanessian (1970) which was adapted to grow in the absence of serum. Kc_0 cells were grown in suspension at 25°C in low phosphate medium as described by Rubin and Hogness (1975). Unless otherwise specified, heat shock cells were obtained by circulating 37°C water around a jacketed spinner flask containing 200 ml of a 25°C culture when it had attained a density of 6–9 × 10^6 cells per ml. 10 min after this temperature jump, 50 mCi of ^{32}P inorganic phosphate or 1–2 mCi of ^{3}H-uridine were added to 200 ml of culture. After 70 min of further incubation at 37°C, the cells were chilled to 0°C and harvested by centrifugation. Cultures at 25°C were also labeled for 70 min at the above cell density and radioactivity.

Nucleic Acid Preparations

^{32}P-labeled cytoplasmic poly(A)$^+$ RNA was prepared from Kc_0 cells labeled at 25 or 37°C as described by Young and Hogness (1977), the heat shock RNA exhibiting a specific radioactivity of approxi-

mately 10^6 cpm/μg. The ^{3}H polysomal poly(A)$^+$ RNA used in Figures 5 and 7 was prepared from polysomes purified by zone sedimentation of a cytoplasmic extract prepared from cells labeled at 37°C as described for the isolation of the cytoplasmic RNA. The polysomes were sedimented into a 0.5–1.5 M linear sucrose gradient in 0.25 M KCl, 0.025 M $MgCl_2$, 0.05 M Tris-HCl (pH 7.4) for 150 min at 25,000 rpm in a Beckman-Spinco SW27 rotor at 0°C. The polysome fractions were pooled and phenol-treated, and the poly(A)$^+$ RNA was isolated as described for the isolation of the cytoplasmic poly(A)$^+$ RNA. All other preparations of labeled polysomal poly(A)$^+$ RNA were prepared from Kc_0 cells that were heat-shocked and labeled as above, except that the temperature jump was from 25 to 36°C, and the 2 mCi of ^{3}H-uridine were added 5 min after the jump. The polysomes from such cells were obtained as a pellet prepared by centrifugation of 6 ml of the cytoplasmic extract through a 3 ml block of 52% (w/v) sucrose in 0.25 M KCl, 0.025 M $MgCl_2$, 0.05 M Tris-HCl (pH 7.4) for 2.5 hr at 45,000 rpm in a Beckman-Spinco Ti50 rotor at 2°C. Each pellet was suspended in 1 ml of extraction buffer [1% Sarkosyl (w/v) in 0.1 M NaCl, 0.03 M Na_2 EDTA, 0.1 M Tris-HCl (pH 8.9); Alwine et al., 1977] containing 300 μg/ml proteinase K. After a 30 min incubation at 37°C, the mixture was extracted 3 times with phenol-CHC_3-isoamyl alcohol, and the poly(A)$^+$ RNA in the aqueous phase was isolated as described by Young and Hogness (1977).

D. melanogaster (Oregon R) embryonic DNA and plasmid DNAs were isolated as described previously (Wensink et al., 1974) with minor modifications (D. J. Finnegan, G. M. Rubin, J. Bower and D. S. Hogness, manuscript in preparation). These DNAs were ^{32}P- or ^{3}H-labeled by nick translation (Rigby et al., 1977). ^{3}H- and ^{32}P-labeled RNAs complementary to plasmid DNAs (cRNAs) were synthesized in vitro as described by Wensink et al. (1974). ^{3}H- and ^{32}P-labeled DNAs complementary to poly(A)$^+$ RNAs (cDNAs) were synthesized in vitro as described by Myers, Spiegelman and Kacian (1977), except that the sodium pyrophosphate was omitted from the reaction mixture, and the reactions were carried out at 42°C instead of 37°C.

Enzymes and Other Materials

The source and method of preparation of the Bam HI, Eco RI, Hind III and Sst I restriction endonucleases used here were described by Rambach and Hogness (1977); Sma I was purified as described by R. Greene and C. Mulder (personal communication), and Sal I as described by J. R. Arrand, P. A. Myers and R. J. Roberts (personal communication). E. coli RNA polymerase (Berg, Barrett and Chamberlin, 1971) and SI nuclease (Vogt, 1973) were gifts from D. Brutlag and T. E. Shenk, respectively, and were prepared according to the indicated references. AMV reverse transcriptase was a gift from R. Padgett, who further purified a preparation from J. Beard (Life Sciences) by passing it over a G-150 Sephadex column. Such a preparation makes full-length copies of heat shock mRNAs (J. T. Lis, unpublished observations). Oligo(dA)- and oligo(dT)-cellulose (T-3) were purchased from Collaborative Research.

Construction and Identification of Hybrid Plasmids

cDm plasmids were constructed by insertion of sheared D. melanogaster DNA at the Eco RI site of the Col EI vector plasmid (Hershfield et al., 1974) by a modification of the poly(dA)-poly(dT) connector method of Wensink et al. (1974), in which poly(dT) tails were added to the Col EI Eco RI termini and poly(dA) tails were added to the D. melanogaster DNA (D. J. Finnegan, G. M. Rubin, J. Bower and D. S. Hogness, manuscript in preparation). Hybrid plasmids were propagated in the E. coli K12 strain HB101 (Boyer and Roulland-Dussoix, 1969) and screened for heat shock genes by colony hybridization with ^{32}P cytoplasmic poly(A)$^+$ heat shock RNA, using a modification of the Grunstein and Hogness (1975) method described by Young and Hogness (1977). Clones containing rDNA which contaminated the set that pass the above screen were identified by a second colony hybridization, using ^{32}P-cRNA to pDm103 (Glover and Hogness, 1977; White and Hogness, 1977) as the probe. cDm plasmids that survived both screens were finally checked for the presence of heat shock genes by hybridization with ^{32}P cytoplasmic poly(A) RNA from both heat-shocked and 25°C cells, usually by quantitative filter hybridization to isolated plasmid DNA fixed to nitrocellulose filters (see below). cDm702, cDm703 and cDm704 were chosen from the set that passed all of the above tests.

mkDm750 consists of the 1.3 kb Hind III fragment of Dm704 (Figure 1) which was inserted into the single Hind III site of the vector plasmid pML21 (Hershfield et al., 1976). pML21 (miniCol EI-kan) consists of the miniCol EI plasmid, pVH51, joined to an Eco RI fragment from pSC105 that determines kanamycin resistance; the Hind III site is located in this kan fragment (Glover and Hogness, 1977). The ligated insertion of Hind III fragments from cDm704 and the subsequent transformation of HB101 by the resulting hybrid plasmids was accomplished essentially as described by Glover and Hogness (1977). mkDm750 was identified by Hind III digestion and electrophoretic analysis of the two resulting fragments, which are the 1.3 kb fragment from cDm704 and linear pML21 DNA.

All plasmids containing Dm segments were propagated in EKI host-vector systems under P2 containment as defined and recommended for this class of experiments by the NIH (June 1976).

Hybridization Procedures

Quantitative filter hybridization of labeled RNA to plasmid DNA fixed to 13 mm Millipore (HAWP) nitrocellulose filters was carried out essentially as described by Glover et al. (1975), except that the solvent (0.1 ml per filter) also contained 1 mg/ml poly(A), and the temperature of hybridization was 50°C. Hybridization of labeled RNA to DNA transferred from agarose gels to nitrocellulose strips (Southern, 1975) was carried out in 5 x SSCP [SSCP = 120 mM NaCl, 15 mM Na citrate, 20 mM Na phosphate (pH 7)], 50% formamide, 1 mg/ml poly(A) and 50 μg/ml E. coli tRNA for 40 hr at 50°C. Strips were washed twice in 5 x SSCP, 50% formamide for 20 min at 45°C, and 3 times in 2 x SSC at room temperature. They were then treated with pancreatic RNAase at 10 μg/ml for 30 min in 2 x SSC at room temperature and washed again in 2 x SSC prior to autoradiography or fluorography.

When labeled cDNA was hybridized to DNA transferred from gels to nitrocellulose strips (Figure 2), the strips were pretreated by soaking in Dx (0.02% ficoll, 0.02% polyvinyl pyrrolidone, 0.02% bovine serum albumin; Denhardt, 1966) plus 5 x SSC, 100 μg/ml sonicated calf thymus DNA and 0.1% SDS at 65°C for 8 hr. Hybridization was carried out for 40 hr at 37°C in Dx, 50% formamide, 5 x SSC, 200 μg/ml poly(U) contained in heat-sealed plastic bags on a rocking platform. Filters were washed for several hours at 44°C in 50% formamide, 5 x SSC, 0.2% SCS, and then twice for 1 hr in 2 x SSC at 60°C and once for 30 min in 0.5 x SSC at 60°C. Hybridization of labeled plasmid DNA to DNA on strips (Figure 8) was carried out in essentially the same manner, except that poly(A) was substituted for poly(U), 0.1% SDS was added to the hybridization solvent and the time of hybridization was 50 hr. An extra wash in 0.1 x SSC at 65°C was also included in the procedure to minimize background.

Electron Microscopy

Heteroduplex formation and sample preparation for the electron microscope were carried out as described by Lis and Schlief (1975), except that the reassociation was at 37 instead of 25°C. The poly(dA) and poly(dT) connectors in the hybrid DNA strands snap back to form hairpin structures at 25°C, but not at 37°C, and these hairpins provide a conformational constraint that inhibits heteroduplex formation.

R loops were formed as previously described (Thomas, White and Davis, 1976; White and Hogness, 1977), except that the incubation of the class II RNA with the Sma I-cut cDm704 DNA (Figure 6) was for 22 hr at 44°C, which is 2° below the Tm for cDm704 DNA in the solvent used [70% formamide, 0.1 M NaCl,

0.01 M EDTA, 0.1 M Tris-HCl (pH 7.4)]. The reaction solution was diluted 10 fold into the hyperphase solution of Lis and Schlief (1975) minus the cytochrome C and held at 23°C for 10 min; cytochrome C was added to 50 μg/ml and then spread for electron microscopy as described previously (Lis and Schlief, 1975).

Standards for DNA Lengths

The standards used in the electrophoretic determination of duplex DNA lengths consist of the fragments generated by Eco RI, Bam HI double digestion of λ phage DNA (Haggerty and Schlief, 1976), whereas that used in the electron microscopic analysis of duplex lengths is pSC101 plasmid DNA (Wensink et al., 1974). In both cases, the standard lengths are based on a value of 46.5 kb for the length of λ DNA (Davidson and Szybalski, 1971). Recent evidence indicates that a better value for the λ length is 49.0 kb (P. Philippsen and R. Davis, personal communication); hence the values for duplex DNA lengths given in this paper should be multiplied by 1.054 to bring them into correspondence with this new λ length. M13 phage DNA (6.6 kb; Marvin and Hohn, 1969) was used as the single-stranded length standard for electron microscopic determinations. The length of each R loop was taken as the average of the lengths of the single-stranded DNA and RNA-DNA duplex elements, using the above single- and double-stranded standards.

Acknowledgments

This work was supported by research grants from the American Cancer Society, the NSF and the NIH. J. T. L. was a Helen Hay Whitney fellow.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 2, 1978

References

Alwine, J. C., Kemp, D. J. and Stark, G. R. (1977). Proc. Nat. Acad. Sci. USA *74*, 5350-5354.

Ashburner, M. (1970). Chromosoma *31*, 356-376.

Berg, D., Barrett, K. and Chamberlin, M. (1971). In Methods in Enzymology, *21D*, L. Grossman and K. Moldave, eds. (New York: Academic Press), pp. 506-519.

Bonner, J. J. and Pardue, M. L. (1976). Cell *8*, 43-50.

Boyer, H. W. and Roulland-Dussoix, D. (1969). J. Mol. Biol. *41*, 459-472.

Davidson, N. and Szybalski, W. (1971). In The Bacteriophage Lambda, A. D. Hershey, ed. (New York: Cold Spring Harbor Laboratory), pp. 45-82.

Denhardt, D. T. (1966). Biochem. Biophys. Res. Commun. *23*, 641-646.

Duesberg, P. and Vogt, P. K. (1973). J. Virol. *12*, 594-599.

Eschalier, G. and Ohanessian, A. (1970). In Vitro *6*, 162-172.

Fellner, P. (1974). In Ribosomes, M. Nomura, A. Tissières and P. Lengyel, eds. (New York: Cold Spring Harbor Laboratory), pp. 169-191.

Fiers, W., Contreras, R., Duerinick, F., Haegeman, G., Iserentant, D., Merregaert, J., MinJou, W., Molemans, F., Racymaekers, A., Vanden Berghe, A., Volckaert, G. and Ysebaert, M. (1976). Nature *260*, 500-507.

Finnegan, D. J., Rubin, G. M., Young, M. W. and Hogness, D. S. (1978). Cold Spring Harbor Symp. Quant. Biol. *42*, in press.

Gilbert, W. (1976). In RNA Polymerase, R. Losick and M. Chamberlin, eds. (New York: Cold Spring Harbor Laboratory), pp. 193-205.

Glover, D. M. and Hogness, D. S. (1977). Cell *10*, 167-176.

Glover, D. M., White, R. L., Finnegan, D. J. and Hogness, D. S. (1975). Cell *5*, 149-157.

Grunstein, M. and Hogness, D. S. (1975). Proc. Nat. Acad. Sci. USA *72*, 3961-3965.

Haggerty, D. M. and Schleif, R. F. (1976). J. Virol. *18*, 659-663.

Henikoff, S. and Meselson, M. (1977). Cell *12*, 441-451.

Hershfield, V., Boyer, H. W., Yanofsky, C., Lovett, M. A. and Helinski, D. R. (1974). Proc. Nat. Acad. Sci. USA *71*, 3455-3459.

Hershfield, V., Boyer, H. W., Chow, L. and Helinski, D. R. (1976). J. Bacteriol. *126*, 447-453.

Ish-Horowicz, D., Holden, J. J. and Gehring, W. J. (1977). Cell *12*, 643-652.

Jacq, B., Jourdan, R. and Jordan, B. R. (1977). J. Mol. Biol. *117*, 785-795.

Jamrich, M., Greenleaf, A. L. and Bautz, E. K. (1977). Proc. Nat. Acad. Sci. USA *74*, 2079-2083.

Jordan, B. R., Jourdan, R. and Jacq, B. (1976). J. Mol. Biol. *101*, 85-105.

Karp, R. W. (1978). Ph.D. Thesis, Stanford University, Stanford, California.

Lewis, M., Helmsing, P. J. and Ashburner, M. (1975). Proc. Nat. Acad. Sci. USA *72*, 3604-3608.

Lifton, R. P., Goldberg, M. L., Karp, R. W. and Hogness, D. S. (1978). Cold Spring Harbor Symp. Quant. Biol. *42*, in press.

Lis, J. T. and Schleif, R. F. (1975). J. Mol. Biol. *95*, 395-407.

Marvin, D. A. and Hohn, B. (1969). Bacteriol. Rev. *33*, 172-193.

McKenzie, S. L., Henikoff, S. and Meselson, M. (1975). Proc. Nat. Acad. Sci. USA *72*, 1117-1121.

Mirault, M.-E., Goldschmidt-Cleremont, M., Moran, L., Arrigo, A. P. and Tissières, A. (1978). Cold Spring Harbor Symp. Quant. Biol. *42*, in press.

Moran, L., Mirault, M.-E., Arrigo, A. P., Goldschmidt-Clermont, M. and Tissières, A. (1978). Phil. Trans. Roy. Soc. B, in press.

Morrow, J. F. (1974). Ph.D. Thesis, Stanford University, Stanford, California.

Myers, J. C., Spiegelman, S. and Kacian, D. L. (1977). Proc. Nat. Acad. Sci. USA *74*, 2840-2843.

Rambach, A. and Hogness, D. S. (1977). Proc. Nat. Acad. Sci. USA *74*, 5041-5045.

Rasch, E. M., Barr, H. J. and Rasch, R. W. (1971). Chromosoma *33*, 1-18.

Rigby, P. W. J., Dieckmann, M., Rhodes, C. and Berg, P. (1977). J. Mol. Biol. *113*, 237-251.

Ritossa, F. M. (1962). Experientia *18*, 571-573.

Rubin, G. M. and Hogness, D. S. (1975). Cell *6*, 207-213.

Rudkin, G. T. (1972). In Results and Problems in Cell Differentiation, *4*, W. Beermann, ed. (New York: Springer-Verlag), pp. 59-85.

Schedl, P., Artavanis-Tsakonas, S., Steward, R., Gehring, W., Mirault, M.-E., Goldschmidt-Clermont, M., Moran, L. and Tissières, A. (1978). Cell *14*, 921-929.

Southern, E. M. (1975). J. Mol. Biol. *98*, 503-517.

Spear, B. B. and Gall, J. G. (1973). Proc. Nat. Acad. Sci. USA *70*, 1359-1363.

Spradling, A., Penman, S. and Pardue, M. L. (1975). Cell *4*, 395-404.

Spradling, A., Pardue, M. L. and Penman, S. (1977). J. Mol. Biol. *109*, 559-587.

Thomas, M., White, R. L. and Davis, R. W. (1976). Proc. Nat. Acad. Sci. USA *73*, 2294-2298.

Tissières, A., Mitchell, H. K. and Tracy, U. M. (1974). J. Mol. Biol. *84*, 389-398.

Van Holde, K. E. and Hill, W. E. (1974). In Ribosomes, M. Nomura, A. Tissières and P. Lengyel, eds. (New York: Cold Spring Harbor Laboratory), pp. 53–91.

Vogt, V. M. (1973). Eur. J. Biochem. *33*, 192–200.

Wensink, P. C., Finnegan, D. J., Donelson, J. E. and Hogness, D. S. (1974). Cell *3*, 315–325.

White, R. L. and Hogness, D. S. (1977). Cell *10*, 177–192.

Young, M. W. and Hogness, D. S. (1977). In Molecular Approaches to Eucaryotic Genetic Systems, ICN-UCLA Symposia on Molecular and Cellular Biology, *VIII*, G. Wilcox, J. Abelson and C. F. Fox, eds. (New York: Academic Press), pp. 315–331.