Proc. Natl. Acad. Sci. USA
Vol. 78, No. 10, pp. 6196–6200, October 1981
Biochemistry

DNA sequence analysis reveals extensive homologies of regions preceding *hsp70* and *αβ* heat shock genes in *Drosophila melanogaster*

(transcription/genetic regulation/promoters/transposition/evolution)

ROGER W. HACKETT AND JOHN T. LIS*

Department of Biochemistry, Molecular and Cell Biology, Division of Biological Sciences, Cornell University, Ithaca, New York 14853

Communicated by David S. Hogness, July 20, 1981

ABSTRACT **Two kinds of RNA are synthesized at the *87C1* chromosomal locus of *Drosophila melanogaster* in response to heat shock. One of these codes for the major heat shock protein, hsp70; the other, *αβ* RNA, derives from tandemly repeated *αβ* units consisting of adjacent *α* and *β* DNA elements and has no identified translation product. Another DNA element, *γ*, flanks the 5′ ends of some *αβ* units. Here we report the complete nucleotide sequence of the 617-base-pair *α* and the 733-base-pair *γ* element as well as a portion of the longer *β* element. Sequence comparisons between the *γ* element and the two *hsp70* genes at *87C1* reveal that the 406 base pairs of *γ* immediately upstream from the 5′ end of the *αβ* unit exhibit 97.5% homology with the sequences at and upstream from the 5′ end of the *hsp70* genes. A similar homology also exists between *γ* and an *hsp70* gene present at another heat shock locus, *87A7*, which contains no *αβ* units. These results, in conjunction with previous observations, strongly suggest that the coordinate induction by heat shock of the *hsp70* and *αβ* genes is a consequence of their homologous 5′ flanking sequences. We propose that this extraordinary degree of sequence conservation stems from the recent transposition of *αβ* DNA to the *87C1* locus, an event that brought *αβ* sequences adjacent to, and under the regulation of, the *hsp70* control element.**

Heat shock of *Drosophila melanogaster* elicits a characteristic response in all tissues that results in the coordinate synthesis of a specific set of proteins. Concomitantly, the synthesis of most other cellular proteins ceases (1). The altered pattern of protein synthesis induced during heat shock reflects regulation at the transcriptional, translational, and RNA processing levels. The predominant heat shock protein (hsp70) derives from a repeated gene located at the two chromosomal sites responsible for the *87A* and *87C* heat shock puffs (2–4). Fig. 1 shows the arrangement of the *hsp70* genes at each chromosomal locus. These genes have common sequence elements within a nontranscribed region situated upstream from the mRNA coding sequences (4, 7). In this paper we shall adopt the nomenclature of Artavanis-Tsakonas *et al.* (4) and refer to those DNA sequences homologous to hsp70 mRNA as z_c (z coding) sequences and to their nontranscribed 5′ flanking region as z_{nc} (z noncoding sequences. The three copies of the *hsp70* gene at *87C* are not the only sequence elements transcribed at this locus in response to heat shock. A set of tandemly repeated units, called *αβ*, are also transcribed at this locus during heat shock. This set is part of an unusual moderately repetitive DNA that is present in 30 copies per haploid genome (6). Half of these copies are arranged in a dispersed manner at a few euchromatic sites and in the heterochromatic regions located near the chromosomal centromeres. The other half are found at the *87C* locus, and only these copies are arranged in tandem arrays. Moreover, these tandemly repeated *αβ* units are the only ones transcribed in response to heat shock (9).

In addition to the repeated *αβ* units, the *D. melanogaster* chromosomal DNA cloned in the cDm704 plasmid (Figs. 1 and 2) contains another repeat unit termed *αγ*, in which the *β* element is replaced by a nonhomologous *γ* element (6). The *γ* element is located at the 5′ end of those sequences homologous to *αβ* RNA. In this report we present the complete nucleotide sequence of a 1.35-kilobase (kb) *Hin*dIII *αγ* unit and a portion of the adjacent *αβ* unit in cDm704. This analysis precisely defines the boundaries of all three elements and allows comparison of the *γ* sequences lying immediately upstream from the *αβ* unit with those lying upstream from the 5′ ends of the *hsp70* genes at both chromosomal loci (10–12). A remarkable sequence homology between these upstream regions is defined and evaluated.

MATERIALS AND METHODS

End-Labeling with Reverse Transcriptase, Fragment Isolation, and DNA Sequence Determination. Plasmid DNA was prepared and phenol-extracted prior to end-labeling with reverse transcriptase (J. Beard, Life Sciences, St. Petersburg, FL) as described (3). After digestion with a second restriction endonuclease to separate the labeled ends, the desired fragments were size-fractionated on polyacrylamide gels and the DNA was extracted (13). DNA fragments were subjected to sequence determination by the Maxam and Gilbert method (14) as modified by Smith and Calvo (15), and the products were separated on 80-cm sequencing gels (15, 16). Approximately 400 nucleotides can be read from these long gels, which simplified the sequencing strategy such that the 1.35-kb *αγ* fragment could be completely solved by using a rather low-resolution restriction map (Fig. 2).

Work with bacteria containing recombinant plasmids was performed under P2/EK1 conditions.

RESULTS

Extensive Sequence Homology Exists Between a Portion of *γ* and Copies of *z* at *87A* and *87C*. We determined the sequence of the 1.35-kb *Hin*dIII fragment subcloned from the hybrid plasmid cDm704 by using the chemical reactions essentially as described by Maxam and Gilbert (14). The nucleotide sequence of the entire *αγ* fragment was compared to the *z* gene sequences of Török and Karch (10) and Karch *et al.* (12) by using the computer program of Queen and Korn (17). The *αγ* sequence is presented in Fig. 3 with a few of its salient features highlighted. The *z* gene sequence of plasmid 56H8 (10), which is derived

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

Abbreviations: kb, kilobase(s); bp, base pair(s).

* To whom reprint requests should be addressed.

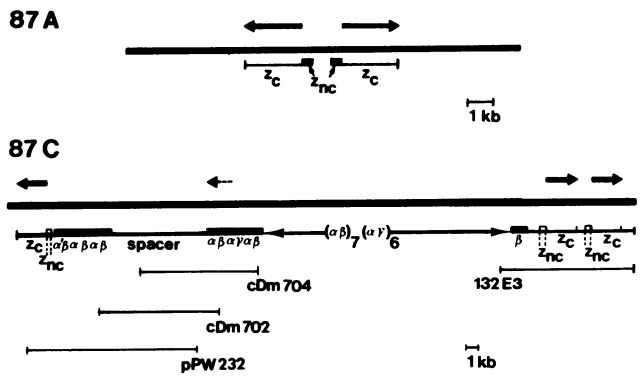

FIG. 1. Genomic organization of heat shock genes at the *87A7* and *87C1* loci. The overall organization of these loci has been described in detail (5). The solid bars represent sections of chromosome 3. Solid arrows above the bars denote the 5′→3′ orientation and extent of the transcribed sequences for the five *hsp70* genes of most strains. The dashed arrow depicts the 5′→3′ orientation of one probable transcribed $\alpha\beta$ unit, which we propose is regulated by the adjacent γ element. The arrangement of previously defined sequence elements is shown below the solid bars, as are cloned segments derived from this region. The overlapping clones cDm702 and cDm704 (6) and pPW232 (2) represent the genomic organization at the proximal end of *87C1*; cloned segment 132E3 (4, 7) represents the distal portion. The z_c element is the region homologous to hsp70 mRNA; z_{nc} is a 5′ flanking element common to *hsp70* genes at both *87A* and *87C*. Although polymorphism in the arrangement exists, we have depicted the organization of the major component of Oregon R DNA (5). Cloned segment 56H8 represents a variant of the *87A* organization shown here. The strain with the 56H8 segment possesses inverted repeat copies of *z* that are farther apart, presumably due to an insertion of a repetitive DNA sequence (8). At *87C* the tandem $\alpha\beta$ repeat units occur between the inverted repeats of *z*. Toward the proximal (left) end of chromosome 3, two such $\alpha\beta$ arrays are separated by approximately 7.2 kilobases (kb) of spacer DNA (6). The left array contains a modified α element, α' (2). Additional $\alpha\beta$ and $\alpha\gamma$ repeats occur within approximately 19 kb of undefined DNA, indicated by $(\alpha\beta)_7$ $(\alpha\gamma)_6$, as estimated by Lis *et al.* (9).

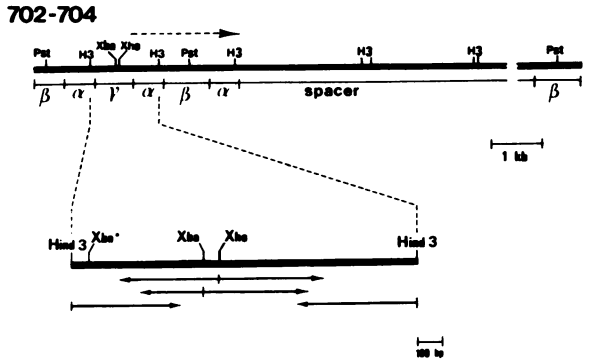

FIG. 2. Restriction map of cloned segments cDm702-cDm704 and sequence determination strategy. The dashed arrow designates the 5′→3′ orientation of the $\alpha\beta$ RNA homologous region. Overlapping segments cDm702 and cDm704 are combined here and redrawn from Lis *et al.* (6). Note that the chromosomal DNA is oriented oppositely from that in Fig. 1. Restriction endonucleases are abbreviated as follows: H3, *Hin*dIII; Pst, *Pst* I; Sal, *Sal* I; Xba, *Xba* I; Xho, *Xho* I. The enlarged segment denotes the 1.35-kb *Hin*dIII fragment from cDm704, and a few key restriction endonuclease sites are shown. Each arrow below the bar represents the direction and extent of a nucleotide sequence determined from a single end-labeling reaction. The sequence of each stretch was determined independently at least twice. The agreement of these independent experiments and of the sequence of opposite strands for some portions permits us to estimate that the reported sequence is >99% accurate. We are especially confident of the sequence containing the important γ_z region (see text). Xba* denotes an *Xba* I site within the sequence that is detectable in genomic DNA but is modified in *Escherichia coli*.

from the *87A* locus, shows extensive homology to a 548-base-pair (bp) region of $\alpha\gamma$ with an average of 83.2% of the nucleotides matching. This region is indicated by brackets in Fig. 3 and extends from positions −484 to 64. The corresponding homologous sequence in cloned segment 56H8 begins upstream from and extends through the entire z_{nc} element, and includes 64 bp of the z_c (RNA homologous) region (10). A similar region of homology exists between $\alpha\gamma$ and both copies of *z* in the cloned segment 132E3, which derives from the *87C* locus. This region starts at −342 and similarly extends to +64, as indicated by the large type in Fig. 3. Although shorter and included within the preceding homology region, these 406 bp exhibit a much greater homology (97.5%) to the two *z* genes in 132E3, where the corresponding region begins near the start of z_{nc}, extends through it, and includes 65 bp of z_c. We shall refer to the 406 bp that are common to both homology regions as the γ_z region because it maps within γ immediately upstream from the 5′ end of the adjacent $\alpha\beta$ unit, as is shown in the following section. The region containing the remaining 327 bp of γ is called γ_r because it contains most of the short repeat elements—direct and inverted—detected in $\alpha\gamma$ by the Queen and Korn (17) program. We ignore these repeats here to focus attention on the more remarkable homologies of the γ_z region.

The γ_z region includes two copies of the Hogness box sequence, 5′ T-A-T-A-A-A-T-A 3′ (18). One of these is located 25 bp upstream from the 5′ ends of the hsp70 transcripts from both *87A* and *87C* (10–12). Both the position of this sequence relative to the start of transcription and the fact that it is flanked by G+C-rich regions (19) indicates that this particular copy is most likely to serve as part of the functional transcriptional start signal for $\alpha\beta$ as well as hsp70 transcripts. The γ_z region also includes a sequence, C-A-A-T-T-C-A, that is similar to the consensus "capping sequence" (20) and is located 29 bp downstream from the Hogness box sequence. These comparisons suggest that transcription of the $\alpha\beta$ gene is initiated at a point in γ_z equivalent to the transcription initiation site for the *hsp70* genes (position +1 in Fig. 3). We therefore predict that the first 64 nucleotides in the $\alpha\beta$ RNA will derive from γ_z and, hence, will mimic those in the hsp70 mRNA. The fact that γ_z includes a 342-bp region homologous to that flanking the 5′ ends of the *hsp70* genes suggests that the regulatory sequences governing transcription initiation are the same for both $\alpha\beta$ and *hsp70* genes and reside within this region.

Sequence Comparison of the $\alpha\beta$ and $\alpha\gamma$ Units Precisely Determines the Boundaries of α, β, and γ Elements. The α element is defined as the region of homology between the $\alpha\beta$ and $\alpha\gamma$ units; the β and γ elements are defined as the nonhomologous regions in the respective units. Homology and nonhomology, and hence the boundaries of the three elements, were originally determined by electron microscopic examination of heteroduplexes between the two units (6). To map the boundaries at the nucleotide level we determined the sequence of appropriate regions in the cDm704 *Hin*dIII fragment that comprises the $\alpha\beta$ unit and compared these sequences with those of the adjacent $\alpha\gamma$ *Hin*dIII fragment described above.

Beginning at the *Hin*dIII sites within the α element of each fragment, we searched in both directions for the position at which the $\alpha\beta$ and $\alpha\gamma$ sequences markedly diverged. These positions define both ends of the α element in both fragments and hence both ends of their respective β and γ elements. Fig. 4 shows these positions in the fragments oriented as in Figs. 2 and 3. The position termed "left junction" is 105 bp from the left end of both fragments and represents the junction between the left end of β or γ and the right end of α; similarly, "right

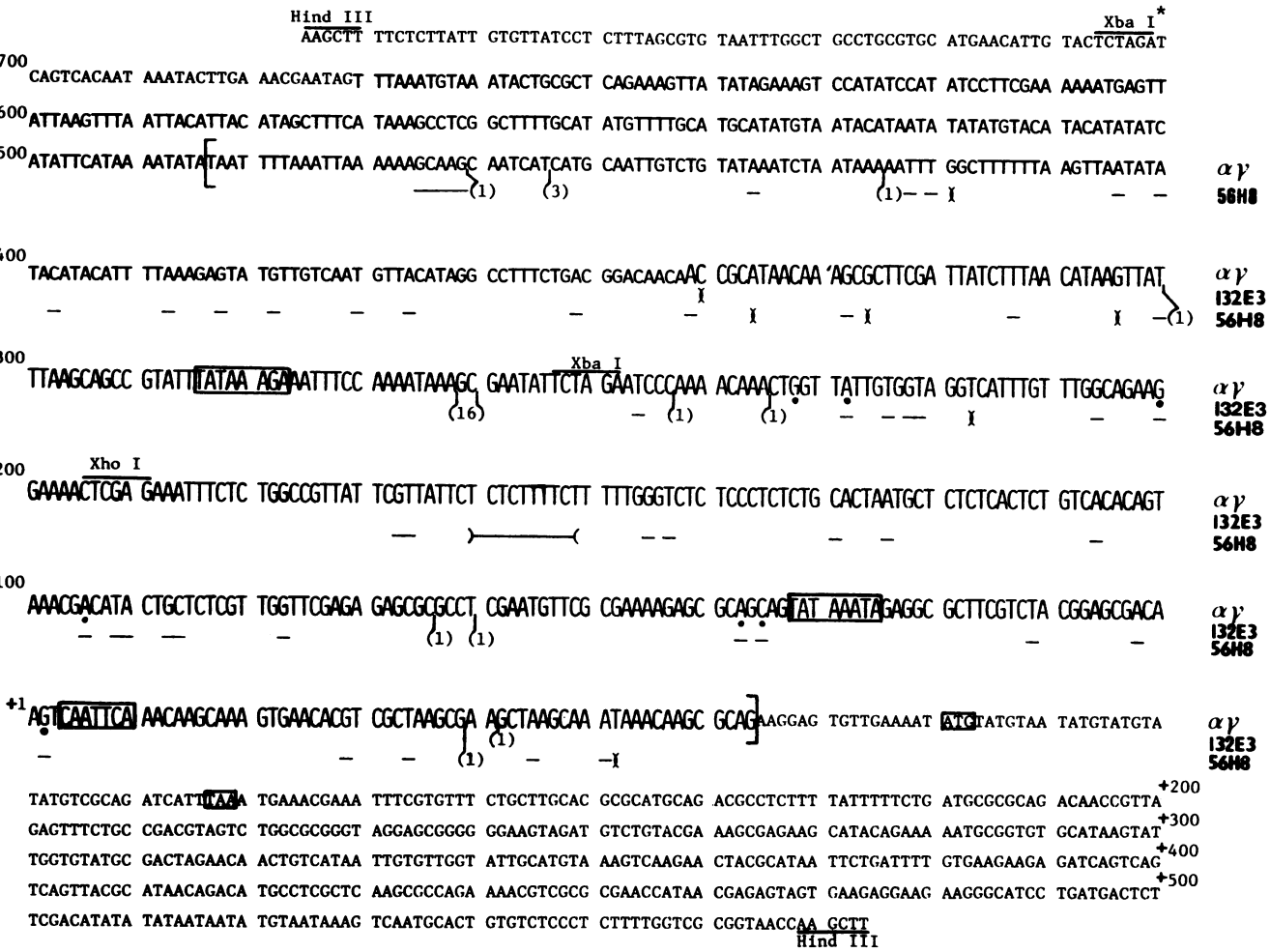

FIG. 3. Nucleotide sequence of the $\alpha\gamma$ fragment. Small type, nucleotides within α; intermediate type, γ_r region; large type, γ_z region. The nucleotide corresponding to the point of hsp70 mRNA transcription initiation as represented by plasmids 56H8 and 132E3 is assigned position +1 (10, 12). Those parts of the $\alpha\gamma$ sequence homologous to z sequences in 132E3 and in 56H8 correspond to the γ_z region and to the larger region delimited by brackets, respectively. For each of these sequence comparisons we indicate below our sequence those nucleotides that differ in 56H8 or 132E3 (see right margin). •, Different nucleotide in 132E3; –, different nucleotide in 56H8; numeral in parentheses, insertion of that number of nucleotides in a z sequence;)(, nucleotide deleted in a z sequence;)——(, position of a 10-bp deletion in 56H8. Only key restriction sites are shown. The two Hogness box sequences are enclosed, as is the first AUG codon downstream from position +1 and the first termination codon in this reading frame. The presumptive "capping sequence" is also boxed. To scan the $\alpha\gamma$ sequence for homologies to 56H8 and 132E3, we used the following parameters with the computer program of Queen and Korn (17): MINMATCH = 10, EXPECT = 10, LOOPLENGTH = 3, MINRATIO = 0.75, MAXLEN = 5000. We allowed the program to generate values for MAXPROB (1.59×10^{-5} and 2×10^{-5} for the two comparisons).

junction," which is 512 bp from the right ends of the fragments, represents the boundary between the right end of β or γ and the left end of α. The right end of the γ element defined by this $\alpha\beta/\alpha\gamma$ comparison corresponds to position +62 in Fig. 3 and therefore falls short by only the A-G dinucleotide (at +63, +64) of the right end of the γ_z region defined by the $z/\alpha\gamma$ comparison of Fig. 3, which is reemphasized at the bottom of Fig. 4. The α and γ elements defined by the junctions in Fig. 4 are 617 and 733 bp long, respectively.

Possible Translation Products of the $\alpha\beta$ RNA. Although $\alpha\beta$ RNA is found in preparations of polysomal poly(A)$^+$RNA of heat shock-induced cells (2, 6), attempts to identify a heat shock protein encoded by this RNA have been unsuccessful (2, 21). It therefore was of interest to examine the sequence given in Fig. 3 for possible translation products. If it is assumed that transcription is initiated in γ_z at position +1 to yield the $\alpha\beta$ RNA, then the first three AUG codons are located 81, 85, and 92 nucleotides from the 5′ end of that RNA, each in a different reading frame. The reading frame of the first of these, which is the most likely initiator of translation, allows for a polypeptide of only 12 amino acid residues. The second is followed immediately by a termination codon, and the third allows for a polypeptide of only 30 amino acid residues. Thus, the two polypeptides allowed by the three reading frames are too small to have been detected reliably in the previous studies. However, we cannot eliminate the alternative that RNA splicing reactions rearrange the sequence in the primary transcript to produce $\alpha\beta$ RNAs with a more extensive open reading frame. Although no evidence for splicing of $\alpha\beta$ RNAs was obtained from electron microscopic observation of their R loops (6), these studies were insufficient to eliminate that possibility, particularly of splicing events near the 5′ end.

DISCUSSION

We have shown that the 406 bp of γ_z, which lies immediately upstream from an $\alpha\beta$ unit, exhibit a near perfect homology with a conserved region in the *hsp70* genes that includes the 345 ± 4 bp immediately upstream from their 5′ end (z_{nc}) and the

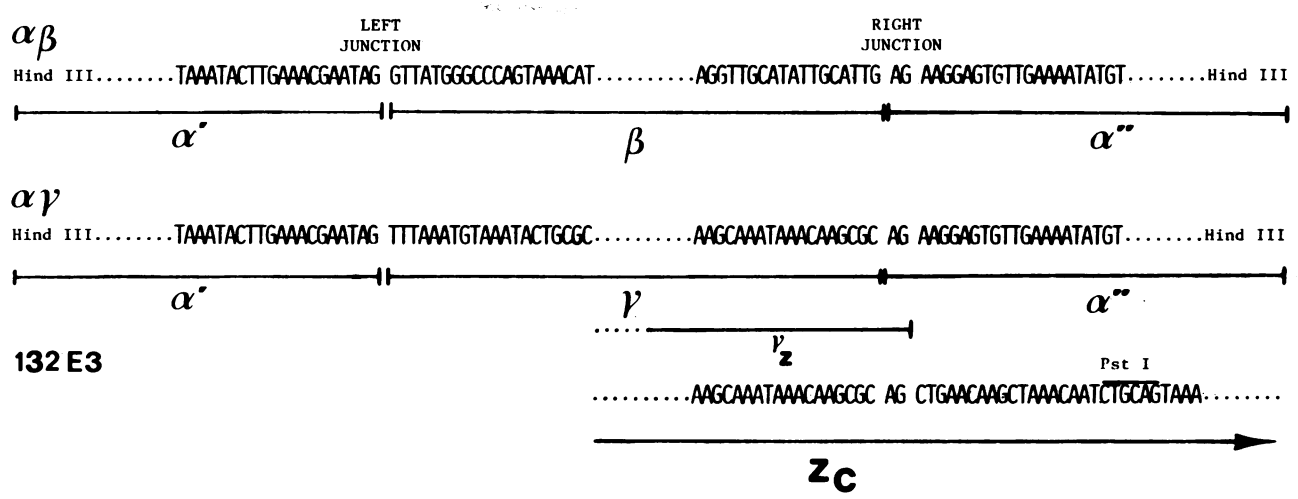

FIG. 4. Comparison of the $\alpha\gamma$ and $\alpha\beta$ sequences at the $\alpha\gamma$ junctions. The sequences are displayed from left to right in accord with their orientations in the cDm704 map in Fig. 2. Because *Hin*dIII cleaves asymmetrically within $\alpha\gamma$, α' designates the 105-bp portion and α'' designates the 512-bp portion. We include a portion of the z_c sequence of Karch *et al.* (12), derived from cloned segment 132E3, to demonstrate that the γ_z element includes the A-G to the right of the γ/α'' junction.

first 64 or 65 bp downstream from that end (the variation is due to whether the comparison is with the *hsp70* gene of *87A* or of *87C*). On the basis of this extensive homology, particularly with z_{nc}, we suggested that the regulatory sequences governing transcription initiation of both the $\alpha\beta$ and *hsp70* genes during heat shock reside within the 406 bp of γ_z. Two sets of observations provide additional support for this proposition: one set concerning the *hsp70* genes and the other, the $\alpha\beta$ genes.

The multiple copies of the *hsp70* genes exhibit a common sequence organization in which the z_{nc} element is highly conserved among all copies in all strains of *D. melanogaster* studied (7, 8). The maximal extent of the regulatory region for the *hsp70* gene has been localized by two deficiency strains to an interval containing z_{nc} and the 5′ half of the transcribed region. In one strain the deletion removes the 3′ half of one *hsp70* gene. Nonetheless, a truncated 40,000-dalton polypeptide is produced in response to heat shock, indicating that sequences downstream from the middle of the gene are not required for regulation (22). The second strain carries a deletion that removes one of the two genes at *87A* and terminates adjacent to the z_{nc} region of the remaining copy (23). The expression of this copy, when analyzed in a background deleted of all *hsp70* genes at *87C*, is normal, indicating that sequences upstream from z_{nc} are not required.

The $\alpha\beta$ DNAs at *87C* are also transcribed in response to heat shock. Although half of the total $\alpha\beta$ DNAs occur as dispersed sequences at other loci within the chromocenter, strains deleted at *87C1* (and hence lacking the tandemly repeated $\alpha\beta$ units) do not synthesize $\alpha\beta$ RNA in response to heat shock (9). Other studies have shown that the closely related species *D. simulans* possesses $\alpha\beta$ DNA sequences that are not located at the *87C* locus and are not transcribed after heat shock (2). These observations and ours taken together strongly suggest that the unique association of $\alpha\beta$ units with the γ_z elements at *87C* in *D. melanogaster* (6) is responsible for their transcription in response to heat shock.

Are these γ_z (or z_{nc}) sequence elements a general feature of heat shock genes in *Drosophila*? *In situ* hybridization using a cloned 174-bp segment of z_{nc} as probe demonstrates that these sequences are not highly conserved at loci other than *87A* and *87C* (24). The striking sequence conservation between different heat shock genes at *87A* and *87C* is therefore not a universal property required for expression of all the genes in the heat shock gene family.

We conclude with a consideration of the origins of the $\alpha\beta$ genes at the *87C* heat shock locus of *D. melanogaster*. Examination of this locus in other species belonging to the *D. melanogaster* subgroup—chiefly *D. simulans* and *D. mauritiana*—reveals a simpler arrangement that lacks the $\alpha\beta$ DNA and resembles the arrangement at the *87A* locus in *D. melanogaster* (2, 25). However, these species do contain $\alpha\beta$ DNA at other chromosomal loci. This suggests that the $\alpha\beta$ DNA was transposed from other loci into the *D. melanogaster 87C* locus by recent evolutionary events—events that created a new arrangement of the $\alpha\beta$ DNA which is inducible by heat shock.

Perhaps the simplest thought is that some arrangement of the $\alpha\beta$ DNA formed a transposable element, similar to the transposable *copia*-like elements of *D. melanogaster* (26–30), which was then inserted into one of a set of *87C hsp70* genes approximately at its +64 position, thereby usurping its z_{nc} regulatory sequences and creating the region we now call γ_z. Although certain arrangements of the $\alpha\beta$ DNA resemble the structure of *copia*-like elements with their terminal direct repeats of several hundred base pairs (e.g., an $\alpha\beta\alpha$ arrangement), the $\alpha\beta$ DNA in *D. melanogaster* does not appear to transpose—at least not at a frequency comparable to that of the *copia*-like elements. Indeed, *in situ* hybridization of an $\alpha\beta$ probe to polytene chromosomes shows that these sequences occupy the same euchromatic sites in each of three different *D. melanogaster* strains (unpublished data); in contrast, several *copia*-like elements exhibit different chromosomal distributions in these same strains.

An alternative possibility is that the $\alpha\beta$ DNA was introduced into *87C* not directly by its own inherent transposability but rather indirectly by virtue of its close linkage to another element capable of transposition. The 7.2-kb spacer segment that separates the two arrays of $\alpha\beta$ units shown in Fig. 1 is of particular interest in this regard. This spacer contains dispersed repeated sequences found at several other chromosomal loci (6) and hybridizes with the clusters of dispersed repeated elements described by Wensink *et al.* (31). One can thus imagine that some of the DNA in the spacer derives from a transposable element that was first inserted adjacent to the $\alpha\beta$ DNA at another locus and then carried this $\alpha\beta$ DNA to *87C* via a subsequent transposition event. This event itself or later rearrangements within *87C*, perhaps associated with the generation of the tandem arrays of $\alpha\beta$ units, might account for the placement of some $\alpha\beta$ units adjacent to the γ_z sequence of an *hsp70* gene and hence

under heat-shock control. In regard to this alternative, we note that normally stable genes can be subject to rearrangement and transposition when residing next to a transposable element, as represented by the TY1-mediated aberrations of the yeast *his4* gene (32, 33) and by the instabilities associated with the w^c and w^a alleles of the *D. melanogaster* white locus (34–36).

The complexity and present indeterminacy of the mechanism for the introduction of $\alpha\beta$ DNA into *87C* should not obscure the important point that the heat shock-induced transcription of $\alpha\beta$ DNA appears to represent an example of the evolutionary recruitment of new genes into a set controlled by a specific regulatory element.

We thank F. Karch, I. Török, and A. Tissières for communicating their extensive sequence data on the *hsp70* genes prior to publication. We thank J. Beard for avian myeloblastosis virus reverse transcriptase, and G. Fink, V. Vogt, J. Hirch, R. Wu, and D. Hogness for comments on the manuscript. We also thank D. Brutlag and B. Fristensky for advice with the computer analysis. This research was supported by National Institutes of Health Grant GM25232 to J.T.L.

1. Ashburner, M. & Bonner, J. J. (1979) *Cell* **17**, 241–254.
2. Livak, K. J., Freund, R., Schweber, M., Wensink, P. C. & Meselson, M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5613–5617.
3. Schedl, P., Artavanis-Tsakonas, S., Steward, R., Gehring, W. J., Mirault, M.-E., Goldschmidt-Clermont, M., Moran, L. & Tissières, A. (1978) *Cell* **14**, 921–929.
4. Artavanis-Tsakonas, S., Schedl, P., Mirault, M.-E., Moran, L. & Lis, J. (1979) *Cell* **17**, 9–18.
5. Ish-Horowicz, D. & Pinchin, S. M. (1980) *J. Mol. Biol.* **142**, 231–245.
6. Lis, J. T., Prestidge, L. & Hogness, D. S. (1978) *Cell* **14**, 901–919.
7. Moran, L., Mirault, M.-E., Tissières, A., Lis, J., Schedl, P., Artavanis-Tsakonas, S. & Gehring, W. J. (1979) *Cell* **17**, 1–8.
8. Goldschmidt-Clermont, M. (1980) *Nucleic Acids Res.* **8**, 235–252.
9. Lis, J. T., Pinchin, S. M. & Ish-Horowicz, D. (1981) *Nucleic Acids Res.*, in press.
10. Török, I. & Karch, F. (1980) *Nucleic Acids Res.* **8**, 3105–3123.
11. Ingolia, T. D., Craig, E. A. & McCarty, B. J. (1980) *Cell* **21**, 669–679.
12. Karch, F., Török, I. & Tissières, A. (1981) *J. Mol. Biol.* **148**, 219–230.
13. Wu, R., Jay, E. & Roychoudhury, R. (1976) *Methods Cancer Res.* **12**, 87–176.
14. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
15. Smith, D. R. & Calvo, J. M. (1980) *Nucleic Acids Res.* **8**, 2255–2274.
16. Yang, R. C. A. & Wu, R. (1978) *Virology* **92**, 340–352.
17. Queen, C. L. & Korn, L. J. (1980) *Methods Enzymol.* **65**, 595–609.
18. Goldberg, M. (1979) Dissertation (Stanford Univ., Stanford, CA).
19. Gannon, F., O'Hare, K., Perrin, F., LePennec, J. P., Benoist, C., Cochet, M., Breathnach, R., Royal, A., Garapin, A., Cami, B. & Chambon, P. (1979) *Nature (London)* **278**, 428–434.
20. Sures, I., Lowry, J. & Kedes, L. H. (1978) *Cell* **15**, 1033–1044.
21. Ish-Horowicz, D., Holden, J. J. & Gehring, W. J. (1977) *Cell* **12**, 643–652.
22. Caggese, C., Caizzi, R., Morea, M., Scalenghe, F. & Ritossa, F. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 2385–2389.
23. Udvardy, A., Sümegi, J., Csordàs Tóth, E., Gausz, J., Gyurkovics, H., Schedl, P. & Ish-Horowicz, D. (1981) *J. Mol. Biol.*, in press.
24. Lis, J. T., Neckameyer, W., Mirault, M.-E., Artavanis-Tsakonas, S., Lall, P., Martin, G. & Schedl, P. (1980) *Dev. Biol.* **83**, 291–300.
25. Leigh Brown, A. J. & Ish-Horowicz, D. (1981) *Nature (London)* **290**, 677–682.
26. Finnegan, D. J., Rubin, G. M., Young, M. W. & Hogness, D. S. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **42**, 1053–1063.
27. Dunsmuir, P., Borein, W. J., Jr., Simon, M. A. & Rubin, G. M. (1980) *Cell* **21**, 575–579.
28. Levis, R., Dunsmuir, P. & Rubin, G. M. (1980) *Cell* **21**, 581–588.
29. Young, M. W. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 6274–6278.
30. Rubin, G. M., Brorein, W. J., Jr., Dunsmuir, P., Flavell, A. J., Levis, R., Strobel, E., Toole, J. J. & Young, E. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **45**, 619–628.
31. Wensink, P. C., Tabata, S. & Pachl, C. (1979) *Cell* **18**, 1231–1246.
32. Chaleff, D. T. & Fink, G. R. (1980) *Cell* **21**, 227–237.
33. Roeder, G. S. & Fink, G. R. (1980) *Cell* **21**, 239–249.
34. Green, M. M. (1967) *Genetics* **56**, 467–482.
35. Green, M. M. (1969) *Genetics* **61**, 429–441.
36. Gehring, W. J. & Paro, R. (1980) *Cell* **19**, 897–904.