The EMBO Journal vol.4 no.11 pp.2949–2954, 1985

Locus 67B of *Drosophila melanogaster* contains seven, not four, closely related heat shock genes

Agnes Ayme[1] and Alfred Tissières

Department of Molecular Biology, University of Geneva, Science II, 30 quai E. Ansermet, 1211 Geneva 4, Switzerland

[1]Present address: Department of Biology, Massachussetts Institute of Technology, Cambridge, MA 02139, USA

Communicated by A.Tissières

The four small hsp genes of *Drosophila melanogaster* as well as three genes regulated during development (genes 1, 2 and 3) are localized at the chromosomal locus 67B. The four small hsp genes share strong sequence homologies between themselves which were detected here by cross-hybridization. Under the same stringency conditions, each of the genes 1, 2 and 3 hybridize to some of the small hsp genes. By DNA sequencing of gene 1, the homology was localized within the same two regions already conserved between the small hsp genes: a central region of 83 amino acids, homologous with the mammalian α crystallin and the first 15 N-terminal amino acids. The transcriptional inducibility of the genes 1, 2 and 3 was also compared with that of the four small hsp genes during various stages of *Drosophila* development at either the normal growth temperature or after a heat shock. We confirm previous reports on the developmental patterns of all seven genes and find moreover that genes 1, 2 and 3 are heat-shock inducible at any of the stages tested. We conclude that genes 1, 2 and 3 are also heat shock genes. Therefore, the locus 67B contains seven, not four, small heat shock genes.

Key words: development/*Drosophila*/heat shock

Introduction

In *D. melanogaster*, seven genes have been identified clustered within a 15 kilobase (kb) region at the cytological locus 67B. Four of them are the small heat shock protein (hsp) genes hsp22, 23, 26 and 27 (Petersen *et al.*, 1979; Craig and McCarthy, 1980; Corces *et al.*, 1980; Wadsworth *et al.*, 1980; Voellmy *et al.*, 1981). These four genes are also transcribed during various specific developmental stages in the absence of heat shock (Sirotkin and Davidson, 1982; Zimmerman *et al.*, 1983; Mason *et al.*, 1984). The three remaining genes were identified by Sirotkin and Davidson (1982) who found them to be regulated during the development of the fly. They provisionally labelled them genes 1, 4 and 5, genes 2 and 3 corresponding to the hsp26 and 22 genes, respectively. Here, as previously (Southgate *et al.*, 1985) we call these three genes 1, 2 and 3.

Incubation of either Schneider 3 *Drosophila* tissue culture cells or isolated imaginal discs with the insect steroid hormone, ecdysterone, induces the abundant transcription of the four small hsp genes. Under the same conditions, however, gene 1 was weakly induced whereas the gene 2 and 3 transcripts were undetectable (Ireland and Berger, 1982; Ireland *et al* , 1982).

Comparison of the predicted amino acid sequences of the four small hsps has revealed the existence of a highly conserved region of 108 residues, shared among all four proteins. The first 83 amino acids of this stretch are also homologous to the mammalian α crystalline B2 chain (Ingolia and Craig, 1982; Southgate *et al.*, 1983). This conserved region as well as the clustering of the four small hsp genes suggests they have a common origin, having arisen from an ancestral gene by duplication and inversion events.

The possible relationships between the four small hsp genes and the three other closely linked developmentally-regulated genes remained, however, at that point an open question. Here, we have first investigated whether genes 1, 2 and 3 share homologous sequences amongst themselves and with any of the four small hsp genes. Both by hybridization experiments and DNA sequencing of gene 1 we find that, to a variable extent, this seems to be the case. Second, we tested the heat shock inducibility of genes 1, 2 and 3 and found that heat shock increases the amount of their mRNAs by 5- to 10-fold depending on the developmental stage which was investigated. It now appears likely that all seven genes at locus 67B are structurally interrelated, being expressed in an independent fashion during both conditions of heat shock stress and normal development.

Results

Cross-hybridization analysis of the small hsp and developmental genes

The organization of the seven genes clustered at the locus 67B is depicted in Figure 1. Plasmids containing the sequences for only one gene were cleaved with restriction endonucleases, the digests run on agarose gels and transferred to nitrocellulose. Single-stranded DNA probes were prepared from M13 clones containing all or the major part of the transcribed region for each of the seven genes. The filters were hybridized to the probes under conditions of reduced stringency (6 × SSC and 40% formamide at 42°C) before being washed sequentially at 65°C, from 2 × SSC to 0.1 × SSC (see Materials and methods for detail). Typical results obtained from hybridization with probes for either hsp22, hsp23 or gene 1 are shown in Figure 2. Similar cross-hybridizations were also obtained with the probes for hsp26, 27, gene 2 and gene 3 (data not shown). The following conclusions about genes' relatedness can be drawn: (i) the hsp 22 probe cross-hybridized weakly with hsp23, 26, 27 and gene 1. The very dark upper band in the hsp22 lane represents the normal hybridization between the M13 probe and pUC vector used in this hsp22 subclone. (ii) The hsp23 probe cross-hybridized to the same extent with hsp26, 27 and gene 1 but not at all to the hsp22 gene. (iii) The hsp26 probe only cross-hybridized with hsp23 and 27, and even with reduced washing stringency, no cross-hybridization with hsp22 and gene 1 was detected. (iv) The hsp27 probe hybridized to the same level with hsp23 and 26 as well as gene 1. (v) The gene 1 probe hybridized very strongly to hsp27, very weakly to hsp23 and no hybridization was detected with hsp22 and hsp26. (vi) The gene 2 probe weakly hybridized with hsp22 and 23, but not with either hsp26, 27 or gene 1. (vii) The gene 3 probe hybridized weakly to the gene 1 and even more weakly to hsp22 and hybridization to the other genes was not observed.

© IRL Press Limited, Oxford, England.

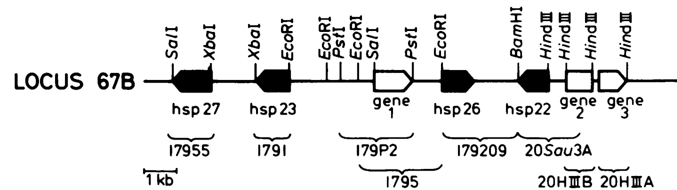

Fig. 1. Organization of the locus 67B with the location of the seven genes. The direction of transcription is indicated by the arrows. The four small hsp genes are shown by black arrows. The open arrows represent the three other genes called here genes 1, 2 and 3 (genes 1, 4 and 5 according to Sirotkin and Davison, 1982). The plasmids used in the cross-hybridization analysis are shown below the map with the restriction sites corresponding to the inserted fragments.

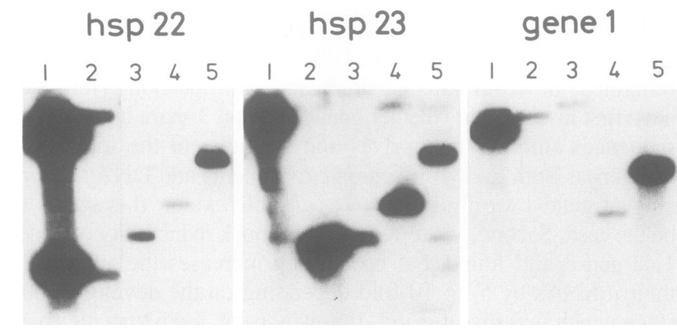

Fig. 2. Typical results of cross-hybridization analysis performed with probes for hsp22, hsp23 and gene 1. The lanes of the three identical filters correspond to digests of different plasmids each specific for one gene: **1)** 20Sau3A for hsp22; **2)** 1791 for hsp23; **3)** 179209 for hsp26; **4)** 17955 for hsp27 and **5)** 1795 for gene 1. The cross-hybridization obtained after washing the filters at the most stringent condition is shown (see Materials and methods). The upper bands represent the vector-containing fragments. The lower bands correspond to the insert containing the coding region of the different genes tested.

Table I. Homology between the small hsp, gene 1 and the α crystalline

Genes	26	23	22	gene 1	α crystalline
27	66 (66)	68 (65)	44 (54)	69 (73)	48
26	–	75 (69)	46 (58)	55 (60)	53
23	–	–	50 (58)	55 (63)	49
22	–	–	–	43 (58)	37
gene 1	–	–	–	–	47

Table I gives the percentage of homology between the four small hsp genes and gene 1 both at the amino acid and, in parentheses, at the nucleotide levels. The comparison is made from residues 59 to 141 in hsp 22, 66 to 148 in hsp23, 84 to 166 in hsp26, 85 to 167 in hsp27, 122 to 204 in gene 1 and 70 to 152 in α crystalline B2 chain. Data for the four small hsp genes are taken from Southgate *et al.* (1983) and for mammalian α crystalline from van der Ouderaa *et al.* (1973).

The extent of homology calculated at the nucleotide level, among the four small hsp genes (Ingolia and Craig, 1982; Southgate *et al.*, 1983; see also Table I) correlates well with this cross-hybridization study. Our results, therefore, probably reflect a true homology between the four small hsp genes and at least the gene 1. This was confirmed for gene 1, by DNA sequencing. The evidence for the presence of homologous regions in the genes 2 and 3 is, however, weaker.

Sequence analysis of gene 1

For the gene 1 sequence, we have determined ~2600 nucleotides of 5′-flanking, transcribed and 3′-downstream sequences (see Figure 3). Coordinate +1 is taken as the A residue of the ATG initiation codon.

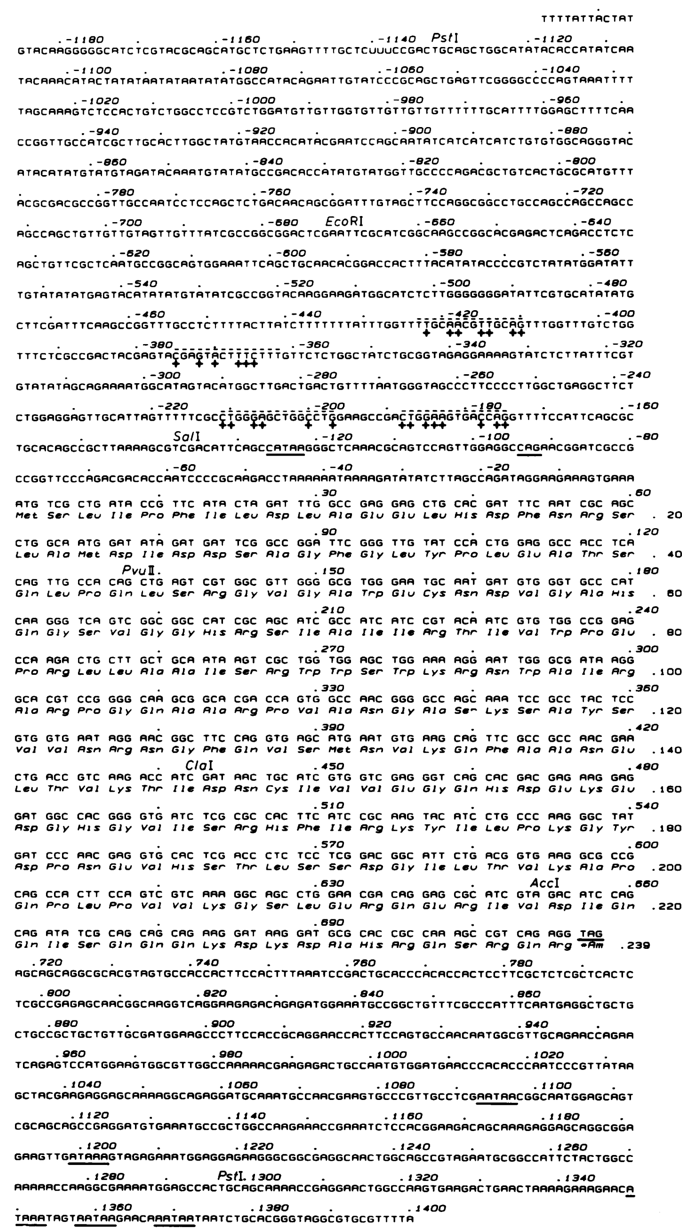

Fig. 3. Complete DNA sequence of gene 1. The coordinate +1 corresponds to the A residue of the ATG initiation codon. The 'CATAA', cap site, termination codon and 'AATAAA' sequences are underlined. The four 'Pelham' sequences are marked by dashed lines and the '+' beside them corresponds to the nucleotides matching with the heat shock consensus sequence. The major restriction sites are also given.

5′ sequences. The start of transcription of gene 1 has been reported by Sirotkin (1982) to lie near the *Sal*I site as shown in Figure 1. We have mapped the precise location of the cap site by primer extension analysis (see Materials and methods, data not shown). We found that the cap site, 5′ CAG 3′ which is comparable with the other *Drosophila* hsp gene cap sites (Hackett and Lis, 1983) is located 45 nucleotides to the right of the *Sal*I site (see Figure 3). The sequence 5′ C A T A A 3′ found at 28 base pairs 5′ to the cap site is considered to represent the canonical 'TATA' box.

Several sequences, sharing homology with the 'Pelham con-

```
4A) hsp27     M - S I I P - L L H L A R E L D H
    hsp26     M - * L S T - * * S * V D * * Q E
    hsp23     M A N * P L - * * S * * D D * G
    hsp22     M A * L P M - F W R M * E * M A
    gene1     M - * L * * F I * D * * E * * H

4B) hsp27     V G K D G F Q V C I D V S Q F K P N E L T V K
    hsp26     * * * * * * * * * M * * A * * * * S * * N * *
    hsp23     I * * * * * * * * M * * * H * * * S * * V * *
    hsp22     * N * * * Y K L T L * * K D Y - - S * * K * *
    gene1     * N R N * * * * S M N * K * * A A * * * * * *
  α crys.     L E * * R * S * N L N * K H * S * E * * K * *

    hsp27     V V D N T V V - V E G K H E E R E D G H G M I
    hsp26     * * * D S I L - * * * * * * * * Q * D * * H *
    hsp23     * Q * * S * L - * * * N * * * * * * D * * F *
    hsp22     * L * E S * * L * * A * S * Q Q * A E Q * G Y
    gene1     T I * * S I * - * * * Q * D * K * * * * * V *
  α crys.     * L G D V I E - * H * * * * * * Q * E * * F *

    hsp27     - Q R H F V R K Y T L P K G F D P N E V V S T
    hsp26     - M * * * * * R * K V * D * Y K A E Q * * * Q
    hsp23     - T * * * * * R * A * * P * Y E A D K * A * *
    hsp22     S S * * * L G R * V * * D * Y E A D K * S * S
    gene1     - S * * * I * * * I * * * * Y * * * * * H * *
  α crys.     - S * E * H * * * R I * A D V * * L A I T * S

    hsp27     V S S D G V L T L K A P P P P S D E Q A K S E
    hsp26     L * * * * * * * V S I * K * Q A V * D K S K *
    hsp23     L * * * * * * * I * V * K * * A I * D K G N *
    hsp22     L * D * * * * * I S V * N * * G V Q E T L K *
    gene1     L * * * * I * * V * * * Q * L P V V K G S L *
  α crys.     L * * * * * * * V N G * R K Q A

    hsp27     R I V Q I Q Q T G - P A H L S V K A P
    hsp26     * * I * * * * V * - * * * * N * * * N
    hsp23     * * * * * * * V * - * * * * N * * E N
    hsp22     * E * T * E * * * E * * K K * A E E *
    gene1     * Q E R * V D I Q Q I S Q Q Q K D K D
```

Fig. 4. The two conserved regions in the protein coding sequence of gene 1 as described in the text. The hsp27 amino acid sequence is taken as reference. Asterisks represent any amino acid residue identical to the hsp27 sequence in the hsp22, hsp23, hsp26, gene 1 and α crystalline B2 chain sequences. Some gaps have been introduced for better alignment of the sequences and are represented as [–]. **A)** first 15 N-terminal residues, **B)** central 108 amino acids region, from amino acids 59 to 167 in hsp22, 66 to 173 in hsp23, 84 to 191 in hsp26, 85 to 192 in hsp27, 122 to 229 in gene 1 and 70 to 152 in α crystalline B2 chain.

sensus sequence' 5′ C T n G A A n n T T C n A G 3′, which is involved in the induction of heat shock genes (Pelham, 1982; Pelham and Bienz, 1982; Mirault *et al.*, 1982; Pelham and Lewis, 1983; Ayme *et al.*, 1985) were found upstream of the 'TATA' box. Two proximal homologies are present in the DNA stretch extending between 82 and 117 bp upstream from the cap site, they share respectively an 8 out of 10 bp match to the consensus (position −188 to −175) and a 6 out of 10 bp match (position −210 to −197, see Figure 3). The same region (from −236 to −182) also contains four direct repeats of the sequence 5′ C T G G A Pu 3′. The two distal 'Pelham boxes' with 7 and 6 out of 10 bp match are located further upstream between positions −268 and −329 from the cap site (see Figure 3).

The first ATG triplet found downstream of the cap site is followed by an open, uninterrupted reading frame of 714 nucleotides encoding 238 amino acids. We assume that this ATG is the translation initiator codon (Baralle and Brownlee, 1978; Kozak, 1978, 1984). The leader sequence with 93 nucleotides is smaller than in the small hsp genes which range from ~150 to 250 nucleotides (Ingolia and Craig, 1982; Southgate *et al.*, 1983) but longer than the average eukaryotic leader sequence (Kozak, 1984). The leader sequence is 53% A+T rich with 39% of A residues, which differs from that reported for the small hsp genes (Ingolia and Craig, 1982; Southgate *et al.*, 1983).

Protein coding region. The 238 amino acids long open reading frame downstream from the ATG initiator codon is blocked in the two other reading frames by multiple termination codons.

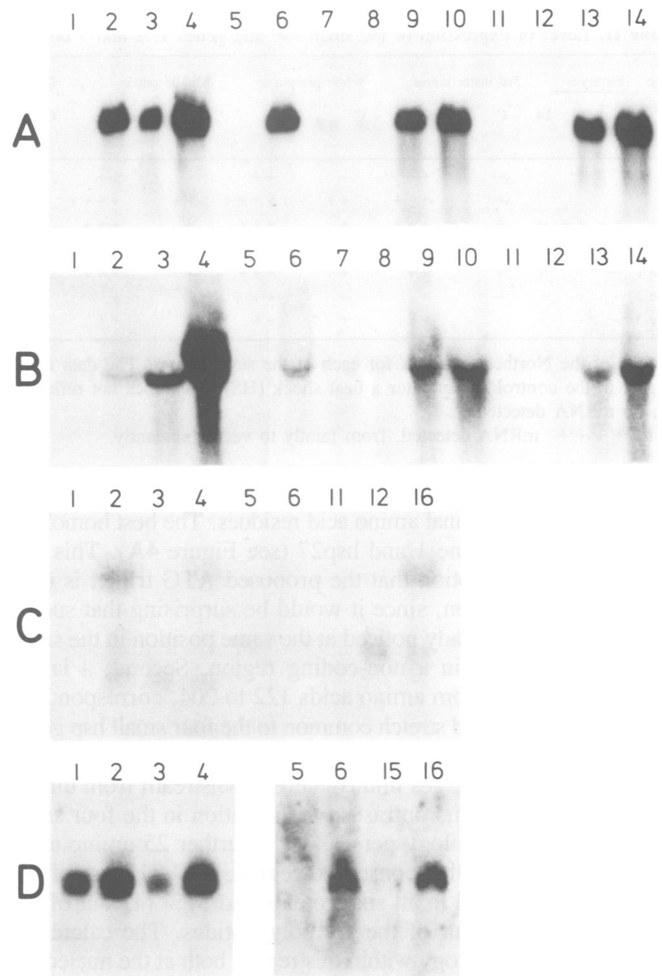

Fig. 5. Level of transcripts from the hsp23 gene as well as from the genes 1, 2 and 3 analysed by Northern blotting. Each lane contains 20 μg of total nucleic acids extracted without (C) or after a heat shock (HS). The numbers above the filters correspond to the following stages: **1)** larvae C, **2)** larvae HS, **3)** pre-pupae C, **4)** pre-pupae HS, **5)** late pupae C, **6)** late pupae HS, **7)** young females C, **8)** young males C, **9)** young females HS, **10)** young males HS, **11)** 3-day-old females C, **12)** 3-day-old males C, **13)** 3-day-old females HS, **14)** 3-day-old males HS, **15)** adults C, **16)** adults HS. **A)** hsp23 transcript, **B)** gene 1 transcript: the first four lanes of the filter corresponded to a 3-day exposure and the remainder to a 10-day exposure, both with intensifying screen, **C)** gene 2 transcript: the filter was exposed for 1 week with intensifying screen. A longer exposure did not reveal any expression in the pupae or the various adult stages. The origin of the two smaller RNAs detected by the gene 2 probe is not known. **D)** Gene 3 transcript: the early and middle pupae stages corresponded to a 5-day exposure and the late and adults to a 10-day exposure, both with intensifying screen.

The derived polypeptide product from gene 1 would be of 26 560 daltons. The coding region of gene 1 shows the same G+C richness previously observed in the small hsp and other *Drosophila* genes (Southgate *et al.*, 1983; Benyajati *et al.*, 1981; Sanchez *et al.*, 1983), due to the preferential use of G and C as the third codon base (data not shown).

A comparison between the coding regions of gene 1 and the four small hsp genes was performed by computer analysis and revealed, both at the nucleotide and amino acid levels, the same two regions of extensive homology already found among the four small hsp genes (Ingolia and Craig, 1982; Southgate *et al.*, 1983). Figure 4 presents the alignment in these two regions, at the amino acid level, for the four small hsp genes, gene 1 and the α-

Table II. Level of expression of the small hsp and genes 1, 2 and 3 during various developmental stages after heat shock

Gene	Embryos			3rd instar larvae		White pre-pupae		Middle pupae		Late pupae		Freshly eclosed				3-day-old adults			
												Females		Males		Females		Males	
	2	3	24	C	HS	C	HS	C	HS	C	HS	C	HS	C	HS	C	HS	C	HS
22				−	++	−	++	−	++	−	++	−	++	−	++	−	++	−	++
23				+	+++	++	++++	++	++++	−	+++	+	+++	+	+++	−	+++	−	+++
26	++	++	−	+	+++	+	+++	+	+++	−	+++	−	+++	−	+++	++	+++	−	+++
27	+	+	−	−	++	+	+++	+	+++	−	+++	−	+++	−	+++	++	+++	−	+++
gene 1	−	−	−	+	++	++	+++	++	+++	−	+	+	++	+	++	−	++	−	++
gene 2	−	+	+	−	+	−	+	−	+	−	+		+		+	−	+	−	+
gene 3	−	−	−	−	+	+	++	+	++	−	+	−	+	−	+	−	+	−	+

Results of the Northern analysis for each of the seven genes. The data represented the variation in the amount of each transcript present in the different tested stages, in the control (C) or after a heat shock (HS). This does not reflect the relative abundance of the transcripts from the different genes.
−: no mRNA detected.
+ to ++++: mRNA detected, from faintly to very abundantly.

crystalline B2 chain. First, a small region of homology resides in the first 15 N-terminal amino acid residues. The best homology is found between gene 1 and hsp27 (see Figure 4A). This finding supports the notion that the proposed ATG triplet is used as the initiation codon, since it would be surprising that such a good homology, already noticed at the same position in the small hsp genes, persists in a non-coding region. Second, a larger region, stretching from amino acids 122 to 204, corresponds to the first 83 amino acid stretch common to the four small hsp genes and the mammalian α-crystalline B2 chain (see Figure 4B). This homology, then, diverges immediately downstream from this 83 amino acid stretch, in contrast to the situation in the four small hsps where the homology persists for a further 25 amino acids (see Figure 4B). In this conserved region, 22% (18 out of 83) residues are identical in all six proteins and 51% (42 out of 83) are shared in five out of the six polypeptides. The calculated percentages of homology within this region both at the nucleotide and amino acid levels revealed a hierarchy in the relatedness between gene 1 and the small hsp genes (see Table I). Gene 1 is highly related to hsp27 with an homology of 69% at the amino acid level and 73% at the nucleotide level against only 55% at the amino acid level with hsp23 and 26.

Apart from these two homologous regions, the amino acid sequence of gene 1 product shares no homology with the small hsps. These non-homologous regions account for the length difference between the five polypeptides.

3′ Trailer and flanking sequences. The termination codon, 5′TAG is followed by several in-phase nonsense codons, the first occurring 16 bp after the former (see Figure 3). The hexanucleotide consensus sequence 5′ AATAAA 3′ is a characteristic feature of the 3′ ends of the eukaryotic mRNAs, located between 11 and 30 bp upstream of the poly(A) addition site (Berget, 1984). Several such sequences are observed downstream from the termination codon in the 3′ sequence of gene 1 (see Figure 3).

Level of transcripts from the locus 67B with or without heat shock at various developmental stages

The presence and relative amount of the transcripts from the four small hsp genes and genes 1, 2 and 3 were analysed throughout the *Drosophila* life cycle either when maintained under normal growth temperature at 20°C (C) or after a heat shock of 1 h at 37°C (HS). Total RNA was extracted from the following *Drosophila* developmental stages: embryos, third instar larvae, white pre-pupae, middle and late pupae, freshly eclosed females and males, 3-day-old females and males. These RNA preparations were then tested by Northern analysis using single-stranded DNA probes. The probes were each unique for hsp22, 23, 26, and 27 as well as for gene 1 transcripts since they do not contain any sequence from the central homologous part of their coding regions which could cross-hybridize (see Materials and methods). The gene 2 and 3 probes, however, contain the entire transcribed regions because their possible cross-hybridizing sequences have not been localized. Cytoplasmic actin mRNA was used as an internal standard, assuming that its steady-state level is constant during a 1 h heat shock treatment (data not shown). Typical Northerns are shown in Figure 5 and the data summarized in Table II.

We first confirm the developmental pattern found for the hsp22, 23, 26 and 27 genes which are normally transcribed from the third instar larval to the middle pupal stages, with the maximum attained in white pre-pupae (Sirotkin and Davidson, 1982; Mason *et al.*, 1984). The level of expression is, however, rather different for the four genes. Hsp23 transcript is the most abundant by 2- to 5-fold as compared with hsp26 and 27 mRNAs. Hsp22 is only weakly detectable. All four gene transcripts are no longer detectable in late pupae. Hsp23 expression is clearly detectable in young adults (Mason *et al.*, 1984), hsp26 and 27 genes are again expressed in 3-day-old females (Zimmerman *et al.*, 1983). Genes 1 and 3 are normally expressed in the same larval and pupal stages as the small hsp genes (Sirotkin and Davidson, 1982) and they are also no longer detected in late pupae. We also found that gene 1 is transcribed in young adults (females and males). Gene 2 is present in 3–24-h embryos, but not during pupation (Sirotkin and Davidson, 1982).

If a hyperthermic shock is applied to animals from the larval to the middle pupal stages, the transcription from the small hsp genes increases significantly over their basal developmentally-controlled level. The hsp 23 gene which gives the higher level of 'developmentally-regulated' transcripts is, comparatively, less 'heat-shock' induced than the hsp22, 26 and 27 genes. The small hsp genes accumulated approximately to the same level after a heat shock in the late pupal and adult stages.

A heat shock, performed at any of the tested stages, induced the transcription from genes 1 and 3. Differences have been observed, however, in the level of accumulated transcripts. In the white pre-pupal and middle pupal stages, the amount of their mRNAs increases significantly over their basal developmental level. During stages when genes 1 and 3 are not expressed or only weakly (third instar larvae, late pupae and adults), the amount of their mRNAs present is at least 10 times lower as compared with white pre-pupal and middle pupal accumulation. This is in contrast with the heat shock expression of the small hsp genes which gives approximately the same level of mRNAs whenever the tested stage. Gene 2 which is not normally detected in

the pupal stages becomes moderately transcribed after heat shock. It is heat-shock-induced at the same level in adult flies.

Discussion

An homologous sequence shared among the four small hsp genes was found, by DNA sequencing, in their coding regions (Ingolia and Craig, 1982; Southgate *et al.*, 1983). The question arose as to whether genes 1, 2 and 3 share common sequences amongst themselves and with the four small hsp genes. Here, the homology among the four small hsp genes was first used to determine conditions allowing cross-hybridization. A specific pattern of hybridization was found which correlates well with the homology calculated from the DNA sequence data (Ingolia and Craig, 1982; Southgate *et al.*, 1983). The same type of analysis was then used to test genes 1, 2 and 3. Gene 1 clearly cross-hybridized to hsp27 and to a lesser extent to hsp23. Genes 2 and 3 share a weak homology with hsp22, gene 3 also weakly cross-hybridizes to gene 1. These results suggest the possible existence in these three genes of regions homologous with the small hsp genes.

From the gene 1 sequence reported here a single uninterrupted open reading frame was found. The derived amino acid sequence corresponded to a polypeptide of 26 560 daltons before any post-translational modifications. Hsp27 has a calculated mol. wt. of 23 620 daltons (Southgate *et al.*, 1983) and by comparison the gene 1 polypeptide would appear on SDS-polyacrylamide gels as a protein ~30 kilodaltons. The comparison of this protein sequence with that of the small heat shock proteins further confirmed the localization and the extent of the homology found by the cross-hybridization analysis. Thus, the 'gene 1 protein' contains the same two conserved regions already noted in the small heat shock proteins: the first 15 N-terminal residues and a central 83 amino acids stretch (see Figure 4). In the small hsps, the central homologous region extends to 108 amino acids of which the first 83 residues are common with the mammalian α-crystalline B2 chain. The remaining 25 amino acids, towards the C-terminal end, are conserved only amongst the four small heat shock proteins (Ingolia and Craig, 1982; Southgate *et al.*, 1983). Gene 1 shares only and specifically the first 83 residues, homologous to the mammalian α-crystalline. Therefore, the small hsp and gene 1 have the same global organization of their coding region. We hypothesize that genes 2 and 3 cross-hybridize to the small hsp genes through the same conserved central domain and therefore have a similar sequence organization. If the conserved domains are, as already proposed, the regions involved in the proteins' function(s) (Ingolia and Craig, 1982; Southgate *et al.*, 1983), we can hypothesize that the polypeptides encoded by genes 1, 2 and 3 may have functions similar to the small heat shock proteins. However, the specific conservation of the last 25 amino acids in the central homologous region of the four small hsps should not be neglected and may confer slightly different properties to the small heat shock proteins.

The degree of homology in the conserved central region, calculated both at the nucleotide and amino acid levels (see Table I), shows that gene 1 is more closely related to hsp27 than to any other small hsp genes. This correlates well with the relationships found between gene 1 and the small hsp genes by the cross-hybridization analysis. The seven genes are probably all derived from an ancestral gene by several duplication and inversion events within the locus. If this hypothesis is correct, then gene 1 must have arisen rather directly from the hsp27 gene or *vice versa*. The specific pattern of preferential cross-hybridization and the calculated degree of homology may, therefore, reflect the order in which the supposed duplication events have occurred.

The expression of the seven genes was investigated during various stages of the *Drosophila* development, both before and after a heat shock. Hsp22, 23, 26 and 27 as well as gene 1 and 3 are normally expressed from the late third instar larval to the middle pupal stages (Sirotkin and Davidson, 1982; Mason *et al.*, 1984). Their levels of expression vary by a factor of 10 with hsp23 being the most expressed and hsp22 the least. We found that gene 1 is also transiently expressed in freshly eclosed adults as it was described for hsp23 (Mason *et al.*, 1984). Gene 2 is specifically expressed in the 3–24-h embryos (Sirotkin and Davidson, 1982). A heat shock at the third instar larval, early and middle pupal stages, increases the amount of mRNAs for the six genes already expressed (the small hsp genes and genes 1 and 3) by a factor of 5 to 25 over their basal expression at these stages. It could be argued that the increase in the detected amount of transcripts occurred by a stabilization of the pre-existing mRNAs. However, the notion of heat shock inducibility for genes 1, 2 and 3 is supported by the following results: genes 1 and 3 are also heat-shock inducible during late pupal and 3-day-old adult fly stages when they are not developmentally expressed and the transcription of gene 2 is induced by a heat shock in all the tested stages.

However, the accumulation of transcripts from genes 1 and 3 after a heat shock is different, whether the genes are already developmentally expressed or not. In the pre-pupal and middle pupal stages, the transcripts from genes 1 and 3 accumulated after heat shock to a level which is 10 times higher than in the late pupae and adults. Therefore, genes 1 and 3 have probably a lower heat shock inducibility than the four small hsp genes.

The 5′ sequences of the hsp83, 70, 68, 26 and 22 genes contain multiple matches to the 'Pelham consensus sequence' which was demonstrated to be responsible for the heat shock induction of the hsp70, 26 and 22 genes (Mirault *et al.*, 1982; Pelham, 1982; Pelham and Lewis, 1983; Ayme *et al.*, 1985). The sequences requirement for optimal induction of the hsp70 gene was investigated by the P-mediated transformation technique. These analyses have shown that two very close 'Pelham consensus' sequences are necessary for the maximal heat shock induction of the reintroduced hsp70 gene (Dudler and Travers, 1984; Simon *et al.*, 1985). Preliminary data for promoter deletion analysis of the hsp26 gene by the same approach give the same type of results (D.Pauli, personal communication). The 5′ region of gene 1 reveals the presence of only two significant (7 or more out of 10 bp match) 'Pelham' boxes located at 230 bp from each other. This relative lack of consensus sequences may explain the low level of heat shock inducibility observed for gene 1. When gene 1 is already expressed, however, the immediate presence of RNA polymerase II and of some transcription factors as well as the modified 'active' chromatin structure may enhance its transcriptional inducibility by heat shock.

The total amount of gene 1 mRNA present after a heat shock in the third instar larval, early and middle pupal stages may also result from a combination of the two processes of transcriptional induction of the gene and increased stability of the pre-existing messengers. The same situation has already been described by Vitek and Berger (1984) in ecdysterone-treated Schneider 3 cells. The hsp22 and 23 genes are transcribed in these cells after the hormone treatment (Ireland and Berger, 1982). If a heat shock is applied to treated cells the amount of mRNAs increased by a combination of induction of the gene's transcription and stabilization of the pre-existing messengers. The hsp transcripts were found to be 2–3 times more stable at 35°C than at 25°C

(Vitek and Berger, 1984).

The relationship between the normal developmental expression and the induction by heat shock remains to be elucidated. The promoter sequences involved in the developmental expression of the seven genes have not yet been identified and the direct implication of the 'Pelham consensus' sequence cannot be excluded. In that case, it remains to be determined whether the same factors are involved in the developmental expression and the heat shock induction. This will probably help in understanding the variation in the heat shock inducibility of genes 1 and 3.

Materials and methods

Cross-hybridization analysis

The different clones used, each specific for only one gene, are listed below: 20Sau3A for hsp22, 1791 for hsp23, 179209 for hsp26, 17955 for hsp27 and 1795 for gene 1. They were previously described in Voellmy *et al.*, 1981; Southgate *et al.*, 1983 and Ayme *et al.*, 1985; see Figure 1. After restriction digests which separate specifically the insert from the vector, the fragments were separated on 1% agarose gels and transferred to nitrocellulose filters (Southern, 1975). The single-stranded DNA probes (see below) were prepared as previously described (Ayme *et al.*, 1985). Hybridization was carried out for 12 h at 42°C in 6 × SSC (20 × SSC is 3M NaCl, 0.3 M Na citrate), 0.1% SDS, 5 × Denhardt (100 × is 2% Ficoll, 2% polyvinylpyrrolidone and 2% BSA), 200 μg/ml denatured salmon sperm DNA and 40% deionised formamide. The washing of the filters was performed at 65°C with a stringency varying from 2 × SSC to 0.1 × SSC for several hours.

Subclone construction and DNA sequencing

The plasmid 179P2 containing gene 1 sequence was cloned from a total *Pst*I digest of the original genomic λ clone 179 (Voellmy *et al.*, 1981) into the pUC8 plasmid vector. The *Hind*III fragments of the 20HIII A and 20HIII B clones were isolated from the plasmid 17920 (Voellmy *et al.*, 1981) and cloned into pUC8 vector. They contain respectively genes 2 and 3 (see Figure 1).

Specific fragments of 179P2 and 1795 were isolated from polyacrylamide gels and subcloned into M13mp10 or mp11 replicative forms (obtained from Biochemicals Ltd). Restriction digests (*Rsa*I, *Alu*I, *Hae*III and *Sau*3A1) of the same plasmids were also randomly cloned into M13. After transformation into JM103, white plaques were screened by dot blotting (Messing and Vieira, 1982). Sequencing was performed by the dideoxynucleotide chain termination technique of Sanger *et al.* (1977, 1980). Reactions were carried out with [α-^{35}S]dATP as labelled deoxynucleotide (Radiochemical Centre, Amersham, England). The dideoxy nucleotides were obtained from Bethesda Research Lab., GmbH. The reactions were run on 6% polyacrylamide, 8.3 M urea, thin sequencing gels which were dried for exposure.

Primer extension analysis

The single-stranded probe *Cla*I-*Pvu*II, contianing the gene 1 sequence from coordinates +437 to +130, was prepared from an M13 clone containing the fragment *Cla*I-*Sal*I (see Figure 3). Hybridization to total pupal RNA and the reverse transcriptase reaction were performed as already described (Ayme *et al.*, 1985) and the extended fragment analysed on thin sequencing gels.

RNA extraction and Northern analysis

Embryos were collected at 0−3 h, 3−6 h and 24 h after egg laying. Third instar larvae were collected as actively climbing larvae. Animals were further staged at puparium formation: white pre-pupae between 0 and 1 h, mid-pupae between 12 h and 2 days, late pupae between 4 and 5 days. Adults were taken within 2 h after eclosion (young) and after 3−4 days. Animals were homogenised in 2−5 ml of 100 mM Tris pH 9, 100 mM NaCl, 20 mM EDTA, 1% Sarcosyl into a 7 ml Dounce homogenizer. Lysis was followed by three phenol extractions (50 v. phenol: 50 v. chloroform: 1 v. isoamyl alcohol). Total nucleic acids were precipitated with 3 v. ethanol overnight at −20°C and resuspended in water. RNA was separated on a 1.2% agarose − 2.2 M formaldehyde gel in 1 × MOPS buffer (10 × is 0.2 M MOPS, 50 mM Na acetate, 10 mM EDTA, pH 7) and directly transferred after electrophoresis on nitrocellulose filters using 20 × SSC. The small hsp and gene 1, 2 and 3 probes are described in the next section. The actin probe corresponded to nick-translation of fragments isolated from the plasmid 5C (Fryberg *et al.*, 1980). Hybridization was performed for 12 h at 42°C in Northern buffer (5 × SSPE, 0.1% Na pyrophosphate, 0.4% SDS, 50% formamide, 8% Dextran sulfate and 250 μg/ml denatured salmon sperm DNA). Filters were then washed at 65°C in 0.1 × SSC, 0.1% SDS for 2−3 h.

M13 clones

The single-stranded DNA probes corresponding to the following restriction fragments were used for (i) the cross-hybridization study and (ii) the Northern analysis. The coordinates for the four small hsp genes are taken from Southgate *et al.* (1983) and from Figure 3 for the gene 1 (+1 corresponds to the A residue of the initiation codon).

(i)	hsp22	*Bam*HI	*Hind*III	+850 to −40
	hsp23	*Xba*I	*Pvu*II	+874 to +135
	hsp26	*Cla*I	*Sac*I	+832 to +308
	hsp27	*Sal*I	*Xba*I	+1200 to −39
	dev1	*Acc*I	*Pvu*II	+645 to +130
	dev2	*Hind*III	*Hind*III	
	dev3	*Hind*III	*Hind*III	
(ii)	hsp22	*Hind*III	*Hinc*II	−40 to −208
	hsp23	*Pvu*II	*Eco*RI	+135 to −108
	hsp26	*Sac*I	*Xba*I	+308 to −234
	hsp27	*Sma*I	*Hinc*II	+31 to −135
	dev1	*Pvu*II	*Sal*I	+130 to −140
	dev2	*Hind*III	*Hind*III	
	dev3	*Hind*III	*Hind*III	

Acknowledgements

We thank Mme C.H.Tonka for her valuable help during DNA sequencing; O.Jenni, Y.Epprecht and F.Veuthey for preparing the figures and D.Pauli, L.Hall, R.Southgate and P.Spierer for their critical reading of the manuscript. This work was supported by the Swiss National Science Foundation.

References

Ayme,A., Southgate,R. and Tissières,A. (1985) *J. Mol. Biol.*, **182**, 469-475.
Baralle,F.E. and Brownlee,G.G. (1978) *Nature*, **247**, 84-87.
Benyajati,C., Place,A.R., Powers,D.A. and Sofer,W. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2717-2721.
Berget,S.M. (1984) *Nature*, **309**, 179-182.
Corces,V., Holmgren,R., Freund,R., Morimoto,R. and Meselson,M. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5390-5393.
Craig,E.A. and McCarthy,B.J. (1980) *Nucleic Acids Res.*, **8**, 4441-4457.
Dudler,R. and Travers,A.A. (1984) *Cell*, **38**, 391-398.
Fyrberg,E.A., Kindle,K.L. and Davidson,N. (1980) *Cell*, **19**, 365-378.
Hackett,R.W. and Lis,J.T. (1983) *Nucleic Acids Res.*, **11**, 7011-7030.
Ingolia,T.D. and Craig,E.A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2360-2364.
Ireland,R.C. and Berger,E. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 855-859.
Ireland,R.C., Berger,E., Sirotkin,K., Yund,M.A., Osterbur,D. and Fristrom,J. (1982) *Dev. Biol.*, **93**, 498-507.
Kozak,M. (1978) *Cell*, **15**, 1109-1123.
Kozak,M. (1984) *Nucleic Acids Res.*, **12**, 857-872.
Mason,P.J., Hall,L.M.C. and Gausz,J. (1984) *Mol. Gen. Genet.*, **194**, 73-78.
Messing,J. and Vieira,J. (1982) *Gene*, **19**, 269-276.
Mirault,M.-E., Southgate,R. and Delwart,E. (1982) *EMBO J.*, **1**, 1279-1285.
Pelham,H.R.B. (1982) *Cell*, **30**, 517-528.
Pelham,H.R.B. and Bienz,M. (1982) *EMBO J.*, **1**, 1473-1477.
Pelham,H.R.B. and Lewis,M. (1983) in Hamer,D. and Rosenber,M. (eds.), *Gene Expression: UCLA Symposia on Molecular and Cellular Biology*, Vol. **8**, pp. 75-86.
Petersen,N.S., Moeller,G. and Mitchell,H.K. (1979) *Genetics*, **92**, 891-902.
Sanchez,F., Tobin,S.L., Rdest,U., Zulauf,E. and McCarthy,B.J. (1983) *J. Mol. Biol.*, **163**, 533-551.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.
Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.
Simon,J.A., Sutton,C.A., Lobell,R.B., Glaser,R.L. and Lis,J.T. (1985) *Cell*, **40**, 805-817.
Sirotkin,K. (1982) in Schlesinger *et al.* (eds.), *Heat Shock from Bacteria to Man*, Cold Spring Harbor Publications, Cold Spring Harbor, NY.
Sirotkin,K. and Davidson,N. (1982) *Dev. Biol.*, **89**, 196-210.
Southern,K. and Davidson,N. (1982) *J. Mol. Biol.*, **98**, 503-517.
Southgate,R., Ayme,A. and Voellmy,R. (1983) *J. Mol. Biol.*, **165**, 35-57.
Southgate,R., Mirault,M.-E., Ayme,A. and Tissières,A. (1985) in Atkinson,B.G. and Walden,D.B. (eds.), *Changes in Eukaryotic Gene Expression in Response to Environmental Stress*, Academic Press, NY, pp. 1-30.
Van der Ouderaa,F.J., de Jong,W.W., Hilderink,A. and Bloemendal,H. (1973) *Eur. J. Biochem.*, **39**, 207-211.
Vitek,M.P. and Berger,E.M. (1984) *J. Mol. Biol.*, **178**, 173-189.
Voellmy,R., Goldschmidt-Clermont,M., Southgate,R., Tissières,A., Levis,R. and Gehring,W.J. (1981) *Cell*, **23**, 261-270.
Wadsworth,S., Craig,E.A. and McCarthy,B.J. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 2134-2137.
Zimmerman,J.L., Petri,W. and Meselson,M. (1983) *Cell*, **32**, 1161-1170.

Received 22 July 1985