Molecular characterization of the 5′ control region and of two lethal alleles affecting the *hsp60* gene in *Drosophila melanogaster*

L. Perezgasga[a], L. Segovia[b], M. Zurita[a,*]

[a] *Departamento de Genética y Fisiología Molecular, Instituto de Biotecnología, UNAM. Apdo. Postal 510-3, Cuernavaca, Morelos 62250, Mexico*
[b] *Departamento de Reconocimiento Molecular y Bioestructura, Instituto de Biotecnología, UNAM. Apdo. Postal 510-3, Cuernavaca, Morelos 62250, Mexico*

Received 11 June 1999

Abstract The chaperonins are evolutionarily conserved essential cellular proteins that help folding newly synthesized or translocated proteins, spending ATP. We present here the molecular analysis of the *hsp60* gene promoter region and of two *Drosophila hsp60* ethyl methane sulfonate embryonic lethal alleles that have an identical phenotype. No heat shock element sequences were found in the 5′ region, supporting previous data (Kozlova, T. et al., 1997) which suggests that mitochondrial Drosophila melanogaster HSP60.1 is not heat inducible. By sequencing the lethal allele's entire open reading frame (ORF), we found a C-T transition in the *hsp60*^{F409} allele that produces a serine to leucine change, apparently distorting the protein equatorial domain structure. No changes were found in the *hsp60*^{G93} ORF. However, an analysis of the heterogeneous nuclear RNA levels showed a reduction of the *hsp60* transcript in *hsp60*^{G93} flies as compared to the wild-type. These data suggest that although the defects in the *hsp60* gene produced by these alleles are at different levels, both behave as null mutations.

Key words: Ethyl methane sulfonate embryonic lethal allele; *hsp60* gene; Chaperonin; *Drosophila melanogaster*

1. Introduction

The chaperonins exist in two distinct evolutionary versions [1,2]. The first family includes the bacterial and organelle chaperonins (mitochondria and chloroplasts), 60 kDa heat shock proteins. The second family, referred to as t-complex polypeptide 1 (Tcp-1), includes the archaeal and eukaryotic chaperonins [3]. The chaperonins of the first family are cellular proteins composed of back to back rings (the *cis* and *trans* rings) of seven rotationally symmetric identical or closely related subunits (~60 kDa) which assist the folding of newly synthesized and newly translocated proteins and promote the refolding of denatured ones in an ATP-dependent reaction [4]. Each of the 14 subunits is formed by three domains: the apical domain, which interacts with the co-chaperonin, the intermediate or hinge domain and the equatorial domain which is a relatively fixed base housing the ATP binding site [5]. The typical representative of this group is *Escherichia coli* GroEL which forms a complex with GroES and is thought to promote productive folding in two major states. In the binding active state, the central cavity of each GroEL ring is open at the end of the cylinder for the ingress of the non-native polypeptides in the *cis* ring [6,7], exposing a hydrophobic lining that seems to bind non-native species through exposed hydrophobic surfaces [8,9]. In the active folding state, binding of the co-chaperonin GroES to the *cis* ring of GroEL in the presence of ATP provokes conformational changes, triggering the folding of the substrate protein in the central channel which is now hydrophilic in character, thus favoring the burial of hydrophobic surfaces in the substrate protein [5]. The hydrolysis of ATP in the *cis* ring weakens GroEL-GroES interaction, priming GroES release while polypeptide folding continues. Binding of ATP to the *trans* ring evicts GroES and the polypeptide from the *cis* ring ($t = \sim 15$ s). This negative allosteric mechanism suggests a model for an ATP-driven folding cycle that requires a double toroid [4,5,10].

In contrast to GroEL, the mammalian mitochondrial HSP60 exists as a single ring [11,12]. This HSP60 with its cognate co-chaperonin HSP10 shows productive folding in vivo due to the low affinity of $HSP60_7$:ADP_7 for HSP10, that eliminates the requirement for ATP binding to a *trans* ring in order to release the co-chaperonin [12].

In a recent work [13], we characterized the first *Drosophila hsp60* homologue. *Drosophila melanogaster* HSP60.1 (DmHSP60) is enriched in the mitochondrial fraction and is most closely related to the mouse mitochondrial HSP60 (72% identity across the entire protein sequence).

We analyze here the *hsp60* promoter region and determine the exact transcription start site, as well as the TATA box, a putative CAAT box and an Oct-1 site. Nevertheless, we do not find functional binding sites for heat shock factor, confirming our previous data [13] that neither the transcript nor the protein are heat inducible. We also present the molecular characterization of two ethyl methane sulfonate embryonic lethal alleles of the *hsp60* gene that have been described previously [14,15]. The molecular analysis of the *hsp60*^{F409} allele shows a C to T transition that produces an evolutionarily conserved serine change to leucine in amino acid 154. The modelling of this mutation suggests that it distorts the equatorial domain structure where the catalytic site is located and, by consequence, affects the folding of the protein substrate. In the case of the *hsp60*^{G93}, we do not find any change in the open reading frame (ORF). Nevertheless, in heterozygous *hsp60*^{G93} flies, we observe a reduction of the *hsp60* heterogeneous nuclear RNA (hnRNA) transcript in comparison with wild-type flies. This suggests that this mutation produces a defect in the transcription or early processing of the mRNA.

*Corresponding author. Fax: (527) (3) 17-23-88.
E-mail: marioz@ibt.unam.mx

Abbreviations: PCR, polymerase chain reaction; RT-PCR, reverse transcriptase polymerase chain reaction; hnRNA, heterogeneous nuclear RNA; ORF, open reading frame; DmHSP60, *Drosophila melanogaster* HSP60.1

PII: S0014-5793(99)00963-1

2. Materials and methods

2.1. Fly strains

The wild-type stock was OreR. The *hsp60*^{F409} and *hsp60*^{G93} alleles were maintained in the presence of the FM6 balancer [15,16]. Flies were reared on standard yeast dextrose medium at 25°C.

2.2. DNA manipulations and polymerase chain reaction (PCR) of the *hsp60*^{F409} and *hsp60*^{G93} alleles

Genomic DNA was independently purified from three different heterozygous females of both alleles and a wild-type female [17] and used as a template in PCR reactions with primers that flank the *hsp60* ORF. Amplification reactions were carried out in a Perkin Elmer 2400 thermal cycler with the Expand High Fidelity PCR System (Boehringer Mannheim), according to the manufacturer's protocol in a 50 μl volume at an annealing temperature of 58°C.

2.3. DNA sequencing

DNA sequencing reactions of the amplified fragment of both alleles or of the 15 kb *Sal*I genomic DNA fragment cloned in the CaSper AUG transformation vector [13] were performed using the Thermo Sequenase radiolabelled terminator cycle sequencing kit (Amersham Life Science) in a Perkin Elmer 2400 thermal cycler according to the supplier's instructions for 45 cycles with primers used to sequence the *hsp60* cDNA clones [13].

2.4. High resolution of the transcription start point

Primer extension analysis was done according to Sambrook [18]. End-labelled BP5-5′r1, complementary to nucleotides +102–+124 on the sequence shown in Fig. 1C, was used as primer in the extension reaction. The extension reactions were performed using total cellular RNA from 0–4 h fly embryos and the products were resolved on a 7% denaturing polyacrylamide gel using the corresponding sequencing ladder as a size marker. We used the same RNA treated with RNAse A before the extension reaction as a control (data not shown).

2.5. Computer modeling of the *hsp60*^{F409} allele and wild-type *Drosophila HSP60* proteins

The DmHSP60 wild-type and mutant sequences were modeled on top of the *E. coli* GroEL structure [5]. The resulting models were minimized with the cvff force field using the Discover program. Figures were produced using the MolMol visualization program and rendered with PovRay 3.0 (tm) [19,20].

2.6. hnRNA reverse transcriptase (RT-) PCR analysis of the *hsp60*^{G93} allele

hnRNA was purified from *hsp60*^{G93}/FM6 and wild-type females' nuclei and used to synthesize cDNA in a 20 μl reaction containing 1×first strand buffer (50 mM Tris-HCl, 75 mM KCl, 3 mM $MgCl_2$, pH 8.3), 10 mM DTT, 200 U of M-MLV reverse transcriptase (GIBCO-BRL), 200 μM dNTPs, 2 mM oligo dT_{15} and 20 U RNAse inhibitor (Amersham Life Science). The mix was initially incubated at 65°C for 5 min before adding the DTT and the enzyme. It was then further incubated at 38°C for 1 h. 3 μl was used for the PCR reaction in 1×Taq buffer (100 mM Tris-HCl, 1.5 mM $MgCl_2$, 50 mM KCl, pH 8.3), 200 μM dNTPs, 0.5 mM of each specific oligo and 1.5 U Taq polymerase (Boehringer Mannheim). All PCR reactions were done in a Perkin Elmer 2400 thermal cycler machine. Reactions were incubated at 95°C for 5 min, 58°C for 40 s before adding Taq polymerase and the reactions were then incubated at 72°C for 1 min for one cycle. Afterwards, samples were held at 94°C for 30 s, 58°C for 40 s and 72°C for 1 min, for 32 cycles. As an internal control, we amplified *rp49* transcript. Reaction products were resolved in a 2% agarose gel in Tris-borate running buffer. After transferring to a nylon membrane, the PCR products were hybridized against BP5-3 and *rp49* cDNA clones [18,13].

2.7. Protein analysis

hsp60^{F409}/FM6 and *hsp60*^{G93}/FM6 and wild-type females were gently homogenized with 20 strokes in a dounce homogenator in the presence of homogenization buffer [21]. The protein concentration of the whole extracts was normalized with the Bradford assay. Proteins were separated on a 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel [22] and then blotted onto a nitrocellulose membrane for immunostaining [23] with the ECL Western blotting protocols (Amersham Life Science) with a 1:1000 dilution of the anti-HSP60 antibody (StressGen), followed by anti-rabbit Ig horseradish peroxidase secondary antibody (Amersham Life Science).

3. Results

3.1. Genomic organization of the Dmhsp60 gene and characterization of the promoter region

From the analysis of the *hsp60* S15 allele, we found an intron of about 3 kb located just before the initiation codon [13]. By sequencing genomic DNA from *hsp60*^{F409}/FM6 and *hsp60*^{G93}/FM6 alleles with primers used to sequence cDNA clones, we found another small intron of 64 bp in the 3′ region of the ORF. These results show that the *hsp60* gene consists of three exons, the first contains the 5′ untranslated region (5′ UTR) separated from the second exon (which has most of the *hsp60* coding sequence) by a 3 kb intron and the last exon that contains the rest of the *hsp60* coding sequence, interrupted by the 64 bp intron, as well as the 3′ UTR (Fig. 1A).

In order to characterize the 5′ control region of the *Dmhsp60* gene, we used the 15 kb *Sal*I genomic DNA fragment cloned in the CaSpeR vector as a template to sequence the *hsp60* promoter region [13]. High resolution primer extension experiments were carried out to define the exact *hsp60* gene transcription start site. The RNA *hsp60* start site (Fig. 1B) was identified using a primer located at nucleotide +102, indicated in Fig. 1C, that produces an extension product of 102 nucleotides. The 5′ control region was then analyzed with the Findpatterns program of the GCG Wisconsin software and the TESS program [24]. We identified a TATA box at −37 bp, a putative CAAT box at −139 bp and a putative Oct-1 site at −225 bp from the transcription start site (Fig. 1C). We did not find any heat shock elements with either of these programs. This supports our previous data [13] that showed that the *hsp60* transcript nor the protein are heat inducible.

3.2. Molecular analysis of the *hsp60*^{F409} and *hsp60*^{G93} alleles

Homozygous *hsp60*^{F409} flies die very early during embryogenesis with no signs of gastrulation. This indicates that the product of the *hsp60* gene is essential for the early stages of fly embryogenesis [13] (Fig. 2). Unfortunately, this phenotype makes working with homozygous embryos very difficult, so all the analyses reported in this work were carried out with heterozygous adult females. We sequenced the entire ORF of three different heterozygous and wild-type females. We found a C to T transition that produces a serine to leucine change in amino acid 154 of the mutant DmHSP60 protein. Serine-154 is a highly conserved amino acid located between α-helices E and F of the reported GroEL structure [5]. We performed a homology modelling of the DmHSP60 wild-type and DmHSP60^{F409} putative structure proteins on the *E. coli* structure to determine if the serine to leucine change produces an important modification in the protein structure (Fig. 2A). In the model, we could see that the equatorial domain of the mutated protein (where the catalytic site is located) is distorted with respect to the wild-type DmHSP60. This distortion would increase the modelled AMP bonding distances with its nearest amino acids (Fig. 2B).

As in the *hsp60*^{F409} flies, *hsp60*^{G93} homozygous mutant animals die early in embryogenesis [15], embryos are amorphous

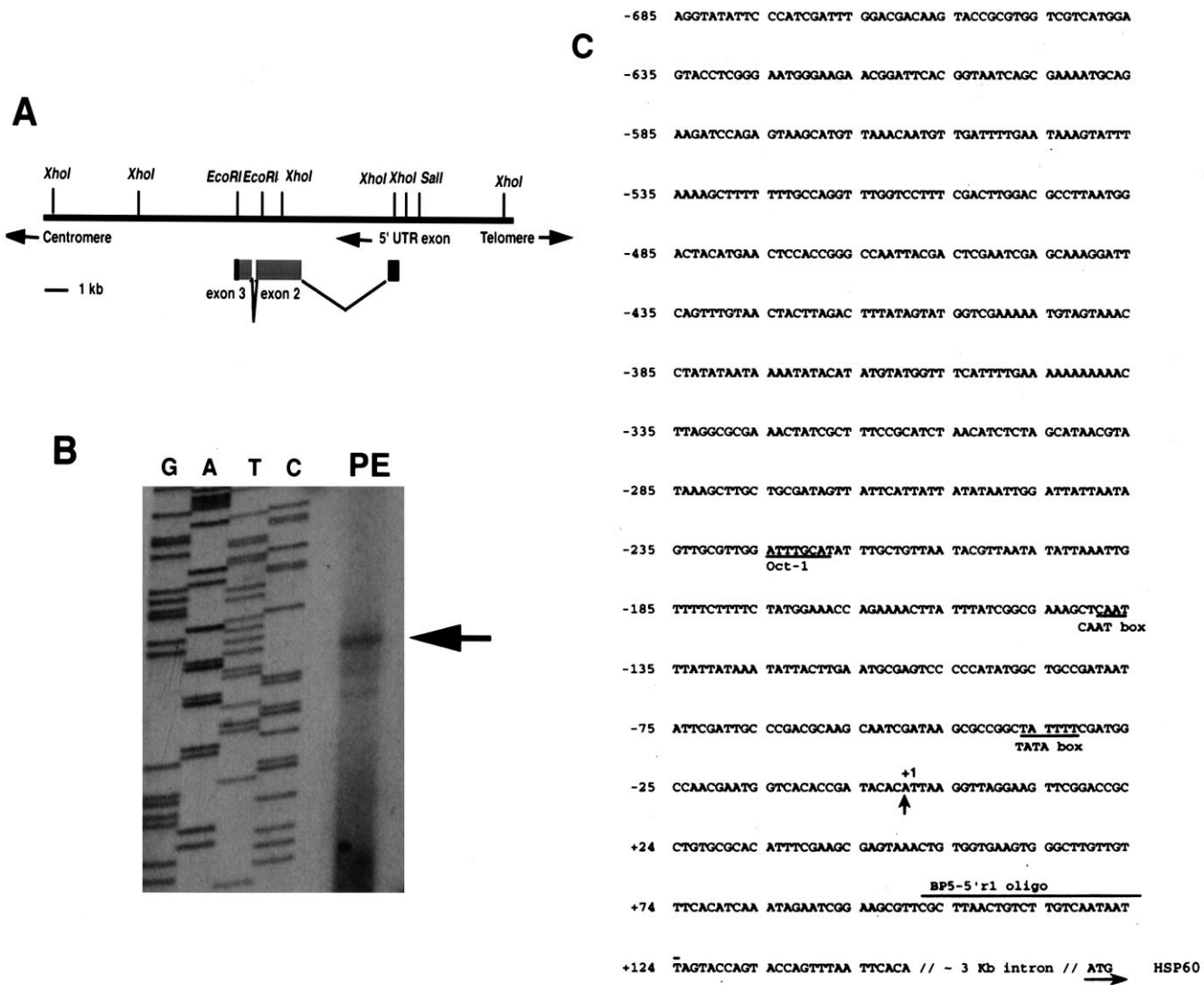

Fig. 1. Genomic organization of the *D. melanogaster hsp60* gene and characterization of its 5′ control region.(A) Oligonucleotide primers deduced from the sequence of the *hsp60* cDNA clones were used to sequence genomic DNA from *hsp60*FM6 females. Black boxes represent the 5′ and 3′ UTRs, while shaded areas correspond to the *hsp60* coding region. (B) Primer extension analysis. Total RNA from 0–4 h embryos was hybridized with ^{32}P-labelled oligonucleotide BP5-5′r1 and incubated with reverse transcriptase to produce cDNA extensions. The products were separated on a 7% polyacrylamide sequencing gel using a sequence reaction of the 15 kb *Sal*I genomic DNA fragment cloned in the CaSpeR AUG transformation vector [13], primed with the same oligonucleotide as the molecular mass marker. Lane PE shows the cDNA extension product. (C) Sequence of the *hsp60* promoter region. Using the 15 kb *Sal*I genomic DNA cloned in the CaSpeR vector as template, we used specific oligonucleotides to sequence the promoter region of the *Dmhsp60* gene and using the Findpatterns program of the GCG Wisconsin software and the TESS program [22], we identified a TATA box at −37 bp of the transcription start site (indicated in the figure with an arrow), as well as a putative CAAT box at −129 bp and a putative Oct-1 site at −215 bp from the transcription start site. We did not find any heat shock element sequence with neither of these programs.

and their nuclei are distributed randomly (data not shown). The *hsp60*^{G93} sequence did not show any change in the ORF. Therefore, this mutation should not be affecting the protein itself but either the mRNA production or its stability. In order to answer this question, we performed a RT-PCR of poly A$^+$ RNA from heterozygous and wild-type females. We did not observe an input difference between the heterozygous and wild-type flies. However, if we used hnRNA in a similar experiment, a clear reduction of the *hsp60* transcript in the mutant *hsp60*^{G93}/FM6 females is found, compared to the wild-type ones. This difference was not seen in the *rp49* transcript, used as internal control. In agreement with the poly A$^+$ RNA RT-PCR experiment performed with the *hsp60*^{G93}/FM6 allele, we detected no difference in the protein quantity between both alleles and the wild-type strain in Western blot experiments (Fig. 3). This result indicates that, although a defect in the transcription or early processing of the *hsp60* mRNA seems to be the cause of the early death of *hsp60*^{G93} homozygous animals, heterozygous females can compensate for the lack of part of the *hsp60* transcript, accumulating the processed transcript to accomplish the required level of protein.

4. Discussion

We have previously described the *Drosophila hsp60* homologue [13]. During this characterization, we found that both

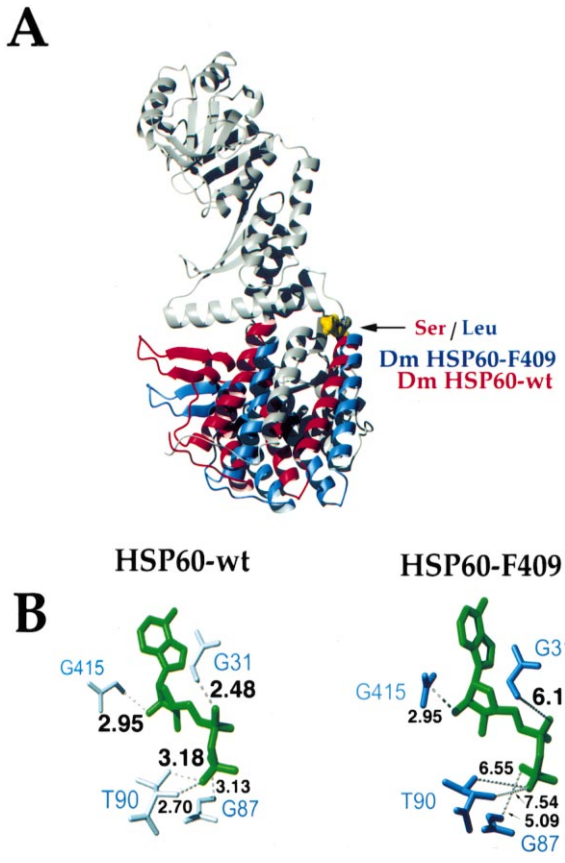

Fig. 2. Computer model of the HSP60 protein from the *hsp60*^{F409} allele. (A) The DmHSP60 wild-type and mutant sequences were modeled on top of the *E. coli* GroEL structure. The resulting model was minimized with the cvff force field using the Discover program. The figure was prepared with MolMol and Povray. DmHSP60 wild-type = red; DmHSP60^{F409} = blue. The arrow indicates the position of the amino acid change. The yellow color shows the increased volume of the leucine compared to serine. (B) Wild-type and mutant DmHSP60 structures were superposed on top of an *E. coli* GroEL subunit to model Mg-ADP binding. Mg-ADP hydrogen bonding was calculated using Ligplot. Figures were prepared with MolMol and Povray. Note the differences between the wild-type (A) and the mutant (B) active site, in the hydrogen bonding distances of the AMP molecule and the amino acids directly joined to it.
←

the gene and the protein are not heat shock inducible (unpublished results). By the analysis of the *hsp60* promoter region, we did not find any functional binding sites for heat shock factor in the 685 bp sequence upstream of the transcription initiation point, which confirms our previous data. Nevertheless, we could identify the TATA box at −37 bp, a putative CAAT box at −139 bp and a putative Oct-1 site at −225 bp from the transcription start site. Even if the molecular nature of embryonic lethal alleles from *D. melanogaster* is different (one probably affecting the protein structure and the other the amount of mRNA), the homozygous embryos of both alleles have identical phenotypes.

The *hsp60*^{F409} allele presents a C to T transition that changes serine-154 to leucine that could promote a distortion in the equatorial domain structure. Heterozygous flies can survive and look like wild-type, suggesting that the *hsp60*^{F409} mutant protein would not be toxic for the HSP60 function.

We did not observe any point mutation in the *hsp60* gene sequence of the *hsp60*^{G93} allele. We did not sequence the 5′ or 3′ UTRs, looking for mutations that could cause the lethality of this allele, due to the fact that these regions tend to be polymorphic.

The hnRNA RT-PCR analysis of the *hsp60*^{G93}/FM6 flies showed a clear reduction in the levels of the *hsp60* transcript compared to the wild-type flies. This difference was not observed in the *rp49* transcript. However, the poly A$^+$ RT-PCR and Western blot experiments showed no differences between heterozygous and wild-type flies. This suggests that in heterozygous flies, *hsp60* RNA accumulates, which alleviates the lack of some of the *hsp60* transcript. Together, these results indicate that the lethality of homozygous *hsp60*^{G93} animals could be in the lack of the *hsp60* transcript and, by conse-

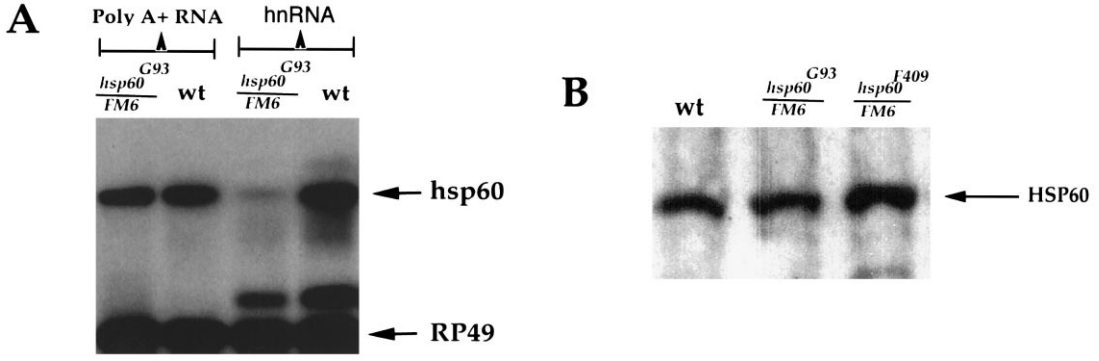

Fig. 3. Molecular analysis of the *hsp60*^{G93} allele. (A) Southern blot of the RT-PCR of poly A$^+$ RNA and hnRNA from heterozygous and wild-type females was performed using specific *hsp60* primers. The blot was hybridized with a 1.8 kb *Eco*RI restriction fragment representing the complete insert of the BP5-3 cDNA clone and with a 650 bp *Eco*RI-*Hin*dIII restriction fragment of the rp49 cDNA. Lanes indicated as poly A$^+$ RNA are the amplification products of the *hsp60* and *rp49* transcripts using poly A$^+$ RNA from mutants and wild-type flies, respectively. Lanes indicated as hnRNA represent the same amplification products with hnRNA from *hsp60*^{G93} and wild-type flies. There is no difference in the poly A$^+$ RNA level between wild-type and mutant females in the *hsp60* nor *rp49* transcripts. Nevertheless, there is a clear reduction in the hnRNA levels of the mutant *hsp60* transcript and the wild-type one. (B) Western blot analysis of whole extracts from wild-type (lane 1), *hsp60*^{F409}/FM6 (lane 2) and *hsp60*^{G93}/FM6 (lane 3) flies immunostained with an anti-HSP60 antibody (StressGen) showed no important differences in the HSP60 levels between the heterozygous and wild-type flies.

quence, of the HSP60 protein. In summary, both mutations can be considered null, one affecting the *hsp60* RNA production and the other producing a non-functional protein.

Acknowledgements: The authors are indebted to Martha Vázquez for discussions and advice. The present work was supported by PADEP-UNAM number 005343 and PAPIIT/UNAM program, Grant number IN200696.

References

[1] Hartl, F.U., Hlodan, R. and Langer, T. (1994) Trends Biochem. Sci. 19, 20–25.
[2] Hartl, F.U. (1996) Nature 381, 571–580.
[3] Gupta, R.S. (1995) Mol. Microbiol. 15, 1–11.
[4] Bukau, B. and Horwich, A.L. (1998) Cell 92, 351–366.
[5] Xu, Z., Horwich, A.L. and Sigler, P.B. (1997) Nature 388, 741–750.
[6] Langer, T., Pfeifer, G., Martin, J., Baumeister, W. and Hartl, F.U. (1992) EMBO J. 11, 4757–4765.
[7] Braig, K., Simon, M., Furuya, F., Hainfeld, J.F. and Horwich, A.L. (1993) Proc. Natl. Acad. Sci. USA 90, 3978–3982.
[8] Braig, K., Otwinowski, Z., Hegde, R., Boisvert, D.C., Joachimiak, A., Horwich, A.L. and Sigler, P.B. (1994) Nature 371, 578–586.
[9] Fenton, W.A., Kashi, Y., Furtak, K. and Horwich, A.L. (1994) Nature 371, 614–619.
[10] Sigler, P.B., Xu, Z., Rye, H.S., Burston, S.G., Fenton, W.A. and Horwich, A.L. (1998) Annu. Rev. Biochem. 67, 581–608.
[11] Viitanen, P.V., Lorimer, G.H., Seetharam, R., Gupta, R.S., Oppenheim, J., Thomas, J.O. and Cowan, N.J. (1992) J. Biol. Chem. 267, 695–698.
[12] Nielsen, K.L. and Cowan, N.J. (1988) Mol. Cell 2, 93–99.
[13] Kozlova, T., Perezgasga, L., Reynaud, E. and Zurita, M. (1997) Dev. Genes Evol. 207, 253–263.
[14] Bgatov, A.V., Pokholkova, G.V. and Zhimulev, I.F. (1986) Biol. Zent.bl. 105, 384–405.
[15] Zhimulev, I.F., Pokholkova, G.V., Bgatov, A.V., Umbetova, G.H., Solovjeva, I.V., Khudyakov, Yu.E. and Belyaeva, E.S. (1987) Biol. Zent.bl. 106, 699–720.
[16] Lindsley, D.L. and Zimm, G.G. (1992) The Genome of *Drosophila melanogaster*, Academic Press, San Diego, CA.
[17] Ashburner, M. (1989) *Drosophila*. A Laboratory Manual, pp. 108–109, Cold Spring Harbor laboratory press, Cold Spring Harbor, NY.
[18] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor laboratory press, Cold Spring Harbor, NY.
[19] Merritt, Ethan A. and Bacon, David J. (1997) Methods Enzymol. 277, 505–524.
[20] Koradi, R., Billeter, M. and Wüthrich, K. (1996) J. Mol. Graph. 14, 51–55.
[21] Birnie, G.D. (1972) Subcellular Components: Preparation and Fractionation, University park press, Baltimore, MD.
[22] Laemmli, E.K. (1970) Nature 227, 680–685.
[23] Burnette, W.N. (1981) Anal. Biochem. 112, 195–203.
[24] Schug, J. and Overton, G.C. (1997) Technical Report CBIL-TR-1997-1001-v 0.0 of the Computational Biology and Informatics Laboratory, School of Medicine, University of Pennsylvania, PA.