Gene, 96 (1990) 295–300
Elsevier

GENE 03753

Sequence of a sea urchin *hsp70* gene and its 5′ flanking region

(Recombinant DNA; heat shock; exon; intron; multigene family)

Maria La Rosa[a], Gabriella Sconzo[a], Giovanni Giudice[a,b], Maria Carmela Roccheri[a] and Marta Di Carlo[b]

[a] Dipartimento di Biologia Cellulare e dello Sviluppo, Università di Palermo, and [b] Istituto di Biologia dello Sviluppo del CNR, Palermo (Italy)

Received by G. Bernardi: 12 December 1989
Revised: 10 April 1990
Accepted: 12 July 1990

SUMMARY

We report the nucleotide sequence of a 4470-bp fragment derived from a sea urchin genomic clone containing part of a heat-shock protein 70 (Hsp70)-encoding gene. This fragment, named *hsp70* gene *II*, contains 1271 bp of the flanking region and 3299 bp of structural gene sequence interrupted by five introns and encoding the N-terminal 371 amino acids (aa) of the protein. The 5′ flanking region contains a putative TATA element, two CCAAT boxes, four heat-shock consensus sequence elements (*hse*) and one consensus sequence for binding of Sp1. Remarkable homologies were observed for deduced aa sequence and intron-exon organization between *hsp70* gene *II* and rat *hsc73* gene.

INTRODUCTION

It is well known that all the organisms investigated so far, plants, vertebrates, invertebrates and bacteria, respond to environmental stresses (e.g., temperature increase, viral infections and other agents such as ethanol or aa analogs) by transiently modifying transcription and translation and suddenly synthesizing new proteins, the so-called Hsp (Levinson et al., 1980; Schlesinger et al., 1982; Li, 1983; Li et al., 1984; LaThangue et al., 1984; Atkinson et al., 1985; Welch and Suhan, 1985; Schlesinger, 1986; Barettino et al., 1988; Lindquist and Craig, 1988; Pelham, 1988). The Hsp appear to protect the cells from injury by environmental stresses. Common to all organisms is the Hsp70, and in eukaryotes the genes encoding the Hsp70 belong to a multigenic family whose nt sequences are highly conserved in all organisms so far analyzed (Hunt and Morimoto, 1985). Some members of this gene family are constitutively expressed and generally contain introns. Other members are instead activated by heat shock or other stresses. The inducible gene activation is characterized by binding of a specific transcription factor, HSTF, to *hse*, in the promoter region (Wu, 1984; Topol et al., 1985; Sorger and Pelham, 1987a; Sorger et al., 1987; Goldenberg et al., 1988). It is well known that Hsp are quickly synthesized to help the stressed cellular metabolism (Pardue, 1988). The proteins constitutively expressed (Hsc) are involved in normal cellular physiology; for example, in cellular transmembranic import in yeast (Deshaies et al., 1988; Chirico et al., 1988).

The embryos of the sea urchin *P. lividus* respond to exposure to heat-shock or to zinc ions with the production of the two major 72.5- and 70-kDa stress proteins only if treated after the hatching blastula stage (Giudice et al., 1980; Roccheri et al., 1981a; 1986; 1988). In earlier stages they do not respond to heat stress as is observed in other embryonic developmental systems (Heikkila et al., 1985). These Hsp are transiently localized in the nucleus and permanently in the cytoplasm (Roccheri et al., 1981b; Sconzo et al., 1985). Several studies have been carried out

Correspondence to: Dr. G. Sconzo, Dipartimento di Biologia Cellulare e dello Sviluppo, 'Alberto Monroy', Università di Palermo, Via Archirafi 22, 90123 Palermo (Italy) Tel. 91-6162632 or 6164501; Fax 91-6165665.

Abbreviations: aa, amino acid(s); bp, base pair(s); Hs, heat shock; Hsc, Hs cognate protein(s); *hse*, heat-shock element(s); Hsp, heat-shock protein(s); HSTF, heat-shock transcription factor; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); p, plasmid; *P.*, *Paracentrotus*; *tsp*, transcription start point(s); ′ (prime), denotes a truncated gene at the indicated side.

aiming at understanding the cellular and developmental mechanisms underlying the Hs response in sea-urchin embryos (Sconzo et al., 1983; 1985; 1986; Giudice, 1988).

Sconzo et al. (1988) have isolated from a sea-urchin genomic library four different *hsp70* genes, as a preliminary approach to the study of the molecular mechanism of the Hs response in these embryos. The aim of the present study was the characterization and sequencing of a sea urchin cloned *hsp70* gene and the structural analysis of its promoter region.

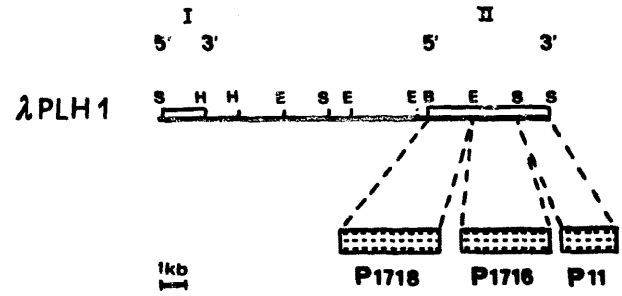

Fig. 1. Restriction map of PLH1 obtained after digestion with the following restriction enzymes: B (*Bam*HI), E (*Eco*RI), H (*Hin*dIII), S (*Sal*I). The open boxes represent the two gene regions. 5′ and 3′ indicate the direction of transcription of the *hsp70*-related genes as determined by hybridization to the specific regions of the *Drosophila hsp70* gene. The dashed boxes indicate the subclones in pUC18 plasmids utilized for the subsequent analysis. Restriction endonucleases were used according to the instructions of the manufacturer (Boehringer-Mannheim Biochemicals). pUC18 recombinants were constructed as described in Maniatis et al. (1982).

EXPERIMENTAL AND DISCUSSION

(a) Characterization and subcloning of an isolated clone

Three different clones containing four *hsp70* gene regions were previously isolated, from a genomic library of *P. lividus* sperm DNA in EMBL3 phage and one of these clones, λPHL1, was thoroughly analyzed (Sconzo et al., 1988). The insert of about 14 kb in length contains two gene regions, *I* and *II*, located at the 5′ and at the 3′ end, respectively (Fig. 1). These were detected by cross-hybridization with 5′ or 3′ fragments of the *Drosophila hsp70* gene probe, and their orientation was determined. From gene region *II* (a *Bam*HI-*Sal*I segment) three subclones p1718, p1716 and p11 in the pUC18 plasmids were obtained. The isolated fragments, when used as probes for Northern-blot analyses (Sconzo et al., 1988) were found to hybridize to the total RNA extracted from heat-treated embryos and, although more weakly, to the total RNA from untreated embryos, revealing the presence of constitutive *hsp70* RNA.

(b) Sequencing of p1718, p1716 and p11 subclones

To analyze the nt sequence of the p1718, p1716 and p11 subclones, a detailed restriction map was determined, thus identifying several restriction sites suitable for subcloning into M13-derived vectors. The arrows in Fig. 2 show the adopted strategy. The entire nt sequence obtained from the subcloned gene *II* was compared to the corresponding *Drosophila hsp70* sequence (Ingolia et al., 1980) by dot matrix analysis (Fig. 3). The results indicate that consecutive sequences of *Drosophila* are homologous to sequences of sea urchin, spaced from each other by nonhomologous regions, which indicate the presence of introns in the *hsp70* gene *II* of sea urchin. The entire nt sequence of the cloned fragments p1718, p1716 and p11 is shown in Fig. 4. The sequence analysis showed that the *hsp70* gene *II* contains five exons, five introns and about 1000 nt upstream from the coding region; the gene is interrupted at the 3′ end of the clone. The canonic consensus of splicing sites was characterized and is shown in Table I. The presence of several introns could indicate that the cloned gene *II* belongs to the constitutive *hsp70* genes, since it has been suggested that the genes encoding *hsp70* are free of introns to escape the RNA-splicing inhibition due to Hs itself (Yost and Lindquist, 1986), although inducible *hsp70* genes which contain introns have also been described in some species (Snutch et al., 1988).

The nt sequence of gene *II* was compared with that of the other *hsc70* genes, as the *hsp70A* of *Caenorhabditis*, the *hsc1* of *Drosophila*, the *hsc73* of rat, the *hsp70* of maize and *hsp70* of petunia (Snutch et al., 1988; Craig et al., 1983; Sorger et al., 1987; Rochester et al., 1986; Winter et al., 1988). The comparison revealed that all the first exons contain nearly the same number of aa deduced from the nt sequences: 65 aa in *Drosophila*, 68 aa in rat, 67 aa in sea urchin, 69 aa

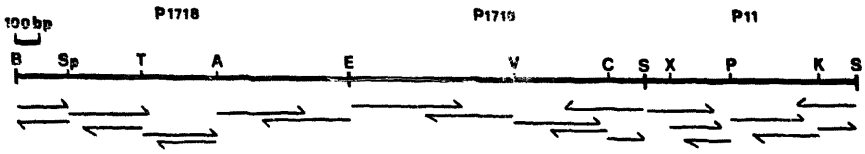

Fig. 2. Restriction map and sequencing strategy for the P1718, P1716, P11 subclones. Restriction enzymes used in subcloning are: Sp (*Sph*I), T (*Stu*I), A (*Acc*I), V (*Pvu*II), C (*Hin*cII), X (*Xba*I), P (*Pst*I), K (*Kpn*I). Other enzymes are listed in Fig. 1, legend. Arrows indicate the direction and extent of nt sequence determination on each M13 subclone. The nucleotide sequence was determined on M13 single stranded templates by the dideoxy chain-termination method (Sanger et al., 1977; 1980) using a modified T7 DNA polymerase (Tabor and Richardson, 1987).

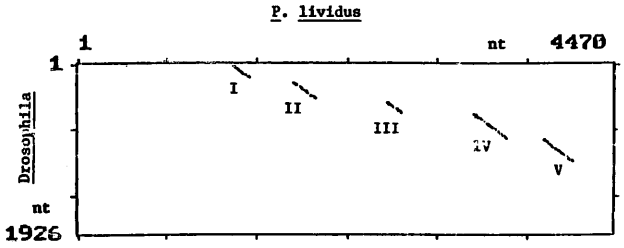

Fig. 3. Homology matrix comparison of the sea urchin-gene *II* and *Drosophila hsp70* gene nt sequences. The sequences were scanned for 30-nt homology window with up to 20-nt mismatch allowed. Regions displaying significant homology include exons I–V. Analyses of the nt sequences were performed with programs of DNASIS (Hitachi Software Engineering).

TABLE I

Introns and exons of *hsp70* gene *II*[a]

Exon No.[b]	Exon size	Donor site	Intron size	Acceptor site
1	201	TG:GTAAGT	297	TTCCAG:AT
2	207	AG:GTTAGT	555	TTACAG:AA
3	153	AG:GTAATT	583	CACTAG:GG
4	288	AG:GTAAGT	304	TTGCAG:AT
5	267	AG:GTATGC	350 + n	—

[a] Canonic consensus of splicing sites in sea urchin *hsp70* gene *II*. The size of the five exons and introns is also shown.

[b] Each number indicates a different exon.

Fig. 4. The nt sequence of gene *II*. The aa sequence deduced from the nt sequence is shown above the coding strand. Lower-case letters represent the intron sequences. The TATA, CAAT, and Sp1 consensus sequences are boxed. The *hse* consensus sequences are underlined. These sequence data will appear in the EMBL/GeneBank/DDBJ Nucleotide Sequence Database under accession No. X16544 *hsp70* gene *II*.

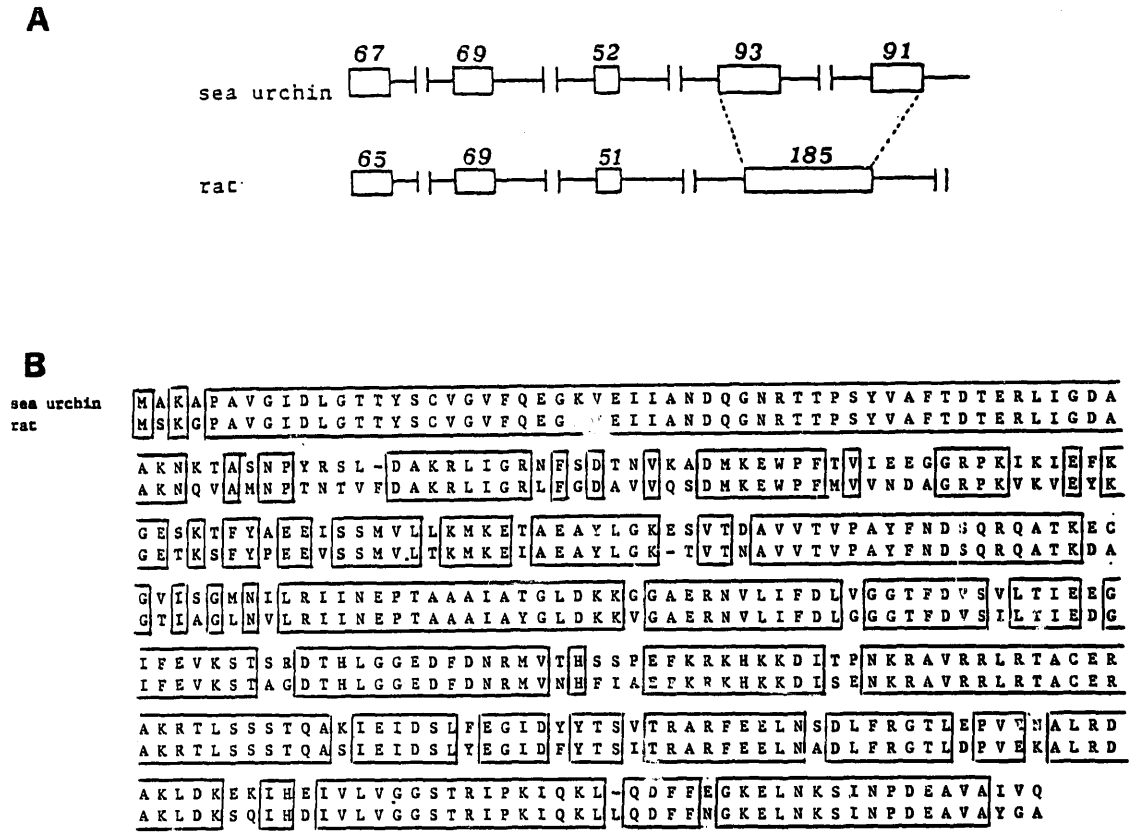

Fig. 5. Organization of genes *II* and *hsc73* and their aa sequence homologies. **(A)** Intron-exon organization of sea-urchin gene *II* and rat *hsc73* gene. The boxes indicate coding region and dashed lines the noncoding regions. The numbers above show the aa present in each exon. **(B)** Comparison of aa sequences coded by genes *II* and *hsc73*. Regions of conserved homology are boxed. Dashes denote the aa deletions at that position.

in *Caenorhabditis*, and 70 aa in maize and petunia. A significant homology in the organization of the other exons and introns was revealed by comparing the sea urchin *hsp70* gene *II* to the rat *hsc73* gene. Fig. 5A shows that the second and the third exons also exhibit great similarity in the number of aa and that the fourth exon of *hsc73* gene includes the fourth and fifth exons of *hsp70* gene *II*. The high extent of conservation of the exon-intron organization between an invertebrate and a vertebrate is remarkable.

The predicted 372-aa sequence of Hsp70 II shares 82.8% homology with the corresponding sequence of rat Hsc73 protein, 79.8% with human Hsp70 (Hunt and Morimoto, 1985), 79.5% with *Xenopus* Hsp70 (Bienz, 1984), 79.4% with *Drosophila* Hsc4 (Perkins et al., 1990), 74.4% with *Drosophila* Hsp70 (Ingolia et al., 1980), 73.9% with yeast Hsp70 (Snutch et al., 1988) and 77.2% with *Caenorhabditis* Hsp70A (Snutch et al., 1988). Among all, the extent of homology between sea urchin and rat (Fig. 5B) is prominent, considering that the % homology of the corresponding aa sequence between human Hsp70 and *Drosophila* Hsp70 is 78.4% (Hunt and Morimoto, 1985).

(c) Characterization of the promoter region

Sequence analysis of the region upstream from the start codon revealed the presence of a putative TATA box, two CCAAT boxes and four *hse* (Fig. 4). The TATA box is at −234 nt from the first ATG at which the translation starts, and supposedly there is a long 5′-end untranslated region as found in other *hsp70* genes.

Two matches to the CCAAT sequence were found at −317 and −339 nt from the first ATG and a good match to the consensus sequence for binding of the Sp1 transcription factor at nt −430.

The *hse* were found in four copies upstream from the *tsp* of gene *II*. These *hse* have an almost perfect match with the palindromic canonic consensus C--GAA--TTC--G (Bienz and Pelham, 1986) and the central GAA--TTC sequence and the palindromic structure are always preserved. The *hse1* is at −263, the *hse2* and *hse3* are adjacent and are placed at nt −773 and −788, respectively, while the *hse4* is at −1160. All positions are calculated from the first ATG.

(d) Conclusions

This paper describes the structural analysis of a *P. lividus hsp70* gene *II*; it is the first *hsp* gene analyzed so far in the sea urchin. A total of 4470 nt were sequenced: 3299 bp of structural gene interrupted by five introns encoding the first 371 aa of the Hsp70, and 1271 nt of the 5′ flanking region. The nt sequence at the 5′ end shows high levels of homology with other eukaryotic *hsp70*-related genes: a putative TATA box, two CAAT boxes, one Sp1-binding site and four *hse* were localized. All these sequences are usually binding sites for the corresponding transcription factors, which may or may not interact simultaneously.

The *hsp70* gene *II*-coding sequence is homologous to both *hsp70* and *hsc70* genes of many different organisms, such as *Drosophila*, yeast, *Caenorhabditis*, human, and *Xenopus*. The most remarkable finding, however, is the homology of the sea-urchin gene *II* to the rat *hsc73* gene. The gene *II* shows a very interesting structure, with five long introns (in the analyzed region). The absence of intervening sequences is instead a general feature of most Hs genes. At present only a few *hsp70*-related genes containing introns have been reported: *Drosophila hsc1* and *hsc2*, rat *hsc73*, *Caenorhabditis hsp70A*, and *hsp70* genes of some plants (maize, petunia).

The location of introns is also interesting from an evolutionary perspective since sea-urchin gene *II*, *hsc1* of *Drosophila*, *hsc73* of rat, *Caenorhabditis hsp70A*, and *hsp70* of maize have the first intron in the same position. Moreover, the position of the other introns is markedly conserved in sea urchin and rat, showing a preservation of the gene sequence and structure between invertebrate and vertebrate.

ACKNOWLEDGEMENTS

This work was supported by funds of CNR and of the MPI (40% and 60%). The invaluable technical assistance of Mr. D. Cascino is acknowledged.

REFERENCES

Atkinson, B.G. and Walden, D.B. (Eds.): Change in Eukaryotic Gene Expression in Response to Environmental Stress. Academic Press, New York, 1985.

Barettino, D., Morcillo, G. and Diez, J.L.: Induction of heat shock response by carbon dioxide in *Chironomus thummi*. Cell Diff. 23 (1988) 27–36.

Bienz, M.: *Xenopus hsp70* genes are constitutively expressed in injected oocytes. EMBO J. 3 (1984) 2477–2483.

Bienz, M. and Pelham, H.R.B.: Heat shock regulatory elements function as an inducible enhancer in the *Xenopus hsp70* gene and when linked to a heterologous promoter. Cell 45 (1986) 753–760.

Chirico, W.J., Waters, M.G. and Blobel, G.: 70K heat shock related proteins stimulate protein translocation into microsomes. Nature 332 (1988) 805–810.

Craig, E.A., Ingolia, T.D. and Manseau, L.J.: Expression of *Drosophila* heat shock cognate genes during heat shock and development. Develop. Biol. 99 (1983) 418–426.

Deshaies, R.J., Koch, B.D., Werner-Washburne, M., Craig, E.A. and Shekman, R.: A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. Nature 332 (1988) 800–805.

Giudice, G.: Heat shock proteins in sea urchin embryos. Develop. Growth Diff. 31 (1989) 103–106.

Giudice, G., Roccheri, M.C. and Di Bernardo, M.G.: Synthesis of 'heat shock' proteins in sea urchin embryos. Cell Biol. Int. Rep. 4 (1980) 69–74.

Goldenberg, C.J., Luo, Y., Fenna, M., Baler, R., Weinmann, R. and Voellmy, R.: Purified human factor activates heat shock promoter in a HeLa cell-free transcription system. J. Biol. Chem. 263 (1988) 19734–19739.

Heikkila, J.J., Kloc, M., Bury, J., Schultz, G.A. and Browder, L.W.: Acquisition of the heat shock response and thermotolerance during early development of *Xenopus laevis*. Develop. Biol. 107 (1985) 483–489.

Hunt, C. and Morimoto, R.: Conserved features of eukaryotic *hsp70* genes revealed by comparison with the nucleotide sequence of human *hsp70*. Proc. Natl. Acad. Sci. USA 82 (1985) 6455–6459.

Ingolia, T.D., Craig, E.A. and McCarthy, B.J.: Sequence of three copies of the gene for the major *Drosophila* heat shock induced protein and their flanking regions. Cell 21 (1980) 669–679.

LaThangue, N.B., Shriver, K., Dawson, C. and Chan, W.L.: Herpes simplex virus infection causes the accumulation of heat shock protein. EMBO J. 3 (1984) 267–277.

Levinson, W., Opperman, H. and Jackson, J.: Transition series metals and sulfhydryl reagents induce the synthesis of four proteins in eukaryotic cells. Biochim. Biophys. Acta 606 (1980) 170–180.

Li, G.C.: Induction of thermotolerance and enhanced heat shock protein synthesis in Chinese hamster fibroblast by sodium arsenite and by ethanol. J. Cell. Physiol. 115 (1983) 116–122.

Li, G.C. and Laszlo, A.: Amino acid analogues while inducing heat shock proteins sensitize CHO cells to termal damage. J. Cell. Physiol. 122 (1985) 91–97.

Lindquist, S.: The heat shock response. Annu. Rev. Biochem. 55 (1986) 1151–1191.

Lindquist, S. and Craig, E.A.: The heat shock proteins. Annu. Rev. Genet. 22 (1988) 631–677.

Maniatis, T., Fritsch, E.F. and Sambrook, J.: Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.

Pardue, M.L.: The heat shock response in biology and human disease: a meeting review. Genes Develop. 2 (1988) 783–785.

Pelham, H.R.B.: Speculations on the functions of the major heat shock and glucose regulated proteins. Cell 46 (1986) 959–961.

Pelham, H.: Coming in from the cold. Nature 332 (1988) 776–777.

Perkins, L.A., Doctor, J.S., Zhang, K., Stinson, L., Perrimon, N. and Craig, E.A.: Molecular and developmental characterization of the heat shock cognate 4 gene of *Drosophila melanogaster*. Mol. Cell. Biol. 10 (1990) 3232–3238.

Roccheri, M.C., Di Bernardo, M.G. and Giudice, G.: Synthesis of heat shock proteins in developing sea urchins. Develop. Biol. 83 (1981a) 173–177.

Roccheri, M.C., Sconzo, G., Di Bernardo, M.G., Albanese, I., Di Carlo, M. and Giudice, G.: Heat shock proteins in sea urchin embryos. Territorial and intracellular location. Acta Embryol. Morphol. Exp. 2 (1981b) 91–99.

Roccheri, M.C., Sconzo, G., LaRosa, M., Oliva, D., Abrignani, A. and Giudice, G.: Response to heat shock of different sea urchin species. Cell Diff. 18 (1986) 131–135.

Roccheri, M.C., LaRosa, M., Ferraro, M.G., Cantone, M., Cascino, D., Giudice, G. and Sconzo, G.: Stress proteins by zinc ions in sea urchin embryos. Cell Diff. 24 (1988) 209–214.

Rochester, D.E., Winer, J.A. and Shah, D.M.: The structure and expression of maize genes encoding the major heat shock protein, hsp70. EMBO J. 5 (1986) 451–458.

Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A.: Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. 143 (1980) 161–178.

Schlesinger, M.J.: Heat shock proteins: the search for functions. J. Cell Biol. 103 (1986) 321–325.

Schlesinger, M.J., Ashburner, M. and Tissières, A. (Eds.): Heat Shock, from Bacteria to Man. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982.

Sconzo, G., Roccheri, M.C., Di Carlo, M., DiBernardo, M.G. and Giudice, G.: Synthesis of heat shock proteins in dissociated sea urchin embryonic cells. Cell Diff. 12 (1983) 317–320.

Sconzo, G., Roccheri, M.C., Oliva, D., LaRosa, M. and Giudice, G.: Territorial localization of heat shock mRNA production in sea urchin gastrulae. Cell Biol. Int. Rep. 9 (1985) 877–881.

Sconzo, G., Roccheri, M.C., LaRosa, M., Oliva, D., Abrignani, A. and Giudice, G.: Acquisition of thermotolerance in sea urchin embryos correlates with the synthesis and age of the heat shock proteins. Cell Diff. 19 (1986) 173–177.

Sconzo, G., LaRosa, M., LaFarina, M., Roccheri, M.C., Oliva, D. and Giudice, G.: Isolation and characterization of a sea urchin *hsp70* gene segment. Cell Diff. 24 (1988) 97–104.

Snutch, T.P., Heschl, M.F.P. and Baillie, D.L.: The *Caenorhabditis elegans hsp70* gene family: a molecular genetic characterization. Gene 64 (1988) 241–255.

Sorger, P.K. and Pelham, H.R.B.: Purification and characterization of a heat-shock element binding protein from yeast. EMBO J. 6 (1987a) 3035–3041.

Sorger, P.K. and Pelham, H.R.B.: Cloning and expression of a gene encoding hsc73, the major hsp70-like protein in unstressed rat cells. EMBO J. 6 (1987b) 993–998.

Sorger, P.K., Lewis, M.J. and Pelham, H.R.B.: Heat shock factor is regulated differently in yeast and HeLa cells. Nature 329 (1987) 81–85.

Tabor, S. and Richardson, C.C.: DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA 84 (1987) 4767–4771.

Topol, J., Ruden, D.M. and Parker, C.S.: Sequences required for in vitro transcriptional activation of a *Drosophila hsp70* gene. Cell 42 (1985) 527–537.

Welch, W.J. and Suhan, J.P.: Morphological study of the mammalian stress response: characterization of changes in cytoplasmic organelles, cytoskeleton, and nucleoli, and appearance of intranuclear actin filaments in rat fibroblasts after heat shock treatment. J. Cell Biol. 101 (1985) 1198–1211.

Winter, J., Wright, R., Duck, N., Gasser, C., Fraley, R. and Shah, D.: The inhibition of petunia hsp70 mRNA processing during $CdCl_2$ stress. Mol. Gen. Genet. 211 (1988) 315–319.

Wu, C.: Two protein-binding sites in chromatin implicated in the activation of heat-shock genes. Nature 309 (1984) 229–234.

Yost, H.J. and Lindquist, S.: RNA splicing is interrupted by heat shock and is rescued by heat shock protein synthesis. Cell 45 (1986) 185–193.