MOLECULAR AND CELLULAR BIOLOGY, June 1990, p. 3232–3238
0270-7306/90/063232-07$02.00/0

Vol. 10, No. 6

NOTES

Molecular and Developmental Characterization of the Heat Shock Cognate 4 Gene of *Drosophila melanogaster*

LIZABETH A. PERKINS,[1]* JOHN S. DOCTOR,[2]† KANG ZHANG,[1] LESLIE STINSON,[2] NORBERT PERRIMON,[1] AND ELIZABETH A. CRAIG[2]

Howard Hughes Medical Institute, Department of Genetics, Harvard Medical School, 25 Shattuck Street, Boston, Massachusetts 02115,[1] and Department of Physiological Chemistry, University of Wisconsin, Madison, Wisconsin 53706[2]

Received 5 October 1989/Accepted 6 February 1990

The *Drosophila* heat shock cognate gene 4 (*hsc4*), a member of the hsp70 gene family, encodes an abundant protein, hsc70, that is more similar to the constitutively expressed human protein than the *Drosophila* heat-inducible hsp70. Developmental expression revealed that *hsc4* transcripts are enriched in cells active in endocytosis and those undergoing rapid growth and changes in shape.

The cells of nearly all organisms have a conserved response to environmental stresses, consisting of the synthesis of several heat shock proteins (hsps) (for review, see reference 15). The most prominent stress protein in the majority of species, hsp70, is highly conserved among procaryotic and eucaryotic species. The *hsp70* gene is a member of a family of closely related genes that includes both stress-inducible genes (*hsp*'s) and genes expressed constitutively during normal development, the heat shock cognate genes (*hsc*'s).

The heat shock cognate proteins appear to be important for normal cellular function (reviewed in references 5, 15, 21, and 27). A number of recent results indicate that the hsp70-like proteins act in an ATP-dependent manner in several cellular compartments. They may function to alter the conformations of proteins or affect protein-protein interactions (9, 18, 25). Additionally, they may play a role in the translocation of polypeptides across specific membranes (4, 7). The abundant cytoplasmic heat shock cognate protein in mammalian cells, hsc70, is involved in the ATP-dependent uncoating of clathrin from endocytotic vesicles (26, 31).

In *Drosophila melanogaster*, the *hsp70* multigene family includes five copies of the heat-inducible *hsp70* gene, one copy of the heat-inducible *hsp68* gene, and seven heat shock cognate genes, *hsc1* through *hsc7*, that are expressed during normal growth (6, 13, 15, 19). The *Drosophila* hsc70 protein, encoded by *hsc4*, is a very abundant polypeptide in all tissues and cells and is localized to a meshwork of cytoplasmic fibers concentrated around the nucleus (19).

Sequence and structure of the *Drosophila hsc4* gene. The *hsc4* gene of *D. melanogaster* was originally isolated on a recombinant plasmid, pMG34 (Fig. 1A and 2), following hybridization with a *Drosophila hsp70* gene (6). The DNA sequence of *hsc4* revealed a single open reading frame of 1,953 base pairs (bp) that potentially encodes a 651-amino-acid polypeptide with an estimated molecular weight of 71,108. S1 nuclease analysis indicated that the protein coding and the 5′ untranslated regions were not contiguous and that an intron was located 5′ of the initiating ATG (data not shown). To confirm the position of this intron, a cDNA, cD12, was isolated and the sequence of the 5′ end was determined. The cDNA sequence diverged from the genomic DNA sequence at the ATG (Fig. 1B). The protein coding and 5′ untranslated regions of the *hsc4* gene were interrupted by a 1.6-kilobase (kb) intron.

The deduced amino acid sequence of the *hsc4* gene (Fig. 3) was 73% identical to that of the *Drosophila* heat-inducible hsp70 (12) and 85% identical to that of the human hsc70 polypeptides (8). Furthermore, *Drosophila* hsc70 was 80% identical to *Caenorhabditis elegans* hsp70A, which is abundant throughout development and only marginally heat inducible (29). An unresolved question is whether constitutively expressed *hsp70*-related genes, such as *hsc4*, and those induced by stress, e.g., *hsp70*, have identical or different functions. The fact that *Drosophila hsc4* is more closely related to vertebrate *hsc70*-like genes than an inducible gene from *D. melanogaster* suggests that the constitutively expressed proteins may be functionally distinct from the stress-induced proteins.

***In situ* hybridization to embryos reveals stage- and tissue-specific enrichment of *hsc4* transcripts.** Northern (RNA blot) analysis demonstrates that the major 2.3-kb *hsc4* transcript is expressed throughout embryonic, larval, pupal, and adult development at relatively constant levels (6) (data not shown). In situ hybridization to wild-type embryos was performed as described by Hafen and Levine (11) or Tautz and Pfeifle (30). Radioactive DNA probes were labeled by nick translation with [^{35}S]dCTP (New England Nuclear Corp.) to a specific activity of approximately 5×10^7 cpm/μg, and the autoradiograms were developed after 2 to 3 days. Nonradioactive probes were prepared essentially by the protocol provided with the nonradioactive labeling and detecting kit (Boehringer Mannheim, catalog no. 1093657). *hsc4* transcripts were localized in a complex spatial and temporal pattern during embryogenesis (Fig. 4 and 5), which was superimposed onto a basal level of expression apparent in virtually all cells of the developing embryo.

Enrichment of *hsc4* transcripts was first observed during late syncytial and cellular blastoderm stages and during early gastrulation in the cytoplasmic compartment between the

* Corresponding author
† Present address: McArdle Laboratory, University of Wisconsin, Madison, WI 53706.

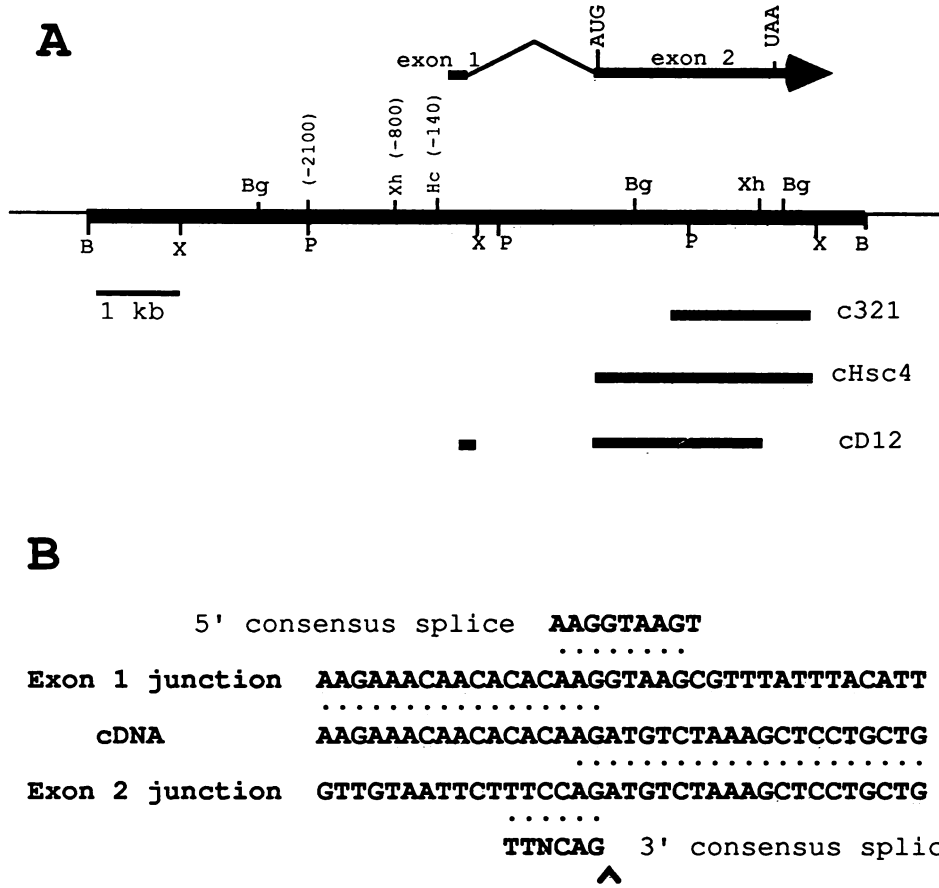

FIG. 1. (A) Restriction map of pMG34, a recombinant plasmid containing the entire *Drosophila hsc4* gene. The direction of transcription, positions of the intron and exons 1 and 2, start (ATG) and stop (TAA) codons of translation, and relevant restriction sites are included (B, *Bam*HI; Bg, *Bgl*II; Hc, *Hin*cII; P, *Pst*I; Xb, *Xba*I; Xh, *Xho*I). Below the restriction map, the approximate extents of three *hsc4* cDNAs are depicted. cDNA clone c321 was isolated from a 3- to 12-h embryonic cDNA library (23) in a differential screen designed to identify genes preferentially expressed in neuroblasts rather than differentiated neurons (L. A. Perkins, A. P. Mahowald, and N. Perrimon, submitted for publication) and was determined to encode *hsc4* sequence based on its localization to 88E on the salivary gland polytene chromosomes, Southern blot analysis with pMG34 as a probe, and partial DNA sequence analysis. cDNA clone cHsc4 was isolated at high stringency from a size-selected 9- to 12-h embryonic library with c321 as a probe (35). cDNA clone cD12 was isolated from an embryonic cDNA library provided by M. Goldschmidt-Clermont with pMG34 as a probe. (B) Location of the intron/exon boundaries in the *Drosophila hsc4* gene. This comparison shows the nucleotide sequence from cDNA cD12 aligned with the genomic DNA sequence from pMG34. This alignment does not permit the unambiguous determination of the precise boundaries of exons 1 and 2, but the predicted splice site (^), based on the eucaryotic consensus (17), is shown.

blastoderm nuclei and the yolk (Fig. 4A and B). These stages are characterized by the rapid assembly of cellular membranes to compartmentalize the nuclei. Tissue enrichment was next observed in neuroblasts of both the head and extending germ band (Fig. 4C, D, and E). Unlike transcripts from *Delta* and members of the *achaete-scute* gene complex, which are enriched in subsets of neuroblasts enlarging within the neurogenic ectoderm (2, 24, 32), *hsc4* transcripts were only observed in newly segregated neuroblasts internal to the ectoderm. Enrichment was clearly observed in neuroblasts from the procephalic neurogenic ectoderm (Fig. 4C and D) and continued to be enriched in specific derivatives of this region (Fig. 4H). Enrichment of *hsc4* transcripts was observed in cells of the embryonic gut from anterior and posterior midgut invagination to hatching (Fig. 4F to H) and transiently in developing mesodermal cells (Fig. 4F and G). Enrichment in the gut occurred while the cells were undergoing numerous cellular processes: mitoses, expansions, stretching, and volumetric growth (3). Enrichment in the mesoderm occurred as the somatic and splanchnic mesoderms became separate layers (Fig. 4F and G) and was readily apparent as the somatic muscles formed and single cells fused into syncytial myotubes and differentiated into somatic muscles (3).

During late embryogenesis, *hsc4* transcripts were most abundant in the garland gland (Fig. 4G and H), an organ postulated to segregate and store waste products (34). Cells from the garland gland are very active in endocytosis via coated vesicles. In fact, electron microscopy reveals the cortex of these cells to be a labyrinth of endocytotic pits or channels that "pinch off" to form clathrin coated vesicles (14; C. Poodry, personal communication). Since a clathrin "uncoating ATPase" activity has been detected in *Drosophila* cells (28), we propose that hsc70 in the garland gland functions in the uncoating of clathrin triskelions.

In conclusion, *hsc4* transcripts are present in most if not all cells during embryonic development but are enriched in cells active in endocytosis and those undergoing rapid growth and changes in shape. Studies in other organisms have demonstrated high levels of hsc70 in rapidly growing

FIG. 2. Complete nucleotide sequence of the *Drosophila hsc4* gene contained on plasmid pMG34. The entire protein-coding region, delimited by start (ATG) and stop (TAA) codons, is 1,953 bp and contains no introns. The predicted 5′ and 3′ splice sites of the intron separating exon 1 and exon 2 and the predicted start of transcription (5′ END) are indicated by arrows. Nucleotides are numbered with reference to the predicted start of transcription (+1). The 5′ untranslated region of the *hsc4* gene is approximately 120 bp (6); intronic sequences are not included in this numbering scheme. A consensus TATA box at −23 to −31, and two regions (−91 to −104 and −144 to −157) with sequence similarity to the consensus heat shock element (20) are observed upstream of the start of transcription. Stars indicate identical matches to this heat shock element, T——GAA——TAA——G. We have marked the approximate 5′ end of the *hsc4* transcript as +1.

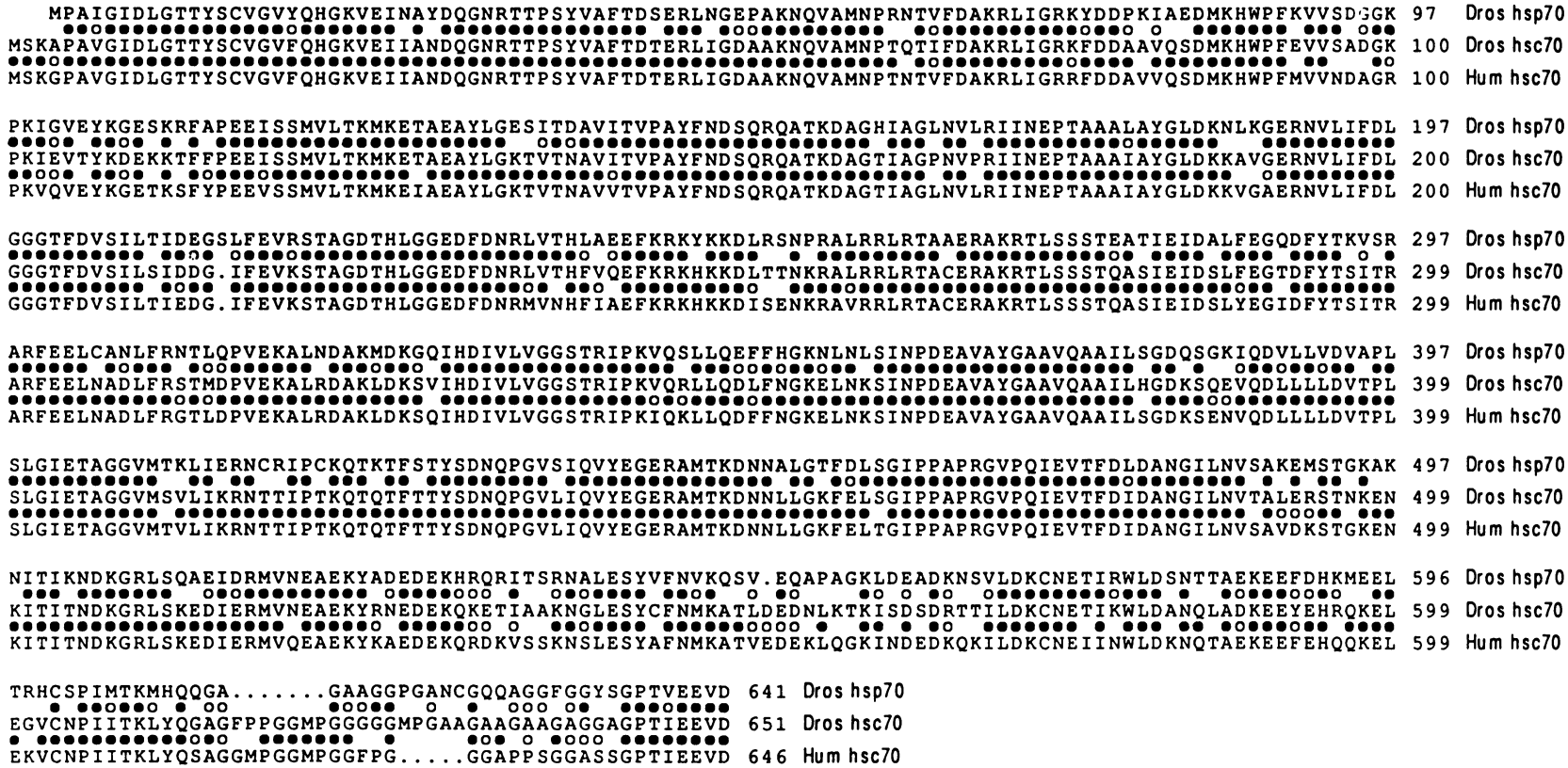

FIG. 3. Comparison of *Drosophila* hsc70, human hsc70, and *Drosophila* hsp70 amino acid sequences. The predicted amino acid sequence encoded by the *Drosophila hsc4* gene (*Dros* hsc70) is compared with the predicted amino acid sequences of human hsc70 (Hum hsc70 [8]) and *Drosophila* hsp70 (*Dros* hsp70 [12]). The sequences were aligned by using the GAP program of the University of Wisconsin Genetics Computer Group with a gap weight of 5.00 and a length weight of 0.30. Identical amino acids (●) and conservative differences (○) are indicated.

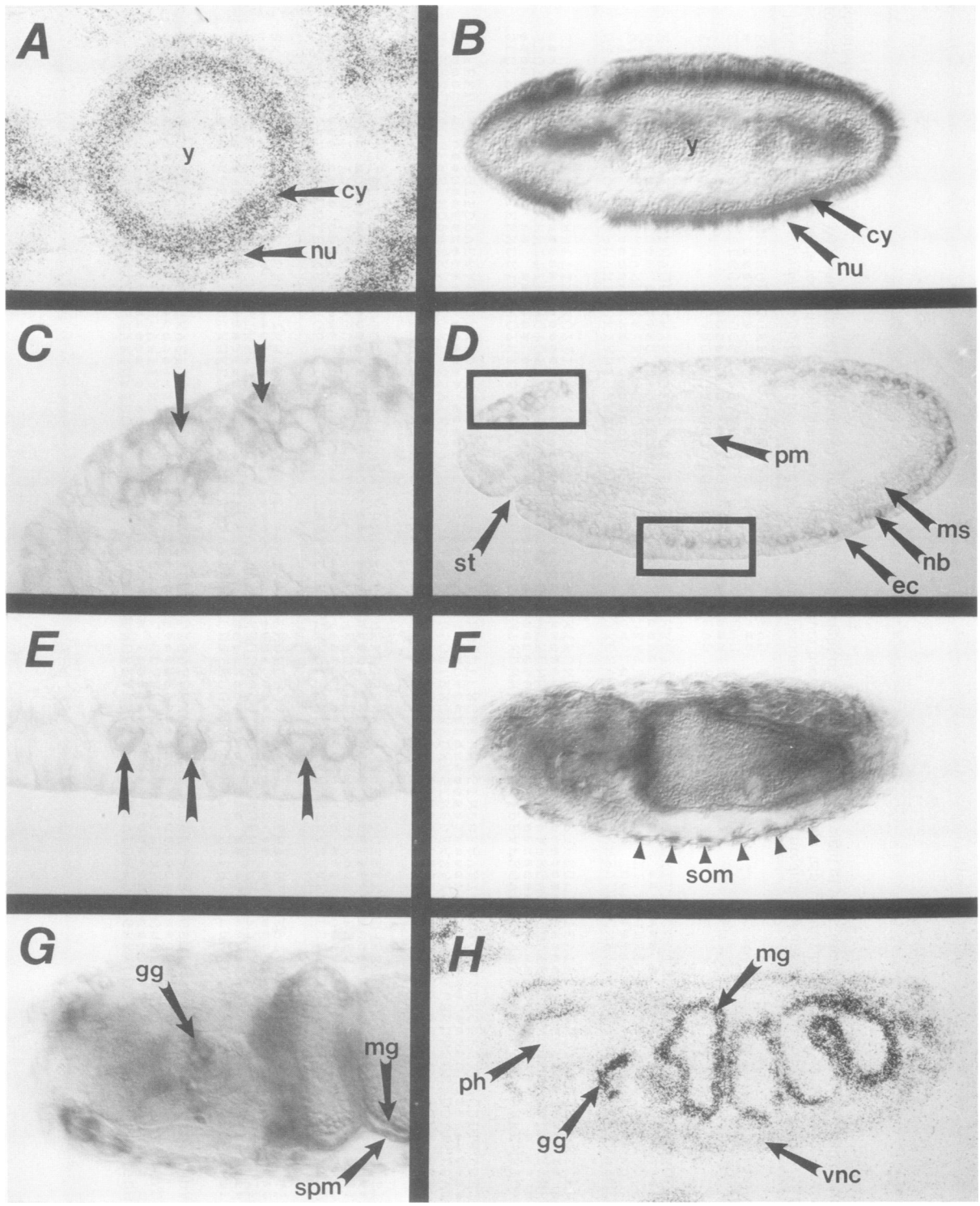

FIG. 4. Expression of *hsc4* transcripts during embryogenesis. (A) Transverse section with ventral at the bottom; (H) parasagittal section with anterior to the left and ventral at the bottom; (B to G) whole-mount embryos labeled with nonradioactive probes (30). From fertilization through the early stages of cleavage, expression of *hsc4* transcripts is uniformly distributed in the embryo (not shown). During the late syncytial and cellular blastoderm stages through early gastrulation, most of the *hsc4* transcript is observed between the peripherally positioned nuclei (nu) and the central yolk (y), i.e., in the cytoplasmic compartment (cy) (A and B). *hsc4* transcripts remain essentially uniform in distribution at the basal level in all embryonic tissues until the germ band is almost fully extended. At this time a punctate band of more intense hybridization internal to the region of the ectoderm (ec) and exterior to the mesoderm (ms), where neuroblasts (nb) have segregated (D, enlarged in E), is detected. With development the intensity of the band increases, presumably due to either increased numbers of cells becoming enriched for the *hsc4* transcript or increased expression of the transcript in the enriched cells. Enrichment is also observed in the procephalic neurogenic regions (D, enlarged in C). Throughout the remaining stages of embryogenesis, the lining of the developing gut is

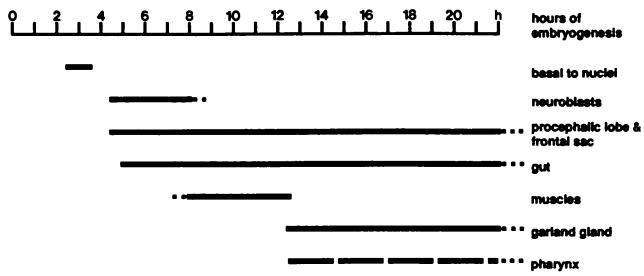

FIG. 5. Schematic summary of tissue-specific enrichment for *hsc4* transcripts during embryonic development. The top scale represents hours of embryogenesis with hatching (h) occurring at 22 h of embryonic development. The solid lines indicate the times at which enrichment was observed in the tissues indicated at the right. The dashed line for the pharynx indicates that *hsc4* transcription was below the basal level of transcription observed in other nonenriched tissues.

embryonic and transformed cells and in some secretory cells (1, 10, 16, 22), suggesting that *hsc4* is a homolog of the mammalian *hsc70* gene. Consistent with this interpretation is the fact that *hsc4* is more closely related to the mammalian hsc70 than to the heat-inducible *Drosophila* hsp70 protein. In addition, like the mammalian hsc70 protein, the *Drosophila hsc4* protein product translocates to the nucleus after thermal stress (19, 33).

We are indebted to Beth Noll for excellent technical assistance, to Mike Slater and Jeff Shilling for assistance in the compilation and analysis of hsc4 sequence, and to Karen Palter for advice and sharing unpublished results. cDNA clone 321 was isolated by L.A.P. while a postdoctoral fellow in the laboratory of A.P. Mahowald. We thank K. Palter and M.L. Pardue for critical reading and suggestions on an early version of this manuscript.

J.S.D. was supported by a postdoctoral fellowship (GM 10131) from the National Institutes of Health (NIH). This work was supported by the Howard Hughes Medical Institute and a grant from the American Chemical Society to N.P. and by Public Health Service grant GM27870 to E.A.C. from NIH.

LITERATURE CITED

1. **Bensaude, O., and M. Morange.** 1983. Spontaneous high expression of heat shock proteins in mouse embryonal carcinoma cells and ectoderm from day 8 mouse embryos. EMBO J. **3**:173–177.
2. **Cabrera, C. V., A. Martinez-Arias, and M. Bate.** 1987. The expression of three members of the *achaete-scute* gene complex correlates with neuroblast segregation in *Drosophila*. Cell **50**: 425–433.
3. **Campos-Ortega, J. A., and V. Hartenstein.** 1985. The embryonic development of *Drosophila melanogaster*, p. 85–165. Springer-Verlag, New York.
4. **Chirico, W. J., G. Waters, and G. Blobel.** 1988. 70K heat shock related proteins stimulate protein translocation into microsomes. Nature (London) **332**:805–810.
5. **Craig, E. A.** 1985. The heat shock response. Crit. Rev. Biochem. **18**:239–280.
6. **Craig, E. A., T. D. Ingolia, and L. J. Manseau.** 1983. Expression of *Drosophila* heat-shock cognate genes during heat shock and development. Dev. Biol. **99**:418–426.
7. **Deshaies, R. J., B. D. Koch, M. Werner-Washburne, E. Craig, and R. Schekman.** 1988. A subfamily of stress proteins facilitates translation of secretory and mitochondrial precursor polypeptides. Nature (London) **332**:800–805.
8. **Dworniczak, B., and M. E. Mirault.** 1987. Structure and expression of a human gene coding for a 71 kd heat shock 'cognate' protein. Nucleic Acids Res. **15**:5181–5197.
9. **Flynn, G. C., T. G. Chappell, and J. E. Rothman.** 1989. Peptide binding and release by proteins implicated as catalysts of protein assembly. Science **245**:385–390.
10. **Giebel, L. B., B. P. Dworniczak, and E. K. F. Bautz.** 1988. Developmental regulation of a constitutively expressed mouse mRNA encoding a 72-kDa heat shock-like protein. Dev. Biol. **125**:200–207.
11. **Hafen, E., and M. Levine.** 1986. The localization of RNAs in *Drosophila* tissue sections by *in situ* hybridization, p. 139–158. *In* D. B. Roberts (ed.), *Drosophila*: a practical approach. IRL Press, Washington, D.C.
12. **Ingolia, T. D., E. A. Craig, and B. J. McCarthy.** 1980. Sequence of three copies of the gene for the major *Drosophila* heat shock induced protein and their flanking regions. Cell **21**:669–679.
13. **Ingolia, T. D., and E. A. Craig.** 1982. *Drosophila* gene related to the major heat shock-induced gene is transcribed at normal temperatures and not induced by heat shock. Proc. Natl. Acad. Sci. USA **79**:525–529.
14. **Kosaka, T., and K. Ikeda.** 1983. Reversible blockage of membrane retrieval and endocytosis in the garland cell of the temperature-sensitive mutant of *Drosophila melanogaster*, *shibire^{ts1}*. J. Cell Biol. **97**:499–507.
15. **Lindquist, S., and E. A. Craig.** 1988. The heat-shock proteins. Annu. Rev. Genet. **22**:631–677.
16. **Morange, M., A. Diu, O. Bensaude, and C. Babinet.** 1984. Altered expression of heat shock proteins in embryonal carcinoma and mouse early embryonic cells. Mol. Cell. Biol. **4**: 730–735.
17. **Mount, S. M.** 1982. A catalog of splice junction sequences. Nucleic Acids Res. **10**:509–518.
18. **Munro, S., and H. R. B. Pelham.** 1986. An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell **46**:291–300.
19. **Palter, K. B., M. Watanabe, L. Stinson, A. P. Mahowald, and E. A. Craig.** 1986. Expression and localization of *Drosophila melanogaster* hsp70 cognate proteins. Mol. Cell. Biol. **6**:1187–1203.
20. **Pelham, H.** 1985. Activation of heat-shock genes in eucaryotes. Trends Genet. **1**:31–35.
21. **Pelham, H. R. B.** 1986. Speculations on the functions of the major heat shock and glucose-regulated proteins. Cell **46**:959–961.
22. **Pinhasi-Kimhi, O., D. Michalovitz, A. Ben-Zeev, and M. Oren.** 1986. Specific interaction between the p53 cellular tumour antigen and major heat shock proteins. Nature (London) **320**: 182–185.
23. **Poole, S. J., L. M. Kauvar, B. Drees, and T. Kornberg.** 1985. The *engrailed* locus of *Drosophila*: structural analysis of an embryonic transcript. Cell **40**:37–43.
24. **Romani, S., S. Campuzano, and J. Modolell.** 1987. The *achaete-scute* complex is expressed in neurogenic regions of *Drosophila* embryos. EMBO J. **7**:2085–2092.
25. **Rothman, J. E.** 1989. Polypeptide chain binding proteins: cata-

highly enriched. By germ band shortening and extending through dorsal closure, the developing somatic (som) and splanchnic (spm) musculature become enriched (F and G). Finally, late in embryogenesis *hsc4* tissue enrichment is observed in the proventriculus, gut (mg), the garland gland (gg), and that region of the frontal sac dorsal to the pharynx which is derived from the procephalic lobe (H). Note that at this stage the lining of the pharynx (ph) shows hybridization below background. This is the only tissue observed during embryogenesis to have less than basal-level hybridization. Other abbreviation: vnc, ventral nerve cord. That the hybridization observed is specific for *hsc4* and not other hsp70-related genes was supported by the fact that a 0.8-kb *Pst*I-*Sal*I fragment taken from the highly divergent 3′ end of *hsc4* (K. Palter, unpublished observations) (panels B to G) showed patterns of hybridization identical to that with cDNAs c321 and cHsc4, which extend into the less divergent 5′ end (panels A and H).

lysts of protein folding and related processes in cells. Cell **59**:591–601.
26. **Rothman, J. E., and S. L. Schmid.** 1986. Enzymatic recycling of clathrin from coated vesicles. Cell **46**:5–9.
27. **Schlesinger, M. J.** 1986. Heat shock proteins: the search for functions. J. Cell. Biol. **103**:321–325.
28. **Schlossman, D. M., S. L. Schmid, W. A. Braell, and J. E. Rothman.** 1984. An enzyme that removes clathrin coats: purification of an uncoating ATPase. J. Cell Biol. **99**:723–733.
29. **Snutch, T. P., M. F. P. Heschl, and D. L. Baille.** 1988. The *Caenorhabditis elegans* hsp70 gene family: a molecular characterization. Gene **64**:241–255.
30. **Tautz, D., and C. Pfeifle.** 1989. A nonradioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals a translational control of the segmentation gene *hunchback*. Chromosoma **98**:81–85.
31. **Ungewickell, E.** 1985. The 70-kd mammalian heat shock proteins are structurally and functionally related to the uncoating protein that releases clatherin triskelia from coated vesicles. EMBO J. **4**:3385–3391.
32. **Vassin, H., K. A. Bremer, E. Knust, and J. A. Campos-Ortega.** 1987. The neurogenic gene *Delta* of *Drosophila melanogaster* is expressed in neurogenic territories and encodes a putative transmembrane protein with EGF-like repeats. EMBO J. **6**: 3431–3440.
33. **Welch, W. J., and J. P. Suhan.** 1985. Morphological study of the mammalian stress response: characterization of changes occurring in intracellular membraneous organelles, cytoskeletal elements, nucleus and nucleoli and appearance of intranuclear actin containing inclusion bodies in rat fibroblasts following heat shock. J. Cell Biol. **101**:1198–1211.
34. **Wigglesworth, V. B.** 1972. The principles of insect physiology, 7th ed., p. 440–442. Chapman and Hall, London.
35. **Zinn, K., L. McAllister, and C. S. Goodman.** 1988. Sequence analysis and neuronal expression of fasciclin I in grasshopper and *Drosophila*. Cell **53**:577–587.