Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 525–529, January 1982
Genetics

Drosophila gene related to the major heat shock-induced gene is transcribed at normal temperatures and not induced by heat shock

(gene isolation/DNA sequence/intervening sequence/cDNA extension)

THOMAS D. INGOLIA AND ELIZABETH A. CRAIG

Department of Physiological Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706

Communicated by James F. Crow, September 14, 1981

ABSTRACT **A gene related to the major heat shock-induced (hsp70) gene of *Drosophila* has been isolated from the sibling species *D. melanogaster* and *D. simulans*. This heat shock-cognate (hsc70) gene is present at cytological locus 70C. The primary sequence of approximately one-third of the protein-coding region has been compared with that of the hsp70 gene; 72% homology of the base sequence and 74% homology of the deduced amino acid sequence was found. In the codon specifying amino acid 66, the hsc70 gene contains an insertion of 1.7 kilobases; the hsp70 genes contain no intervening sequences. The sequence at the 5′ and 3′ junctions of the insertion is similar to that found in many intervening sequences. cDNA extension experiments indicate that the hsc70 gene is transcribed at normal temperatures in adult flies and that transcription is not enhanced by heat treatment.**

Heat shock and certain other stimuli result in a dramatic change in the pattern of gene expression in *Drosophila*. Transcription of most genes is suppressed and expression of a small set of relatively inactive genes is enhanced (for review, see ref. 1). A similar response is observed in cells of a wide variety of species including yeast (2), *Dictyostelium* (3), chicken, and man (4). Little is known about the physiological role of the heat shock response or the heat shock genes. It is known that many of the heat shock polypeptides reside in the nucleus (5, 6) and that one of the chicken polypeptides will associate with the *src* gene product of Rous sarcoma virus (7).

The major heat shock polypeptide (hsp) expressed in the sibling species *D. melanogaster* and *simulans* is hsp70. Also, the mRNA encoding hsp70 is the most abundant species accumulating after treatment (8, 9). Copies of the gene for hsp70 are present at two cytological locations, 87A7 and 87C1. The total copy number per haploid genome is about five: two at 87A and three at 87C (10–13). The primary sequences of the gene copies from both chromosomal locations are similar, ≈94% conserved (14, 15). No intervening sequences have been found in genes from either location (12, 14, 16). We report the isolation and characterization of a heat shock cognate (hsc) gene related to the hsp70 gene.

MATERIALS AND METHODS

General Methods. Restriction enzyme digests, agarose and acrylamide gel electrophoresis, DNA·DNA hybridizations, plasmid DNA isolation, labeling by DNA kinase, Maxam and Gilbert DNA sequence analysis (17), and *in situ* hybridizations were carried out as described (12, 14).

cDNA Extension Experiments. Primers for the synthesis of cDNA were isolated and endlabeled with DNA kinase by standard procedures (14). After concentration by ethanol precipitation, the DNA was mixed with 50 μg of total RNA isolated by the guanidine thiocyanate method (18) from flies maintained at 25°C or from flies heated for 40 min at 37°C before harvest. The DNA was denatured by heating to 80°C for 1 min in 90% formamide. The nucleic acids were hybridized in 80% formamide/0.4 M NaCl/10 mM 1,4-piperazinediethanesulfonic acid, pH 6.8/2.5 mM EDTA and either maintained for 2 hr at 48°C (termed stringent) or allowed to slowly cool from 48°C to 35°C over a 2-hr period (termed nonstringent). The nucleic acid was precipitated twice and then added to the avian myeloblastosis virus reverse transcriptase (RNA-dependent DNA nucleotidyltransferase) reaction mixture [50 mM Tris·HCl, pH 8.3/10 mM $MgCl_2$/10 mM dithiothreitol/1 mM dGTP/1 mM dATP/1 mM dCTP/1 mM TTP containing actinomycin D at 50 μg/ml and 12 units of reverse transcriptase (total vol, 20 μl)]. After phenol extraction, the DNA was subjected to electrophoresis on 5% acrylamide/urea gels.

Isolation of MG21 and SG52. SG52 was isolated from a library of *Hin*dIII-cleaved *D. simulans* chromosomal DNA inserted into the *Hin*dIII site of pBR322. It was identified by hybridization to a cDNA clone of the hsp70 mRNA (HS4) as described (12). Subsequently, MG21 was isolated from a library of Schneider cell DNA by hybridization to the *Bgl* II fragment isolated from the intervening sequence of SG52 (see Fig. 1). Recombinant DNA studies were carried out in accordance with National Institutes of Health guidelines for recombinant DNA research.

RESULTS

Primary Sequence of Genes Analyzed. Primary sequence analysis was carried out on recombinant clones G13, SG52, and MG21. G13, described previously (14), is a hsp70 gene present at cytological locus 87A. SG52 and MG21 are similar heat-shock cognates isolated from *D. simulans* and *D. melanogaster* DNA, respectively. Partial restriction maps of the clones are shown in Fig. 1.

Primary sequences were determined by using the chemical cleavage method (17). The regions analyzed are indicated in Fig. 1 as enlarged portions of the restriction maps. The best alignment of the primary sequences of the related genes is shown in Fig. 2.

Comparison of hsp70 and hsc70 Primary Sequences. Previously, we reported the primary sequence of the protein coding region of one copy of the 87C hsp70 gene from clone G3 (14). One reading frame is devoid of stop codons for >2000 base pairs (bp); the region between the first ATG and the first stop codon could encode a protein of 70,270 daltons. Comparison of the first 303 bases of this coding region between an 87C copy (G3) and an 87A copy (G13) shows 18 differences or 94% homology

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

Abbreviations: hsp, heat shock polypeptide; hsc, heat shock cognate; bp, base pair(s).

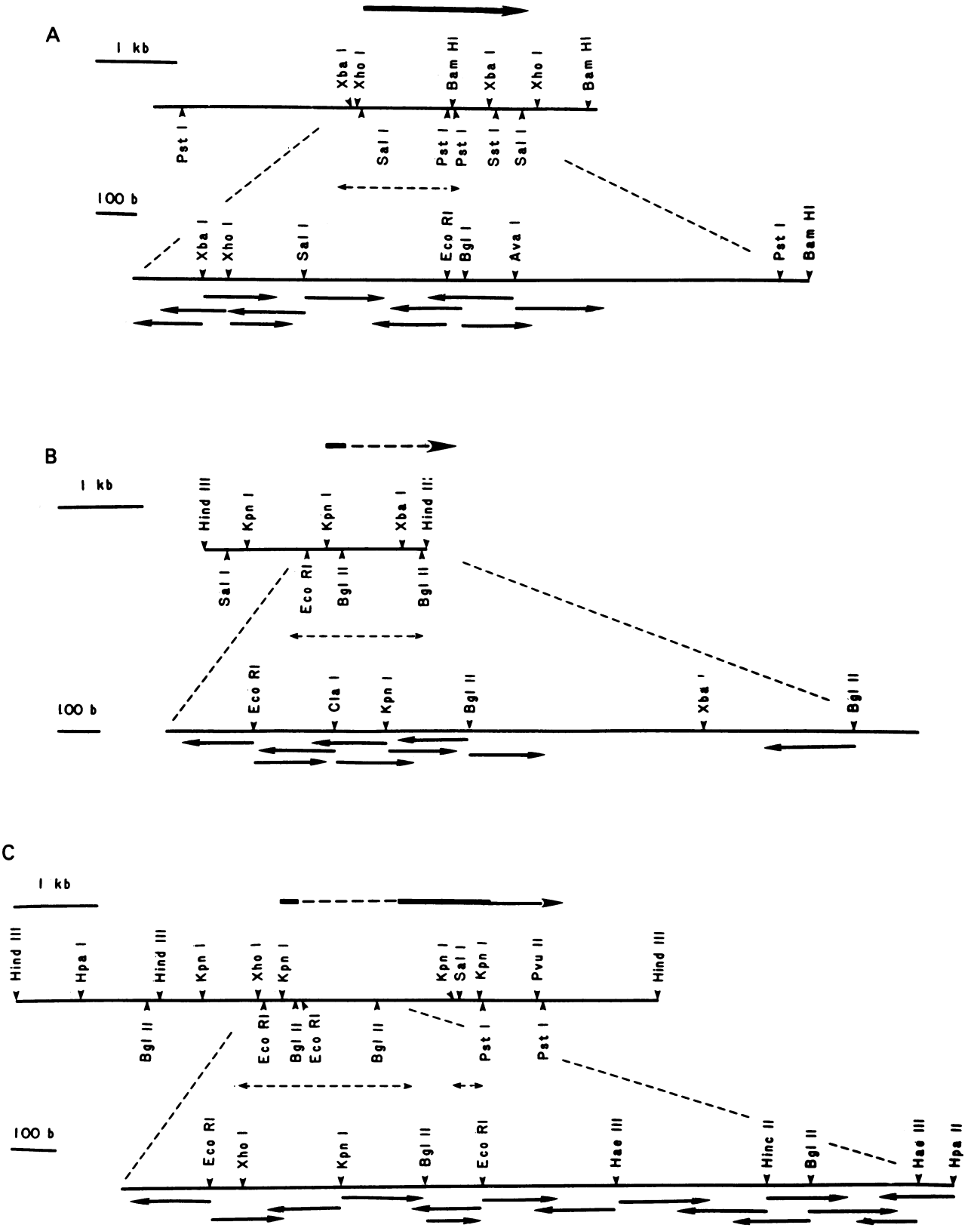

FIG. 1. Restriction map and strategy used to determine nucleotide sequences of regions of G13 (*A*), MG21 (*B*), and SG52 (*C*). Restriction maps were deduced as described (12). Solid regions of the single headed arrows above the maps indicate regions homologous to hsp70 mRNA; dashed regions represent regions containing intervening sequences; Thin solid regions of the single-headed arrow of SG52 indicate regions of homology predicted if no other intervening sequences are present. Double-headed dashed arrows underneath the maps represent regions for which sequences were determined. DNA sequence analysis was carried out from the sites indicated by arrows on the enlarged portions of the maps. The lengths of the arrows correspond to the number of nucleotides actually sequenced from each start. The sequence of the small segment of SG52 not included in the expanded map was determined from the *Sal* I and *Kpn* I sites. b, Base(s); kb, kilobase(s).

a

```
                                                                          aa1                                                         aa16
SG52  GATCTTTCGGCGGGACCTTGGACGCGGTGGTGGGGTCTAATTAAAGCCACAGCCGACATGCCCAAA TTA CCA GCG GTT GGC ATT GAT CTT GGT ACC ACG TAC TCG TGT GTC GGT
G21   ....C...................T............C..CC........................ ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
87A   AT.GAAGACAAGAAGAGAACTCTGAATACT.TCAACAAGTCTTT.C.G.GGAGGA..TCA.A..C. ATG ..T ..T A.. ..A ..C ... ..G ..C ... ..C ... ..C ..C ..G ...
87C   AT.GAA.ACAAGAAGAGAACTCTGAATACT.TCAACAAGT.ACCGAGA.AGAAGA..TCA.A..CA ATG ..T ..T A.. ..A ..C ... ..G ..C ... ..C ... ..C ..C ..G ...

                                                                                                                                   aa49
SG52  GTC TTT CAG CAT GGC AAG GTG GAG ATC ATC GCC AAC GAC CAG GGC AAT CGC ACC ACG CCC AGC TAT GTG GCT TTC ACG GAG TCG GAG CGT CTG ATC GGC
G21   ... ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
87A   ... .AC ..A ... ... ... ... ... ..T ... ... ... ... ... ... ..C ... ... ... ..G TC. ..C ... ... ... ..A ..T ... ..A ..C ..C ... ...
87C   ... ..C ... ... ... ... ..T ... ..T .A. ... T.T ... ... ... ..C ... ... ... ..G TC. ..C ... ... ... ..A ..C ... ..A ..C ..C .AT ..T

                                                              aa65
SG52  GAT GCG GCC AAG AAC CAG GTG GCC ATG AAT CCC AAC AAC ACG ATC TTT GGTGAGTCTCGTTCGCATTGTCCCTGCCACAGATCTTAGTTGATCTCTGCCTGGGTCTCGATCTCTC
G21   ... ... ... ... ... ..A ... ... ... ... ... ... ... ... ... ... ..............................................G.........T.......T........
87A   ... C.. ... ... ... ... ... ... ... ..C ... .GA ... ..A G.G ... .
87C   ..A C.. ... ... ... ... ... ... ... ..C ... .GA ... ..A G.G ... .

SG52  GCCATGATCACTTTAAATTACTGGCAACATTTTCATGAATGAACTGTGCGCATATTTTCTTTATTATCTCGCTTTCGAGTTTATTTTAAGAA-TTCTCACTCACTGGCGCCAAAAAATGGATTTCACTCCC
G21   ..............-.............................................................................G...A.......................-..............

                                           1,500 b
SG52  AATGCGAGAGAAAATCTCGCCTCCCACTAAAGTCCA  - - -  AAATATTTCT-AAAATGTTTCTTTCGAATACAAGGACTTTAATTCCAACTATGCCGAACATTCTAACCAAACCAAAATTAGTCA
G21   .T..T...............................  - - -  ..........C....-........A...T...A........--------..C..A......................GC..A.

SG52  ATTTG---TCAACTATTATTTCTTTTACTTTATGTAAATAGTATTCCAGCATTCGTTTAGTCC--TTTCTCTAAACATTTTCCATGC----TCCCATCCATAAGATTTGCTGAGCAGATCTAATGCAAGCC
G21   .....CAA.-..A.............TAACA.............A..................TT....C.................ACCCA........A....................

SG52  TTCTTTCTATTTTTTGTACTTTCATTTCAGTCTCAGGGTCCGTATAGAAATCCCAATAGAAAGCCCAGACCTGCCCAGAAACACAAAACAAGCCCAAAATTCCAGACACACAATGCCACATATTTATTATC

                                                                                                         aa66
SG52  CCAACACAACCCGTATATGTTAGCTAATAAAGTAAATATGTTGTCCAAGTGCGCTGATGTTGTTTCATTGGATTCTGATCGAGTGATTCTGACTTGCAGAT GCC AAG CGG CTG ATT GGC CGC CGT
87A                                                                                                        .C ... ... ..A ..C ..C ... ..A AAA
87C                                                                                                        .C ... ... ..A ..C ..C ... ..A AAA

                                                                                                       aa101      aa200      aa203
SG52  TTC GAC GAC GCC ACC GTC CAG TCG GAC ATG AAG CAT TGG CCC TTC GAG GCG TTC GCC GAG AAC GGA AAG CCG CGC ATC CGG  - - -  GGT ACC TTC GAC
87A   .A. ... ... C.. .AG A.. GCA GA. ... ... ... ..C ... ..T ... AA. .T. G.A AG. ..C GG. ... ... ..C AAG ... G..  - - -
87C   .A. ... ..T C.. .AA A.. G.. GA. ... ... ... ..C ... ..T ... AAA .TT G.A .G. ..T GG. ... ... ..C AAG ... G..  - - -  ..C ... ... ..T

                                                                                                                                  aa236
SG52  GTA TCC GTG CTG ACC ATC GAG GAT GGC --- --- TTC GAG GTG AAG GCC ACC GCT GGC GAC ACC CAT CTC GGT CGC GAG GAC TTT GAT AAC CGG CTG GTG
87C   ..C ... A.C ... ... ... ..C ..G ..A TCT CTG ... ... ... CGC T.. ..A ..C ..A ... ..A ..C T.G ..C ... ... ... ... ..C ... ... ..A ..C

                                                                                                                                  aa269
SG52  AAC CAC CTC GTC CAG GAG TTT CAG CGC AAA CAT GGG AAG GAT CTG GGC CAG AAC AAG CGG GCT TTG AGG CGC CTC CGC ACC GCC TGC GAA AGG GCA AAA
87C   .C. ... ..G .CG G.. ... ..C A.. ... ..G T.C AA. ... ... ... C.. TCC ... CCT ..C ..C C.A C.A ..A ... A.A ..A ..A GCT ... C.. ..C ..G

                                                                                                                                  aa302
SG52  CGC ACC CTC TCC TCG TCC ACG CAG GCC AGC ATC GAG ATC GAT TCC CTC TTT GAG GGC GTC GAC TTC TAC ACC TCG GTG ACA CGG GCT CGC TTC GAG GAG
87C   ... ..A ... ... ..T AG. ... G.. ... .C. ... ... ... ..C GCA T.G ... ... ... CAA ... ... ... ... AAA .A. .GC ..T ..C A.G ..T ... ...

                                 aa311
SG52  TTG AAT GGG GAT CTG TTC CGA GGT ACC
87C   C.G TGC .C. A.C ..C ... ..C AAC ...
```

b

```
                  aa1                                                                                                             aa30
SG52  met pro lys leu pro ala val gly ile asp leu gly thr thr tyr ser cys val gly val phe gln his gly lys val gly ile ile ala asn asp gln
87C   --- --- --- met         ile                                                     tyr                             asn     tyr

                                                                                                                                  aa63
SG52  gly asn arg thr thr pro ser tyr val ala phe thr gly ser gly arg leu ile gly asp ala ala lys asn gln val ala met asn pro asn asn thr
87C                                                   asp                 asn     gly pro                             arg

                                                                                                                                  aa96
SG52  ile phe asp ala lys arg leu ile gly arg arg phe asp asp ala thr val gln ser asp met lys his trp pro phe glu ala phe ala gly asn gly
87C   val                                     lys tyr         pro lys ile ala gly                             lys val val gly asp gly

                      aa101            aa200                                                                                      aa223
SG52  lys pro arg ile arg     - - - -     gly thr phe asp val ser val leu thr ile glu asp gly --- --- phe glu val lys ala thr ala gly asp
87C           lys     gly     - - - -                     ile                     asp gly     ser leu             arg ser

                                                                                                                                  aa256
SG52  thr his leu gly gly glu asp phe asp asn arg leu val asn his leu val gln glu phe gln arg lys his gly lys asp leu gly gln asn lys arg
87C                                                       thr         ala glu         lys         tyr lys             arg ser     pro

                                                                                                                                  aa289
SG52  ala leu arg arg leu arg thr ala cys glu arg ala lys arg thr leu ser ser ser thr gln ala ser ile glu ile asp ser leu phe glu gly val
87C                                   ala                                         glu     thr                 ala                 gln

                                                                                  aa311
SG52  asp phe tyr thr ser val thr arg ala arg phe glu glu leu asn gly asp val phe arg gly thr
87C                   lys     ser                         cys ala asn             asn
```

FIG. 2. (*a*) Comparison of primary sequences and proposed amino acid (aa) sequences of hsp70 and hsc70. The 87C hsp70 sequence is taken from ref. 14. The DNA sequences are aligned with those of SG52; ·, aligned DNA that contains the same base as SG52. A triplet is underlined if the base change(s) results in an aa change. aa1, first ATG of the hsp70 genes; bases in italic lettering are intervening sequences. (*b*) Comparison of portions of deduced aa sequences of hsc70 and hsp70. Amino acids encoded by hsp70 are given only in positions where differences occur.

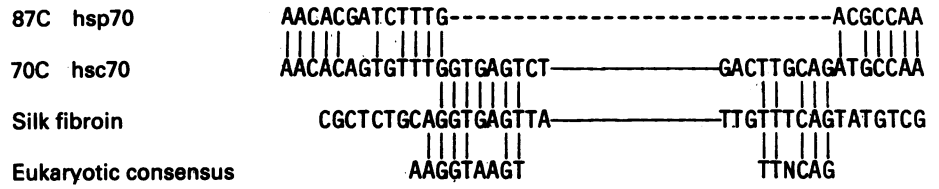

FIG. 3. Comparison of sequences at exon–intron junctions. The hsp70 and hsc70 genes are aligned near the intron–exon junction. The silk fibroin junction (19) and the eukaryotic consensus junction (20) are aligned with hsc70 to give the best homology. Solid line, intron.

(Fig. 2). These 18 base changes result in 3 of a possible 101 amino acid differences or 97% homology between the proposed amino acid sequences.

Regions of homology among the hsc70 and the hsp70 genes were identified. When optimally aligned, SG52 and MG21, the *D. simulans* and *D. melanogaster* hsc genes, respectively, had only three mismatches in the region from +1 to +192. All three are changes at the third position of codons, and none would result in a change in the amino acid sequence. Therefore, we conclude that SG52 and MG21 are analogous genes from the sibling species.

A more extensive comparison was made between the protein coding regions of the 87C hsp70 (G3) and SG52. The base sequences that would encode amino acids 1–101 and 200–311 were compared, and several features were noted. First, the homologous regions are interrupted in the codon for amino acid 66 by an insertion of 1.7 kilobases in the SG52 gene. The sequence at the 5′ and 3′ junctions of the insertion in the hsc70 sequence shows a striking similarity to the conserved sequences found at many intron–exon borders (Fig. 3). Comparison of portions of the intervening sequence found in both SG52 and MG21 shows much less homology (especially toward the 3′ end) than that between the protein coding regions. This decreased homology is characterized by small deletions and insertions as well as base substitutions (Fig. 2).

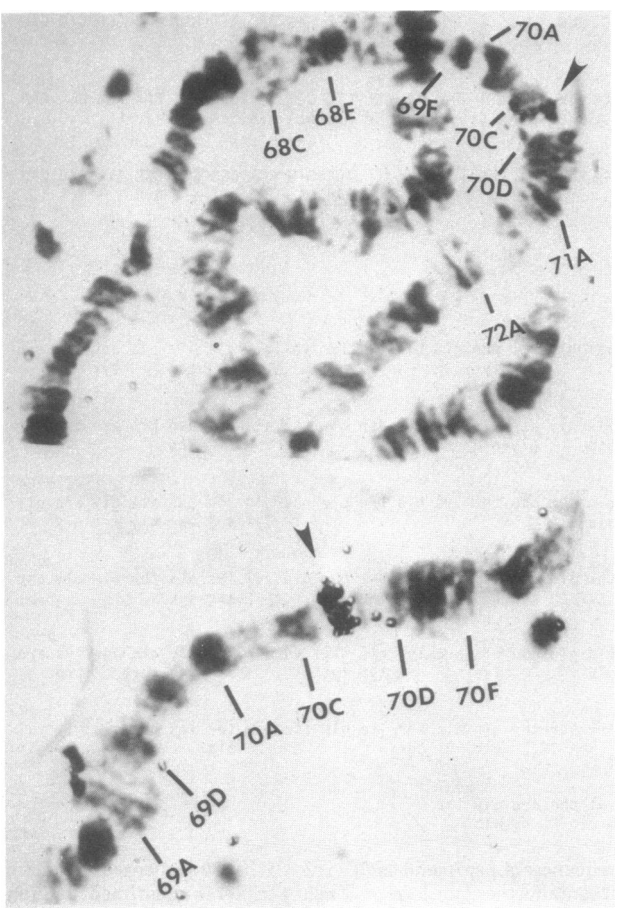

FIG. 4. *In situ* hybridization of MG21 to salivary gland polytene chromosomes of *D. melanogaster*. [^{3}H]cRNA was synthesized from MG21; 300,000 cpm; 7-day exposure.

Cytological Localization of the hsc70 Gene. The hsc70 gene was localized on the *D. melanogaster* genome by hybridization of ^{3}H-labeled MG21 to salivary gland polytene chromosomes. [^{3}H]cRNA synthesized from MG21 hybridized strongly to the cytological locus 70C (Fig. 4). Longer exposures showed minor sites of hybridization in chromosomal subdivisions 87–89, including 87A and 87C (unpublished observations). This result is consistent with earlier findings that sequences encoding hsp70 are found at chromosomal loci 87A and 87C (10–13).

cDNA Extension Experiments to Determine Presence of hsc70 Transcripts. cDNA extension experiments were carried out to determine whether the hsc70 gene is transcribed. ^{32}P-5′-labeled primers from the NH_2-terminal end of the protein coding regions of G13 and MG21 were prepared. A 104-bp *Hinf*/*Kpn* I fragment corresponding to positions +27 to +131 and a 61-bp *Bgl* II/*Bst*NI fragment corresponding to positions +90 to +141 were isolated from MG21 (locus 70C) and G13 (locus 87A), respectively. The primer fragments were denatured and hybridized to RNA, and the primer DNA was elongated by using reverse transcriptase. The size of the cDNA synthesized was determined by electrophoresis on a 6% acrylamide/8 M urea gel using labeled denatured *Hae* III fragments of pBR322 as a standard. The result of an experiment using the MG21 primer after hybridization to RNA from heat shocked flies under stringent and nonstringent conditions is shown in Fig. 5 (lanes d and e, respectively). Hybridization under stringent conditions promotes the extension to cDNA that migrates as a single band with a size of ≈510 bases. Based on our previous sequence and 5′-end analysis of hsp70 genes from 87A and 87C (14), we conclude that this cDNA is ≈80 bases longer than expected for the heat-induced mRNA transcribed from the hsp70 genes. If the hybridization is carried out under nonstringent conditions, a second, very intense, band of the expected size of 425 bases is observed.

To confirm that the intense lower molecular weight band represents the extension using the normal heat-shock mRNA as template, we used the primer from G13 (locus 87A). Since the end of this primer is located at +141 compared with +131 for the MG21 (locus 70C) primer, an extension 10 bases larger, or 435 bases, is expected. Lanes f and g of Fig. 5 show the bands resulting after hybridization under stringent and nonstringent conditions, respectively; the expected intense band, ≈435 bases long, was observed after both hybridizations. These results indicate that a mRNA transcribed from MG21, or a similar gene, is present in flies after heat treatment and that the 5′ end of the mRNA is ≈85 bases longer than that of the hsp70 transcript. The same experiment was carried out using RNA from untreated flies (lanes b and c). A band of 510 bases was found after hybridization under stringent (and nonstringent) conditions. No band corresponding to the heat-inducible transcript was seen. The data suggest that the 70C hsc gene is transcribed at 25°C and not induced by heat treatment. Whether transcrip-

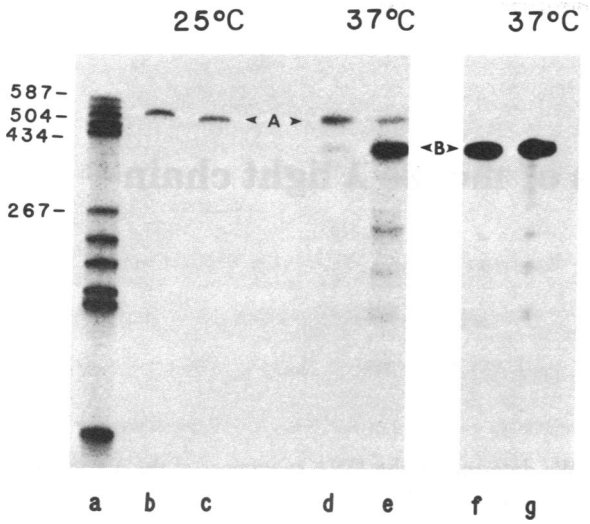

FIG. 5. Analysis of RNA transcribed from hsc70 genes using cDNA extensions. A 104-bp *HinfI/Kpn* I fragment from SG52 (lanes b–e) and a 61-bp *BstNI/Bgl* II fragment from G13 (lanes f and g) were end labeled at the *Kpn* I and *Bgl* II sites, respectively, and used as primers. Lanes: a, *Hae* III-cleaved pBR322; d–g, hybridization to RNA isolated from heat shocked flies; b and c, hybridization to RNA from untreated flies. Hybridization was carried out under stringent (lanes b, d, and f) and nonstringent (lanes c, e, and g) conditions. A, 510-base fragment; B, 425-base fragment. Numbers on the left represent fragment sizes (in bases).

tion of the 70C hsc gene is repressed after heat treatment is not addressed by these experiments, since only steady-state levels of RNA are measured.

DISCUSSION

A gene from cytological locus 70C on the left arm of chromosome 3, which is ≈76% homologous to the genes of the heat-inducible proteins found on the right arm of chromosome 3 at cytological loci 87A and 87C, has been isolated from *D. melanogaster*. This cognate is found in both *D. melanogaster* and *D. simulans*, suggesting that this gene existed before speciation and is not unique to one strain or cell line. Evidence presented here suggests that this gene is transcribed at the normal growth temperature of 25°C and that, unlike the hsp70 genes from 87A/C, transcription is not enhanced by heat treatment. This is consistent with the fact that 70C is not the site of a heat-shock puff nor the site of detectable transcription after heat shock as analyzed by *in situ* hybridization or autoradiography (21–23).

The primary sequence of the 70C gene resembles that of a functional gene. Distinguishing characteristics of pseudogenes such as frame-shift mutations and large deletions are not found in the 70C gene. Primary sequences of the protein coding regions of the 70C and 87C genes, which could be translated into ≈21,000 daltons of protein, were compared. Although no mutations that would disrupt the reading frame of the mRNA were found, one deletion of the 70C gene relative to the 87C gene occurred; this was a 6-bp deletion that would cause a loss of two amino acids but still maintain the reading frame. The pseudogenes that have been analyzed in the most detail, those of globin, contain several small deletions or base substitutions in a shorter segment of DNA than has been analyzed here, which would result in either a change in reading frame or chain termination (24, 25). The functionality of the 70C gene will not be established until it is demonstrated that a protein product is encoded by the 70C gene.

Comparison of primary sequence data suggests that the 70C gene contains an intervening sequence in the codon for amino acid 66. The junctions of the insertion sequence with the hsp70 coding regions show similarity to intron–exon junctions of many eukaryotic genes. Only one of the seven heat-induced genes of *Drosophila*, hsp83, contains an intervening sequence (16). It is of interest that the hsp83 protein is the one heat-inducible protein found in detectable quantities at normal growth temperature (20). Perhaps only those heat-shock genes or heat-shock cognates that are expressed significantly at 25°C will be found to contain intervening sequences.

The function of the heat-shock proteins has not been determined. It is known that the heat-shock proteins, except hsp83, migrate to the nucleus (5, 6). Hsp70 has been hypothesized to be a nonhistone chromosomal protein (5). The experiments reported here establish the existence of a gene related to the hsp70 gene that is expressed at the normal growth temperature but not induced by heat shock. The protein product of this hsc gene may perform a function analogous to that of hsp70. This is the first indication that a function similar to that performed by hsp70 may be required for normal growth of the fly. Supporting this hypothesis is our finding that the *Drosophila* genome contains at least two other hsc genes, at least one of which is normally expressed but not heat inducible (unpublished observations). Further characterization of the hsc genes should help elucidate the role of the heat-shock proteins.

We thank Brian McCarthy in whose laboratory initial experiments were performed; S. Vernick for excellent technical assistance; S. Tobin, L. Manseau, and R. Kreber for help with the *in situ* experiments; and R. Metzenberg for thoughtful reading of the manuscript. This research was supported by a National Institutes of Health grant to E.A.C. and a fellowship from the Leukemia Society of America to T.D.I.

1. Ashburner, M. & Bonner, J. (1978) *Cell* **15**, 1277–1286.
2. Miller, M. J., Xuong, N. & Geiduschek, E. P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5222–5225.
3. Loomis, W. T. & Wheeler, S. (1980) *Dev. Biol.* **79**, 399–408.
4. Kelley, P. & Schlessinger, M. (1978) *Cell* **15**, 1277–1286.
5. Velazquez, J. M., DiDomenico, B. J. & Lindquist, S. (1980) *Cell* **20**, 679–689.
6. Arrigo, A.-P., Fakan, S. & Tissieres, A. (1980) *Dev. Biol.* **78**, 86–103.
7. Oppermann, H., Levinson, W. & Bishop, J. M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1067–1071.
8. McKenzie, S. L., Henikoff, S. & Meselson, M. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1117–1121.
9. Spradling, A., Penman, S. & Pardue, M. L. (1975) *Cell* **41**, 395–404.
10. Henikoff, S. & Meselson, M. (1977) *Cell* **12**, 441–451.
11. Ish-Horowicz, D., Holden, J. J. & Gehring, W. (1977) *Cell* **12**, 643–652.
12. Craig, E. A., McCarthy, B. J. & Wadsworth, S. W. (1979) *Cell* **16**, 575–588.
13. Ish-Horowicz, D., Pinchin, S. M., Schedl, P., Artavanis-Tsakovas, S. & Mirault, M.-E. (1979) *Cell* **18**, 1351–1358.
14. Ingolia, T. D., Craig, E. A. & McCarthy, B. J. (1980) *Cell* **21**, 669–679.
15. Torok, I. & Karch, F. (1980) *Nucleic Acids Res.* **8**, 3105–3123.
16. Holmgren, R., Livak, K., Morimoto, R., Freund, R. & Meselson, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5390–5394.
17. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **75**, 5613–5617.
18. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **24**, 5294–5299.
19. Lewin, B. (1980) *Cell* **22**, 324–366.
20. Suzuki, Y., Tsuad, M., Tsujimoto, Y., Okshima, Y. & Griza, P. E. (1978) *Carnegie Inst. Washington Yearb.* **78**, 84–100.
21. Ashburner, M. (1970) *Chromosoma* **31**, 356–376.
22. Bonner, J. J. & Pardue, M. L. (1976) *Cell* **8**, 43–50.
23. Spradling, A., Pardue, M. L. & Penman, S. (1977) *J. Mol. Biol.* **109**, 559–587.
24. Vanin, E. F., Goldberg, G. I., Tucker, P. W. & Smithies, O. (1980) *Nature (London)* **286**, 222–225.
25. Nishioka, Y., Leder, A. & Leder, P. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2806–2809.