DNA
Volume 8, Number 4, 1989
Mary Ann Liebert, Inc., Publishers
Pp. 233-243

Characterization of the hsp70 Multigene Family of *Caenorhabditis elegans*

MARK F.P. HESCHL* and DAVID L. BAILLIE

ABSTRACT

Our laboratory has been characterizing the hsp70 multigene family from the nematode *Caenorhabditis elegans* as the first step to the genetic characterization of the heat shock response in a relatively simple multicellular eukaryote. Two gene members, *hsp-1* and *hsp-2*ps have already been characterized (Snutch *et al.*, 1988; Heschl and Baillie, 1989). The third gene member, *hsp-3*, is expressed constitutively and is non-heat inducible; its mRNA is most abundant at the L1 larval stage. The *hsp-3* protein (hsp70C) shares a high degree of identity with the rat grp78 protein and has a long, hydrophobic leader sequence. The carboxyl terminus of hsp70C has the putative ER-retention signal, KDEL. The fourth gene member, *hsp-6* is expressed constitutively and moderately heat inducible. A partial *hsp-6* protein (hsp70F) sequence shares a higher degree of identity with the *Escherichia coli* dnaK protein than with eukaryotic hsp70 proteins. The predicted amino-terminal half of the hsp70F polypeptide also contains a long, amphiphilic leader sequence similar to mitochondrial import leader sequences. These two genes encode proteins that potentially cross intracellular membranes. We compared the 5'-flanking DNA from the *C. elegans hsp-3* gene to fragment containing enhancer activity from the rat grp78 gene regulatory region (Lin *et al.*, 1986). A 23-nucleotide sequence was conserved between the two promoter regions. This sequence shares approximately 80% identity between these two evolutionary distant organisms. A comparison to other hsp70 genes did not reveal any conservation of this 23-nucleotide sequence. We propose that this sequence may be involved in a unique aspect of the regulation of the *C. elegans*' grp78-like gene and the rat grp78 gene.

INTRODUCTION

ORGANISMS OFTEN RESPOND to sudden elevation of temperature or stressful circumstances by synthesizing a specific set of proteins called heat shock proteins (hsp). Simultaneously, the synthesis of most proteins normal to development is inhibited. The sizes and structures of these stress-induced proteins appear to be highly conserved throughout evolution. In all eukaryotic genomes, there are several closely related sequences encoding the 70,000-dalton hsps (hsp70; reviewed in Craig, 1985; Lindquist, 1986). These sequences encode proteins normally found in unstressed cells and organisms. These proteins have been termed hsc70 (heat shock cognates; Ingolia and Craig, 1982; Craig *et al.*, 1983). In addition, some of the hsc70 genes may be regulated developmentally (Milarski and Morimoto, 1986; Zakeri and Wolgemuth, 1987).

The various hsp70-related genes are organized into an hsp70 multigene family. The hsp70 multigene family has been well characterized in *Saccharomyces cerevisiae* (Craig and Jacobsen, 1984, 1985; Craig *et al.*, 1987; Werner-Washburne *et al.*, 1987) and identified in humans (Mues *et al.*, 1986). In *S. cerevisiae*, the hsp70 multigene family can be subdivided into five subfamilies based on both genetic analysis and nucleotide sequence identity. Two of these subfamilies, SSA and SSC, are essential for viability (Werner-Washburne *et al.*, 1987; Craig *et al.*, 1987). Members of the SSA subfamily encode proteins that may be involved in precursor protein import into the ER and mitochondria (Deshaies *et al.*, 1988; Chirico *et al.*, 1988). A fifth subfamily has been defined recently and is represented by *KAR2*, the yeast grp78-like gene (M. Rose, personal communication). In mammals, the grp78 gene has been isolated and characterized (Munro and Pel-

Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6.
*Present address: Department of Physiological Chemistry, University of Wisconsin—Madison, 1300 University Avenue, Madison, WI 53706.

ham, 1986; Ting and Lee, 1988). The grp78 proteins had been studied previously because of their ability to bind to immunoglobulin heavy chains prior to their secretion. Synthesis of the mammalian ER-localized grp78 protein is enhanced when cells are deprived of glucose or stimulated with calcium ionophores (for review, see Lee, 1987).

The nematode *Caenorhabditis elegans* is a relatively simple eukaryote, ideal for the study of both genetics and the biochemistry of genes and proteins. With this in mind, our laboratory has been characterizing the *C. elegans* hsp70 multigene family as the basis for a genetic analysis of the hsp70 genes in a higher eukaryote. An initial characterization of the heat shock response in *C. elegans* has been done (Snutch and Baillie, 1983) and several of the hsp70 genes have been isolated (Snutch *et al.*, 1988). This family has been estimated to have at least nine members. Two of the genes have been characterized previously. *hsp-1* (LGIV) is a constitutively expressed, heat-inducible gene (Snutch *et al.*, 1988) and *hsp-2*ps (LGX) is a pseudogene of *hsp-1* (Heschl and Baillie, 1989). The *hsp-1* gene encodes the hsp70A protein, which shares a high degree of identity with the *Drosophila melanogaster* heat-inducible hsp70 protein and members of the *S. cerevisiae* SSA gene subfamily.

In this study, we have characterized and identified the *hsp-3* gene and the *hsp-6* gene from *C. elegans*. The *hsp-3* gene, encoding hsp70C, is expressed constitutively and non-heat inducible under the conditions tested. The *hsp-3* mRNA is found at maximum levels in the L1 larvae stage and decreases more than sixfold in the adult nematode. *hsp-6*, encoding hsp70F, has low basal mRNA levels and exhibits a two- to three-fold increase in mRNA after a temperature upshift (Snutch *et al.*, 1988). We find that hsp70C shares a high degree of identity with the rat grp78 protein and appears to be the *C. elegans* equivalent of grp78. A stretch of nucleotides in the 5′ regulatory region of the *hsp-3* gene is also found in the rat grp78 gene regulatory region. We suggest that this highly conserved element may be involved in a unique aspect of expression shared between the two grp78 genes. The hsp70F shares more identity with the *dnaK* protein from *Escherichia coli* Bardwell and Craig, 1984) than with eukaryotic hsp70 proteins. The predicted hsp70F amino acid sequence contains a long amino-terminal peptide characteristic of a mitochondrial import leader sequence.

MATERIALS AND METHODS

Construction of plasmids for sequencing

The phage containing the hsp70 genes (*hsp-3*, hsl 140 and *hsp-6*, hsl B4; Snutch, 1984; Snutch *et al.*, 1988) were digested with *Eco* RI or *Hin*d III and subcloned into the plasmid vectors pUC19 (Norrander *et al.*, 1983) or Bluescript⁺(Stratagene). The plasmids containing the hsp70 genes were identified (*hsp-3* 5′ end, pCes433 and pCes434, *hsp-3* 3′ end, pCes428; *hsp-6* 5′ end, pCes405) and their orientations determined based on the pattern derived from restriction digests. Overlapping plasmid deletions were made using either exonuclease III (Henikoff, 1987) or restriction enzymes.

Plasmid DNA preparation

Plasmid DNA for deletion and sequencing reactions was prepared using the mini alkali lysis method (Maniatis *et al.*, 1982), except that two phenol/sevag extractions were done and after the first ethanol precipitation, the air-dried pellet was resuspended in 0.25 *M* sodium acetate and reprecipitated with ethanol. After digestion with RNase, the plasmid DNA samples wee precipitated with polyethylene glycol (Hattori and Sakaki, 1986).

DNA sequencing and sequence analysis

Dideoxy sequencing was performed on denatured plasmid DNA (Chen and Seeburg, 1985; Sanger *et al.*, 1980). DNA sequences were analyzed using the computer program Microgenie (Beckman). Visual inspection and preparation of the DNA sequences for publication was done with the aid of the computer program ESEE (E. Cabot, personal communication).

RESULTS

Characterization of the hsp-3 *gene from* C. elegans

A restriction map of the *C. elegans hsp-3* gene is shown in Fig. 1. The complete sequence of the *hsp-3* gene and its flanking DNA is presented in Fig. 2. The *hsp-3* gene contains three introns (Figs. 1 and 2) of 46, 238, and 104 nucleotides. In *C. elegans*, the intron boundaries are highly conserved, easily recognized, and characterized by the consensus sequences AG/GTAAGT (5′ splice site) and TTTTCAG/G (3′ splice site) (Karn *et al.*, 1983; Spieth *et al.*, 1985). These consensus sequences were used to aid in the identification of the *hsp-3* intron boundaries as well as gaps and shifts in the amino acid sequence in a comparison to the *hsp-1* gene sequence and predicted amino acid sequence (Snutch *et al.*, 1988). One of the intron positions was confirmed by sequencing over the appropriate region surrounding the third intron from a partial cDNA. Although the assignment of intron positions based on the consensus sequences is reliable in *C. elegans*, the confirmation of the first two intron positions awaits sequencing of a full-length cDNA. None of the *hsp-3* introns were in the same position as the *hsp-1* introns (Snutch *et al.*, 1988). Just upstream from the start of translation is a 3′ splice site from −23 to −31 (Fig. 2, block 3) which may be involved in a *trans* splicing reaction (Krause and Hirsh, 1987).

There is a long 3′ untranslated sequence of 623 nucleotides. We estimate the size of the mRNA to be approximately 2.7 kb based on the positioning of the poly(A) ad-

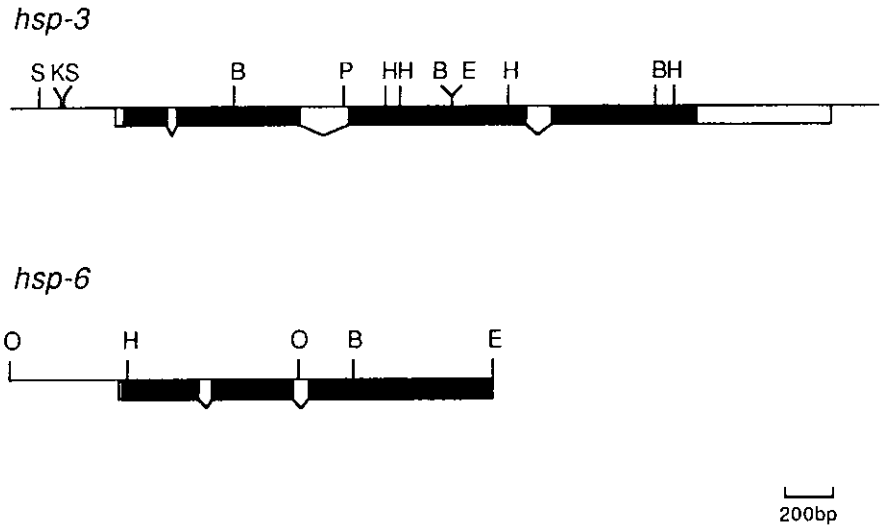


FIG. 1. Restriction maps of the *hsp-3* and *hsp-6* genes from *C. elegans*. The genes are aligned to show maximum homology. The coding regions are shaded and the transcribed, untranslated regions are unshaded. Introns are shown by breaks in the coding region. B, *Bam* HI; E, *Eco* RI; H, *Hin*d III; K, *Kpn* I; P, *Pst* I; S, *Sal* I; *X, Xba* I.

dition signal (Fig. 2, block B) and an approximation of the transcriptional start site (see below). This is in good agreement with the 2.6-kb size predicted from Northern blots (Snutch *et al.*, 1988).

Within the long trailer sequence is the septemer TTTTTTC tandemly repeated three times (Fig. 2, block A). This septemer was searched for in other sequences. Several TTTTTTC sequences were found within the terminal arms of the *C. elegans* transposable element *Tc1* (Rosenzweig *et al.*, 1983) but not in the open reading frame. Similar copies of the septemer sequence were found tandemly repeated in the 3′ untranslated region of the *Trypanosoma cruzi* hsp83-like gene (Dragon *et al.*, 1987), the 3′ untranslated region of a mouse male germ line specific hsp70 gene (Zakeri *et al.*, 1988) and just 3′ of the *plasmodium falciparum* TRAP gene (Robson *et al.*, 1988).

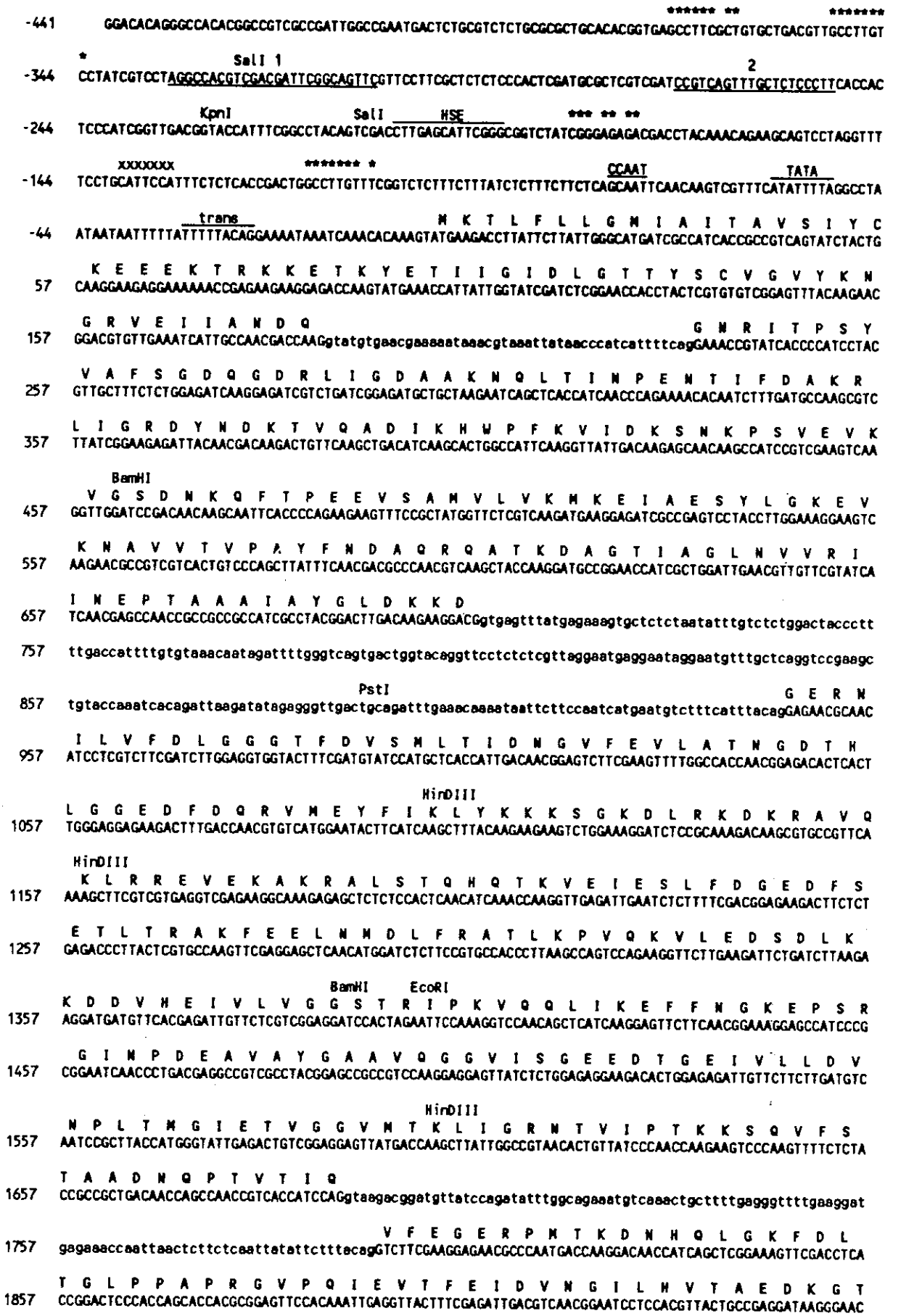
To identify the *hsp-3* gene product, a comparison was made between the predicted amino acid sequences of the *hsp-3* gene product (hsp70C) and hsp70A from *C. elegans* (Snutch *et al.*, 1988) and grp78 from rat (Munro and Pelham, 1986) (Fig. 3). This comparison revealed a striking degree of similarity with the rat grp78 amino acid sequence (77%). This is in contrast to the 59% identity shared with hsp70A. The hsp70C carboxyl terminus has the characteristic sequence KDEL found at the rat grp78 carboxyl terminus (Munro and Pelham, 1986). This sequence is required for retention of the protein in the ER (Munro and Pelham, 1987). Overall, the last 70 amino acids, excluding the last four, exhibit marginal conservation between hsp70C and grp78. This has been observed with other members of the hsp70 family. The hsp70C amino terminus is much longer when compared to hsp70A, but of similar length to the grp78 leader peptide (Fig. 3). The hsp70C putative signal peptide has features characteristic of secretory leader sequences. It contains a positively charged amino terminus followed by a hydrophobic central region followed by two small nonpolar residues with a single residue between. This is similar to the rat grp78 protein (Munro and Pelham, 1986). Cleavage of the hsp70C leader peptide probably occurs between I17 and Y18 after the two nonpolar residues. Although the structural features may be conserved between the hsp70C and grp78 proteins, the amino acid sequences of the leader peptides are not conserved (Fig. 3).

The size of the hsp70C protein before cleavage has been estimated to consist of 661 amino acids with a molecular weight of 73,339 daltons. After cleavage, the hsp70C protein would have a predicted molecular weight of 71,579 daltons. The hsp70C protein does not have any glycosylation sites (D-X-S/T) in agreement with the rat grp78 sequence (Munro and Pelham, 1986).

The high degree of similarity shared with the rat grp78 protein in both structural characteristics and amino acid sequence has led us to conclude that the *hsp-3* gene is probably the grp78 gene equivalent in *C. elegans*. Based on this, the *hsp-3* gene product is probably transported into and retained in the ER.

The hsp-3 *and rat grp78 5′ regulatory regions contain conserved sequences*

The *hsp-3* 5′ regulatory region was searched for sequences known to act as regulatory elements in other genes. We detected one copy of a heat shock element (HSE) (Pelham, 1982) from −205 to −195 (Fig. 2). Several sequences similar to the enhancer core sequences of E1A (Hearing and Shenk, 1983) and SV40 (Weiher *et al.*, 1983) are also found in the 5′ regulatory region of the hsp70C gene (Fig. 2). Similarities to E1A and SV40 core enhancer sequences are also


-441
GGACACAGGGCCACACGGCCGTCGCCGATTGGCCGAATGACTCTGCGTCTCTGCGCGCTGCACACGGTGAGCCTTCGCTGTGCTGACGTTGCCTTGT
SalI 1
2
-344
CCTATCGTCCTAGGCCACGTCGACGATTCGGCAGTTCGTTCCTTCGCTCTCTCCCACTCGATGCGCTCGTCGATCCGTCAGTTTGCTCTCCCTTCACCAC
KpnI
SalI
HSE
-244
TCCCATCGGTTGACGGTACCATTTCGGCCTACAGTCGACCTTGAGCATTCGGGCGGTCTATCGGGAGAGACGACCTACAAACAGAAGCAGTCCTAGGTTT
-144
TCCTGCATTCCATTTCTCTCACCGACTGGCCTTGTTTCGGTCTCTTTCTTTATCTCTTTCTTCTCAGCAATTCAACAAGTCGTTTCATATTTTAGGCCTA
CCAAT
TATA
trans
-44
ATAATAATTTTTATTTTTACAGGAAAATAAATCAAACACAAAGTATGAAGACCTTATTCTTATTGGGCATGATCGCCATCACCGCCGTCAGTATCTACTG
57
CAAGGAAGAGGAAAAAACCGAGAAGAAGGAGACCAAGTATGAAACCATTATTGGTATCGATCTCGGAACCACCTACTCGTGTGTCGGAGTTTACAAGAAC
157
GGACGTGTTGAAATCATTGCCAACGACCAAGgtatgtgaacgaaaaataaacgtaaattataacccatcattttcagGAAACCGTATCACCCCATCCTAC
257
GTTGCTTTCTCTGGAGATCAAGGAGATCGTCTGATCGGAGATGCTGCTAAGAATCAGCTCACCATCAACCCAGAAAACACAATCTTTGATGCCAAGCGTC
357
TTATCGGAAGAGATTACAACGACAAGACTGTTCAAGCTGACATCAAGCACTGGCCATTCAAGGTTATTGACAAGAGCAACAAGCCATCCGTCGAAGTCAA
BamHI
457
GGTTGGATCCGACAACAAGCAATTCACCCCAGAAGAAGTTTCCGCTATGGTTCTCGTCAAGATGAAGGAGATCGCCGAGTCCTACCTTGGAAAGGAAGTC
557
AAGAACGCCGTCGTCACTGTCCCAGCTTATTTCAACGACGCCCAACGTCAAGCTACCAAGGATGCCGGAACCATCGCTGGATTGAACGTTGTTCGTATCA
657
TCAACGAGCCAACCGCCGCCGCCATCGCCTACGGACTTGACAAGAAGGACGgtgagtttatgagaaagtgctctctaatatttgtctctggactaccctt
757
ttgaccattttgtgtaaacaatagattttgggtcagtgactggtacaggttcctctctcgttaggaatgaggaataggaatgtttgctcaggtccgaagc
PstI
857
tgtaccaaatcacagattaagatatagagggttgactgcagatttgaaacaaaataattcttccaatcatgaatgtctttcatttacagGAGAACGCAAC
957
ATCCTCGTCTTCGATCTTGGAGGTGGTACTTTCGATGTATCCATGCTCACCATTGACAACGGAGTCTTCGAAGTTTTGGCCACCAACGGAGACACTCACT
HinDIII
1057
TGGGAGGAGAAGACTTTGACCAACGTGTCATGGAATACTTCATCAAGCTTTACAAGAAGAAGTCTGGAAAGGATCTCCGCAAAGACAAGCGTGCCGTTCA
HinDIII
1157
AAAGCTTCGTCGTGAGGTCGAGAAGGCAAAGAGAGCTCTCTCCACTCAACATCAAACCAAGGTTGAGATTGAATCTCTTTTCGACGGAGAAGACTTCTCT
1257
GAGACCCTTACTCGTGCCAAGTTCGAGGAGCTCAACATGGATCTCTTCCGTGCCACCCTTAAGCCAGTCCAGAAGGTTCTTGAAGATTCTGATCTTAAGA
BamHI
EcoRI
1357
AGGATGATGTTCACGAGATTGTTCTCGTCGGAGGATCCACTAGAATTCCAAAGGTCCAACAGCTCATCAAGGAGTTCTTCAACGGAAAGGAGCCATCCCG
1457
CGGAATCAACCCTGACGAGGCCGTCGCCTACGGAGCCGCCGTCCAAGGAGGAGTTATCTCTGGAGAGGAAGACACTGGAGAGATTGTTCTTCTTGATGTC
HinDIII
1557
AATCCGCTTACCATGGGTATTGAGACTGTCGGAGGAGTTATGACCAAGCTTATTGGCCGTAACACTGTTATCCCAACCAAGAAGTCCCAAGTTTTCTCTA
1657
CCGCCGCTGACAACCAGCCAACCGTCACCATCCAGgtaagacggatgttatccagatatttggcagaaatgtcaaactgcttttgagggttttgaaggat
1757
gagaaaccaattaactcttctcaattatattctttacagGTCTTCGAAGGAGAACGCCCAATGACCAAGGACAACCATCAGCTCGGAAAGTTCGACCTCA
1857
CCGGACTCCCACCAGCACCACGCGGAGTTCCACAAATTGAGGTTACTTTCGAGATTGACGTCAACGGAATCCTCCACGTTACTGCCGAGGATAAGGGAAC

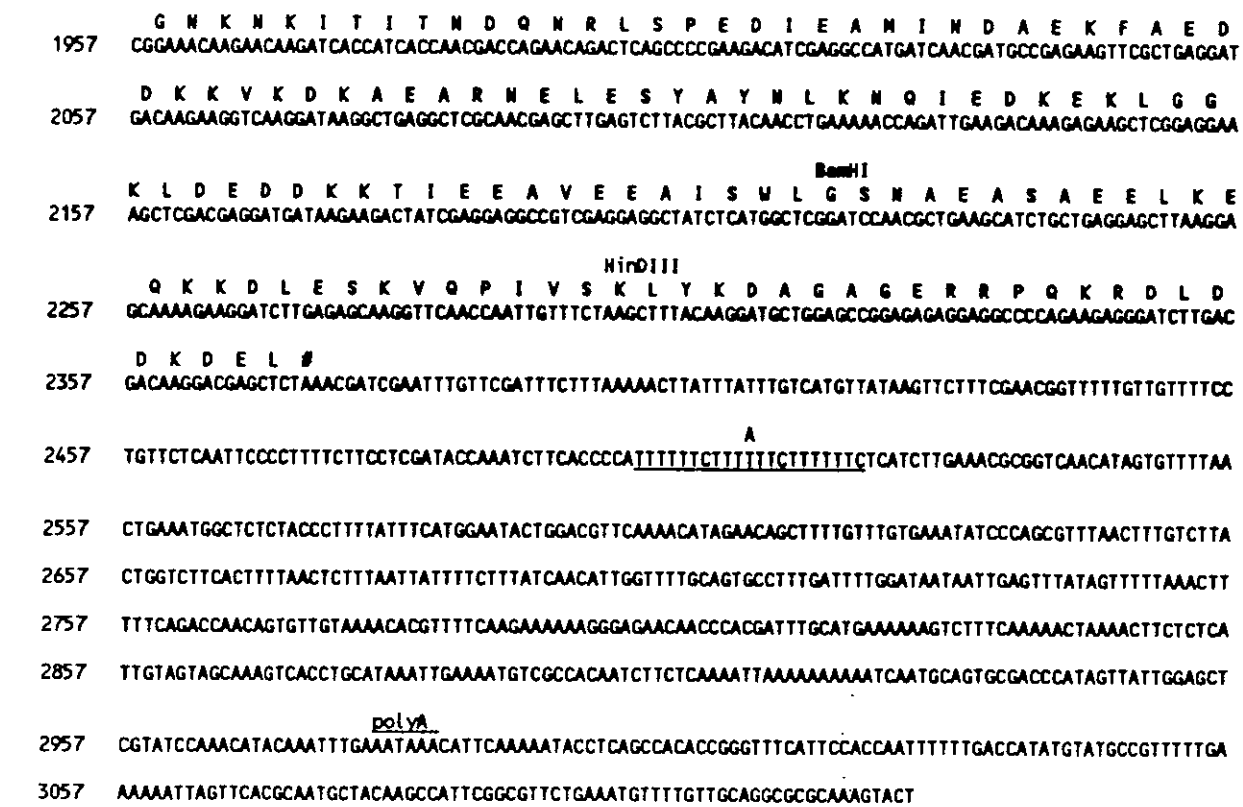


FIG. 2. Nucleotide sequence of the *hsp-3* gene. Introns are shown in lower-case letters. Numbering is with respect to the start of translation. The amino acid sequence is shown above the nucleotide sequence. Sequences homologous to the E1A core enhancer sequence ($^{A}/_{C}GGAAGTG^{A}/_{C}$; Hearing and Shenk, 1983) are indicated with an asterisk and the SV40 core enhancer sequence ($GTGG^{A}/_{T}{}^{A}/_{T}{}^{A}/_{T}G$; Weiher *et al.*, 1983) is indicated with an x. The HSE is shown as are the putative CCAAT and TATA boxes. The 3′ splice site at −23, which could serve as a *trans* splice site (Krause and Hirsh, 1987), is indicated. Block A represents the three-time repeated septemer sequence TTTTTTC. The poly(A) addition signal is also shown. Blocks 1 and 2 represent sequences found in the rat grp78 regulatory region (see Fig. 4).

FIG. 3. Comparison of the amino acid sequences of hsp70C, hsp70A (Snutch *et al.*, 1988), and the rat grp78 (Munro and Pelham, 1986). Amino acid matches are indicated by a dot (.) and amino acid gaps are shown with a dash (-).

observed in the rat grp78 regulatory region (Lin *et al.*, 1986). There appears to be a TATA box from −58 to −51 and a CCAAT box from −78 to −74 (Fig. 2).

It has been reported that the 5′ regulatory region of the rat grp78 gene has an enhancer-like activity (Lin *et al.*, 1986). We reasoned that if the HSEs are highly conserved between evolutionarily distant species (Pelham, 1982, 1985) and if the grp78 enhancer-like activity is important for the expression of the grp78 genes, then this element should be conserved between *Caenorhabditis* and rat. The grp78 regulatory regions from *C. elegans* and rat exhibit some conservation of nucleotide sequence. In addition to identities to the viral enhancers, there is a 23-nucleotide block in which 18 nucleotides are shared between the *hsp-3* and rat grp78 regulatory regions (Fig. 4). This conserved element in the rat grp78 gene is located within the restriction fragment reported to contain an enhancer-like activity (Lin *et al.*, 1986; Chang *et al.*, 1987). This 23-nucleotide sequence was compared to other hsp gene 5′ regulatory regions to determine if the sequence was a stress general signal or specific for the grp78-related genes. A search involving the *C. elegans hsp-1* (Snutch *et al.*, 1988), *hsp-6* (this report), and *hsp-16* genes (Russnack and Candido, 1985; Jones *et al.*, 1986), a *Drosophila* hsp70 (Ingolia *et al.*, 1980), and a human hsp70 gene (Hunt and Morimoto, 1985) did not reveal any significant homologies to the 23-nucleotide sequence. Therefore, we conclude that this 23-nucleotide sequence is specific for the grp78 genes.

Characterization of the hsp-6 *gene*

A restriction map of the 2.9-kb *Eco* RI restriction fragment containing the *hsp-6* gene is shown in Fig. 1. The region of the *hsp-6* gene sequenced represents the first two-thirds of an hsp70 gene based on sequence comparison to the *hsp-1* gene (Snutch *et al.*, 1988) and the *hsp-3* gene (Fig. 1). The sequence of the partial *hsp-6* gene fragment is presented in Fig. 5. Unfortunately, several screenings of several different genomic libraries did not result in the isolation of the full-length *hsp-6* gene. Within the *hsp-6* coding region thewre are two putative introns of 66 and 49 nucleotides. These two introns are not in the same position as the introns from *hsp-1* and *hsp-3*.

In the 5′-flanking region is an HSE (Pelham, 1982, 1985) from −316 to −303 (Fig. 5). This HSE is also part of a two palindromic 10-bp sequences NTTCNNGAAN required for full heat inducibility of the heat shock-inducible genes (Xiao and Lis, 1988). The HSE is flanked by the septemer TTTTTTC. There are several more copies of this sequence as well as several degenerate copies upstream from the coding region. Downstream from the HSE is a region centered around −278 to −272 that could function as

```
                          1
hsp70C   -330  TCGTCCTAGGCCACGTCGACGATTCGGCAGTT CGTTCCTTCGC   -288
               ||  ||||| | ||||  || ||||||  |  | || |
grp78    -180  GCGGAGGAGGCCGCTTCGA AT CGGCAGCGGCCAGCGTTGGT   -140
```

FIG. 4. A comparison of the regulatory regions of the *hsp-3* gene and the rat grp78 gene (Lin *et al.*, 1986). Numbering of *hsp-3* is with respect to the start of translation. Numbering of the rat grp78 gene is with respect to the start of transcription (Lin *et al.*, 1986). Block 1 represents the stretch of nucleotide identity shared between the *hsp-3* regulatory region and the rat grp78 regulatory region, a region which has demonstrated enhancer-like properties (Lin *et al.*, 1986).

```
          EcoRI
-1307  GAATTCTTGCATCAATTCCCGTCGTTTTGCAATAAGTTCCGACGCGGAAAGATGCGAAAAATCTTCCCGATGTGGCTCATCCTCTTCTCCTTCTTCTTCT

-1207  TCATCACCGCTTTTTTCTTCTTCGTCTTCCTTAACTTCTTTTTCTCGTGTCGGACGAATCAGTGTTCCATCTTTTAGCTTAATCGGCAGAAGTTCTTCAT

-1107  GGCCTTCTTCGAGTAGTCCTGCGAATTTTCTCTTTTTCTTCTCAACATCATCTTCATCGTCCTGTCTGATTCTACACCATTCTTTCTGCAATTTAACGTT

-1007  ATTAAGTTTTATTAACTAATTTTCAAACAACAATTACTTTTCTCAAGCCGCCTTTTTGCGTTGGCAAATGCGCTGTGTCCCTTCGATGTCGTCGTCCATC

 -907  ATGTCCAATGCGAAACTTCATCACTATTCTCATAATGAGTCTTGTCTGTCGGCGATGAAATCATCTTCAATCTGAAAATTATTGGATAAATTGTTTTATT

                                    PstI
 -807  TTCGAAAAATTGCTCACATTAACAGCTCTCTGCCTCAAATCTGCAGATTTTCTTGATTTAATATCTGCCATTCCATCGCGAATCTCTCTCGACAGCTTTT

                                                                                    BamHI
 -707  TCCGCTTTGCAAGGCGATTTAAATGTTTCGTAGTTTTTCTTGTTTTATTTGTCTTCAGCATTTTCAGCTTCTCTTCGCGAGAAGCGGAGACAGGATCCCA

 -607  TGTTTTTCTGCAAAATAATCCATTGATTTAACACCTCGTAAATAATTTAAAAAAGAGTTAAATTTAATTGCAACCCTATTTGTAAAAAGAAAACTCATTT

 -507  TCGCCAAAAATAAAGCAAAAATAATTCAAGAGAAAAACGCGCCGGTGTTGCGATTGGGGCGTAACTGCAATGTGTGCGCACACAATCTCAACAAGCGCTG

 -407  CGAGACCCGCCGCCTGACCGTAATGTGAAATGGGCGGAGACGAGAAGTTTTTTTCTGTTTGAAAGTTGATGCAAAAGCCCGTGATTCTTTTTTTCGAGAA

       HSE                    TATA                  v         v
 -307  ATTTCTCGAGTTTTTTCCAACGAAAAATTCATTAAATTTAAACCTTTTAGCTCTCCTTTCCAATCTTTTGCATCATTATTCTCCTAAACTTGGCATATTC

 -207  AGTGGAAATGATGCAAAATGACCCTGACTTTTGTTATCAAAAATAACAAGAAAATTGTCCCGTTTAACGGTTGAAAAGCAAATTTTGTGTCATTTTGTTT

                                                                                    trans
 -107  AGGAAATGTCAAAATAAGCTCAAAAACCGATTACAAATTATATTTTACTGCTTTTTATCCTATTTTCTCGCGTTTTCGTTCATGAATGCAATTTTCTTTC
```

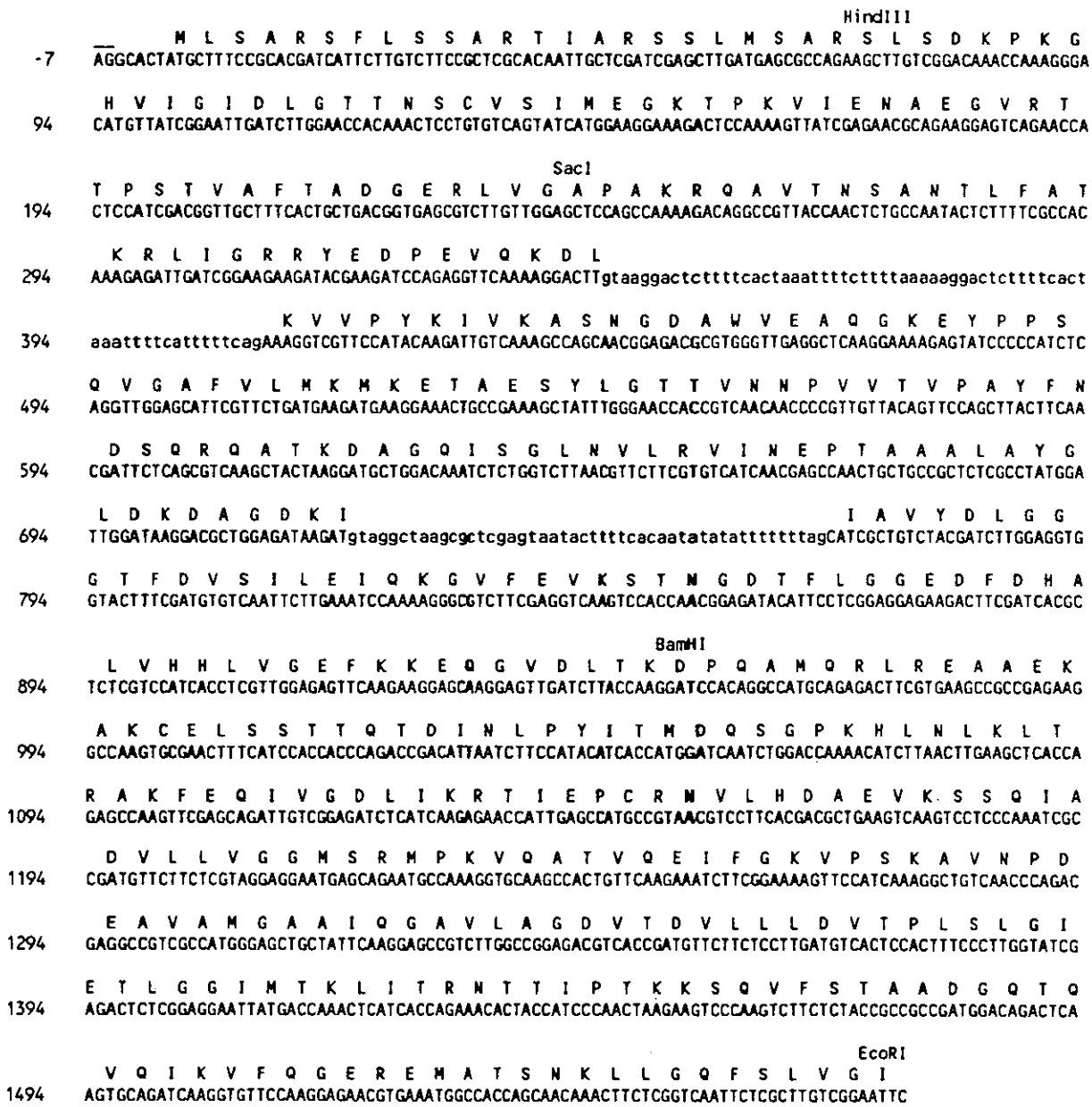


FIG. 5. Nucleotide sequence of the first two-thirds of the *hsp-6* gene. Introns are shown in lower-case letters. Numbering is with respect to the start of translation. A putative TATA box is shown as well as potential transcription start sites (v) and potential *trans* splice site. A dimer of the HSE (Xiao and Lis, 1988) is indicated. The end of the sequence presented here does not represent the most 3′ end of the *hsp-6* gene (see text for explanation).

a TATA box. Approximately 30 nucleotides downstream from this region is a sequence that shares identity with a transcription initiation/cap site sequence (derived from an analysis of the *C. elegans* major sperm protein gene family, CATAATCTTCA where A is the probable site of transcription initiation; Klass *et al.*, 1988) from −246 to −236 and −216 to −206. There is a 3′ splice site from −16 to −6 which is a feature of the *trans* splicing event (Krause and Hirsh, 1987).

To determine the identity of the *hsp-6* gene, a comparison of the hsp70F amino acid sequence to several other known hsp70 amino acid sequences was made (Fig. 6). hsp70F shares limited amino acid identity with hsp70A (53%; Snutch *et al.*, 1988) and hsp70C (53%) and the *Hsc1* (48%), *Hsc2* (43%; Craig *et al.*, 1983), and heat inducible hsp70 (52%; Ingolia *et al.*, 1980) gene products from *D. melanogaster*. Surprisingly, a comparison of hsp70F to the *dnaK* protein (Bardwell and Craig, 1984) revealed an identity of 67%. The hsp70 proteins listed above, over the same region covered by hsp70F, share approximately 48–54% identity at the amino acid level with the *dnaK* protein and 61–86% identity at the amino acid level amongst themselves. Therefore, we conclude that the *hsp-6* gene is more like the *dnaK* gene than any eukaryotic hsp70 gene identified to date.

hsp70F has a 29-amino-acid leader sequence when com-

```
hsp70F   MLSARSFLSSARTIARSSLMSARSLSDKPKGHVIGIDLGTTNSCVSI-MEGKTPKVIENAEGVRTTPSTVAFTAD-GE   76
dnaK                                  M.KI.....A.-.D.T..R.L.....D.....II.Y.Q.-..   47
hsp70A                               MS.HNAV.....Y...GVF.H..VE-I.A.DQ.N.....Y....-.-T.   48
hsp70C   MKTLFLLGMIAITAVSIYCKEEEKTRKKETKYETI.....Y...GVYKN.RVE-I.A.DQ.N.I...Y...SG.Q.D   79
hsc1                                  M..LPAV.....Y...GVFQH..VE-I.A.DQ.N.....Y....-E-S.   48
hsc2                                  MG.IPA.....Y...GV-WQNSKVEI.A.DQ.N.....Y...N-E-T.   48

hsp70F   RLVGAPAKRQAVTNSANTLFATKRLIGRRYEDPEVQKDLKVVPYKIV-KASNGD-AW-VEAQ--G--KEYPPSQVGAFVL  149
dnaK     T...Q.....PQ.....I.....FQ.E...R.VSIM.F..I-A.D...-..-..VK--.-Q.MA..-.IS.E..  120
hsp70A   ..I.DA..N.VAM.PH..V.DA.....KFD..A..S.M.HW.F.VI-S.EGAK-PK-.QVEYK.EN.IFT.EEISSM..  125
hsp70C   ..L.DA..N.LTI.PE..I.DA.....D.N.KT..A.I.HW.F.VID.SNKPSVEVK.GSD--N--.QFT.EE.S.M..  155
hsc1     ..I.DA..N.VAM.PN..I.DA.....FD.AT..S.M.HW.FEAF--.G..K-PR-IR                       104
hsc2     ..I.D...N.VAM.AK..V.DA.....KFD..KI.E...LW.F.VI--NEK.K-PK-I.                       104

hsp70F   MKMKETAESYLGTTVNNPVVTVPAYFNDSQRQATKDAGQISGLNVLRVINEPTAAALAYGLDKDA-GDKIIAVYDLGGGT  228
dnaK     K...K...D...EP.TEA.I.....A.....R.A..E.K.I.....GT-.NRT.....  199
hsp70A   L...K...AF.EP..KDA.....T.....A.A.....I.....I.....KGH.ERNVLIF.....  205
hsp70C   V....I.....KE.K.A.....A.....T.A....V.I.....I.....KD-.ERN.L.F.....  234

hsp70F   FDVSILEI--QKG--VFEVKSTNGDTFLGGEDFDHALVHHLVGEFKKEQGVDLTKDPQAMQRLREAAEKAKCELSSTTQT  304
dnaK     ..I..I..DEVD.EKT...LA....H.....SR.INY..E....D..I..RN..L.....K.....I....AQ..  279
hsp70A   .....T.--ED.--I.....A...H.....NRM.N.FCA...RKHKK..ASN.R.LR...T.C.R.NET...SC.A  281
hsp70C   ....M.T.--DN.--....LA.....H.....QRVMEYFIKLY..KS.K..R..KR.V.K..REV....RA..TQH..  310

hsp70F   DINLPYITMDQSGPKHLNLKLTRAKFEQIVGDLIKRTIEPCRNVLHDAEVKSSQIADVLLVGGMSRMPKVQATVQEIF-G  384
dnaK     .V.....A.AT....M.I.V....L.SL.E..VN.S...LKVA.Q..GLSV.D.D..I....QT...M..KK.A.F.-.  359
hsp70A   S.EID-SLFE--.IDFYT-NI...R..ELCA..FRS.MD.VEKS.R..KMDK..VH.IV....ST.I....KLLSDL.S.  357
hsp70C   KVEIE-SLF.--.ED-FSET.....ELNM..FRA.LK.VQK..E.SDL.KDDVHEIV....ST.I....QLIK.F.N.  386

hsp70F   KVPSKAVNPDEAVAMGAAIQGAVLAGD----V-T-DVLLLDVTPLSLGIETLGGIMTKLITRNTTIPTKKSQVSFTAADG  457
dnaK     .E.R.D.....I...V..G..T..-----.-K-.....M..V..T..AK.....H.....E.N  432
hsp70A   .ELN.SI.....L.Y...V.A.I.S..KSEA.-Q-.L.....A.....A..V..A..K.....TA.T.T.YS.N  435
hsp70C   .E..RGI.....Y...V..G.IS.E----ED.GEIV....N..TM....V..V.....G...V.....N  462

hsp70F   QTQVQIKVFQGEREMATSNKLLGQFSLVGI.....  487
dnaK     .SA.T.H.L....KR.AD..S....N.D.....  462
hsp70A   .PG.L.Q.YE...A.TKD.N...K.E.S.....  465
hsp70C   .PT.T.Q..E...P.TKD.HQ..K.D.T.L.....  492
```

FIG. 6. Comparison of the hsp70F amino acid sequence to the protein products of *dnaK*, *hsp-1* (hsp70A), *hsp-3* (hsp70C), *Hsc1*, *Hsc2*, and hsp70. Matches are indicated with a dot (.) and gaps with a dash(-). The amino acid sequences presented here do not represent the complete amino acid sequences for any of the proteins listed above (see text for explanation).

pared to the *dnaK* protein (Fig. 6). The hsp70F leader sequence is composed primarily of uncharged amino acids with a few hydrophobic and basic amino acids but no acidic amino acids. Within this 29-amino-acid leader sequence, 10 of the residues are serine and threonine. This sequence would not be as hydrophobic as the hsp70C leader sequence since there are few hydrophobic residues. Instead, the hsp70F leader sequence is quite similar to a mitochondrial matrix import leader sequence. These sequences are characterized by their lack of acidic amino acids and the presence of basic amino acids as well as extensive stretches of uncharged amino acids and a high content of serine and threonine residues (van Loon *et al.*, 1986; Colman and Robinson, 1986). Therefore, it seems likely that the hsp70F protein is transported into the mitochondria.

DISCUSSION

We have characterized two members of the hsp70 multigene family from *C. elegans*. One of these genes is a member of a grp78-like subfamily, *hsp-3*, encoding hsp70C; the other is the *hsp-6* gene encoding hsp70F. hsp70F shares a high degree of sequence similarity with the *E. coli* dnaK protein. The conservation of the amino acid sequence with *E. coli* is greater than the conservation seen with other eukaryotic hsp70s (67% *vs.* 43–53%). In addition, the predicted hsp70F protein sequence has a peptide leader characteristic of a mitochondrial import leader sequence. If the hsp70F protein is transported into the mitochondria this would explain the high degree of homology shared between hsp70F and *dnaK* since it is believed that mitochondria arose as a symbiotic relationship between bacteria and primitive

eukaryotic cells. Recently, it has been demonstrated that the *S. cerevisiae SSC1* protein is imported into the mitochondria and also shares a high degree of identity with the *dnaK* protein (J. Kramer and E.A. Craig, personal communication).

Unfortunately, we have been unable to isolate a full-length copy of the *hsp-6* gene even though several genomic libraries have been screened, including those constructed from partial *Eco* RI and *Mbo* II digests. This may be due to infrequent *Eco* RI or *Mbo* II sites at the 3′ end of the *hsp-6* gene, or the region may contain a sequence rendering the 3′ end of the *hsp-6* gene unclonable, or we have just had bad luck in our attempts to isolate the 3′ half of the *hsp-6* gene. We are attempting to isolate a cDNA corresponding to the *hsp-6* gene and then using this as a probe to screen a genomic library. Although the lack of a complete sequence for the *hsp-6* gene is disappointing, it does not detract from our suggestions that the *hsp-6* gene product hsp70F is transported into the mitochondria.

The *hsp-3* gene has one copy of an HSE in the 5′ regulatory region. The presence of an HSE is consistent with the observation that the mammalian grp78 genes are slightly heat inducible (Attenello and Lee, 1984; Lin *et al.*, 1986). This would suggest that the *hsp-3* gene should be heat inducible, although no increase in the mRNA concentration has been detected under the conditions tested (Snutch *et al.*, 1988). It may be that the *hsp-3* gene is transiently expressed for only a short time after heat shock. However, Xiao and Lis (1988) have shown that for full heat-inducibility of a gene, two repeats of the 10-mer NTTCNNGAAN must be present. If this is true for *C. elegans* as well, then the *hsp-3* gene would not be expected to respond significantly to a temperature shift.

The presence of the septemer sequence TTTTTTC in the 3′ untranslated trailer sequence of the *hsp-3* gene is intriguing. A sequence of similar nature is found in the 5′ regulatory region of the *hsp-6* gene (this report) and terminal arms of the *Tc1* transposable element (Rosenzweig *et al.*, 1983). A hexamer sequence of similar nature (TTTTTC) is tandemly repeated four times in the 3′ untranslated regions of the *T. cruzi* hsp83-like genes (Dragon *et al.*, 1987), an octamer sequence (TTTTTTTC) is tandemly repeated five times just 3′ of the TRAP gene from *P. falciparum* (Robson *et al.*, 1988) and degenerate copies of the septemer are repeated in the 3′ untranslated region of a mouse male germ line-specific hsp70 gene (Zakeri *et al.*, 1988). It is interesting to note that the hsp83-like gene and the TRAP gene encode proteins that are presented as antigens and that all three genes are expressed at specific stages of development. The *hsp-3* gene also appears to be developmentally regulated with its mRNA being most abundant at the L1 larval stage (Snutch *et al.*, 1988). Despite these observations, the significance of these conserved repeats is not known.

The C. elegans hsp-3 *5′ regulatory region shares identity with the rat grp78 regulatory region*

An analysis of the 5′-flanking sequence of the *hsp-3* gene revealed homologies to sequences known to function as enhancer sequences and HSEs. In addition, a 23-nucleotide element was detected in the rat grp78 5′ regulatory region. This element is conserved between the *hsp-3* gene of *C. elegans* and its homolog from *C. briggsae* (Heschl and Baillie, in preparation). The conserved element from the rat grp78 regulatory region is protected by a protein during nuclease footprinting studies (Resendez *et al.*, 1988). Because *C. elegans* and the rat are evolutionarily highly diverged and this sequence is so highly conserved, this 23-nucleotide element must be important for some aspect of regulation specific to the grp78 genes. Although an increase in *hsp-3* mRNA synthesis in response to glucose deprivation has not been demonstrated in the nematode *C. elegans*, based on the high degree of similarity between hsp70C and the glucose-regulated rat grp78 protein, this 23-nucleotide sequence could represent a regulatory element unique to the regulation of the grp genes.

The hsp70 multigene family of C. elegans

hsp70 multigene families consisting of eight or more members have been identified in *S. cerevisiae*, *Drosophila* (see Craig, 1985; Lindquist, 1986), and man (Mues *et al.*, 1986). *C. elegans* has at least nine members (Snutch *et al.*, 1988) two of which, the *hsp-1* and *hsp-2*ps genes, have been described elsewhere (Snutch *et al.*, 1988; Heschl and Baillie, 1989). Some of the *S. cerevisiae*, *Drosophila*, and human hsp70 genes have been sequenced, identified, and studied in detail (Craig, 1985; Lindquist, 1986), but only the unicellular eukaryote *S. cerevisiae* hsp70 multigene family has been analyzed genetically (Craig and Jacobsen, 1984, 1985; Craig *et al.*, 1987; Werner-Washburne *et al.*, 1987). *C. elegans* has a relatively complex developmental pathway but the amenability of this organism to genetic analyses makes it seem likely that a study of the hsp70 genes in *C. elegans* may uncover additional, valuable information about this gene family.

A comparison of the nucleotide sequences of the *C. elegans* hsp70 genes characterized to date to each other suggests that the *C. elegans* hsp70 genes can tentatively be assigned to gene subfamilies. The *hsp-1* gene potentially represents one gene family analogous to members of the SSA gene subfamily from yeast (Snutch *et al.*, 1988). The grp78-like subfamily consists of at least two members, the constitutive *hsp-3* gene and the highly heat-inducible *hsp-4* gene (unpublished results). The third subfamily has as its sole member the *hsp-6* gene, probably homologous to the yeast SSC subfamily (Craig *et al.*, 1987; J. Kramer and E.A. Craig, personal communication). A more precise assignment of the *C. elegans* hsp70 genes to specific subfamilies awaits the isolation of the remaining gene members of the family, complete isolation of genes partially characterized, and a genetic characterization of the hsp70 genes.

Mutations of the *hsp-3* gene may be difficult to isolate. There are two grp78-like proteins in *C. elegans*, *hsp-3* and the partially characterized, highly heat-inducible *hsp-4* (unpublished results). The *hsp-4* gene product, hsp70D, has the sequence HDEL at the carboxyl terminus instead of KDEL, like the yeast *KAR2* sequence (M. Rose, personal

communication). A null mutation in the *hsp-3* gene might be lethal like the disruption mutant of *KAR2*. However, if there is an interaction between the two grp78-like genes as seen in the SSA and SSB subfamilies of *S. cerevisiae* hsp70 genes (Craig and Jacobsen, 1984; Werner-Washburne *et al.*, 1987), then no lethal mutation could be recovered for the individual grp78-like genes. Alternatively, nonlethal, hypomorphic mutations of the *hsp-3* and *hsp-4* genes, possibly defective in protein secretion (for example, cuticle proteins), or temperature-sensitive mutations may be isolated.

In the *hsp-6* analog from *E. coli*, *dnaK*, only temperature-sensitive mutations have been isolated (Saito and Uchida, 1977; Itikawa and Ryu, 1979; Neidhardt *et al.*, 1984), suggesting that the *dnaK* gene product might be essential for cell growth. The *SSC1* gene, when disrupted, prevents growth (Craig *et al.*, 1987), suggesting that *SSC1* is essential for cell viability. If hsp70F is actively transported into the mitochondria, like the *SSC1* protein, then a mutation in *hsp-3* would most probably be an early embryonic lethal. Alternatively, hypomorphic mutations with a slow growth phenotype might also be recovered.

The genetic analysis of the *C. elegans* hsp70 genes will be enhanced greatly by the physical mapping of the hsp70 genes by cosmid assignment, *in situ* hybridizations, or the identification of restriction fragment length differences (RFLD). The *hsp-1* gene has been mapped to LGIV (right) with the aid of RFLD between two strains of *C. elegans* (Snutch *et al.*, 1988). The *hsp-3* gene has been localized to a cosmid contig of approximately 150 kb of DNA (as described in Coulson *et al.*, 1986). This contig has been mapped to the middle of LGII by *in situ* hybridization (D. Albertson, personal communication). Assignment of the *hsp-6* gene to cosmids and eventually a chromosomal region awaits either the isolation of phage DNA containing the respective genes and flanking DNA or the detection of RFLDs. Localization of the *C. elegans* hsp70-like genes will make it possible to generate and isolate mutants lacking a functional hsp70 gene. This would provide a convenient starting point for the genetic analysis of the heat shock response in the multicellular organism *C. elegans*.

ACKNOWLEDGMENTS

We would like to thank A. Rose, B. Honda, M. Smith, M. Werner-Washburne, and C. Nicolet for their comments and suggestions during the various drafts of this manuscript. This work was supported by a Medical Research Council of Canada Studentship and Simon Fraser University Research Fellowships to M.F.P.H. and a grant from Natural Sciences and Engineering Research Council of Canada to D.L.B.

REFERENCES

ATTENELLO, J.W., and LEE, A.S. (1984). Regulation of a hybrid gene by glucose and temperature in hamster fibroblast. Science **226**, 187-190.

BARDWELL, J.C.A., and CRAIG, E.A. (1984). Major heat shock gene of *Drosophila* and the *Escherichia coli* inducible *dnaK* gene are homologous. Proc. Natl. Acad. Sci. USA **81**, 848-852.

CHANG, S.C., WOODEN, S.K., NAKAKI, T., KIM, Y.K., LIN, A.Y., KUNG, L., ATTENELLO, J., and LEE, A.S. (1987). Rat gene encoding the 78kDa glucose regulated protein GRP78: Its regulatory sequences and the effect of protein glycosylation on its expression. Proc. Natl. Acad. Sci. USA **84**, 680-684.

CHEN, E.J., and SEEBURG, P.H. (1985). Supercoil sequencing: A fast and simple method of sequencing plasmid DNA. DNA **4**, 165-170.

CHIRICO, W.J., WATERS, M.G., and BLOBEL, G. (1988). 70K heat shock related proteins simulate protein translocation into microsomes. Nature **332**, 805-810.

COLMAN, A., and ROBINSON, C. (1986). Protein import into organelles: Hierarchial targeting signals. Cell **46**, 321-322.

COULSON, A., SULSTON, J., BRENNER, S., and KARN, J. (1986). Toward a physical map of the genome of the nematode *C. elegans*. Proc. Natl. Acad. Sci. USA **83**, 7821-7825.

CRAIG, E.A. (1985). The heat shock response. Crit. Rev. Biochem. **18**, 239-280.

CRAIG, E.A., and JACOBSEN, K. (1984). Mutations of the heat inducible 70 kilodalton genes of yeast confer temperature sensitive growth. Cell **38**, 841-849.

CRAIG, E.A., and JACOBSEN, K. (1985). Mutations in cognate genes of *Saccharomyces cerevisiae hsp70* result in reduced growth rates at low temperatures. Mol. Cell. Biol. **5**, 3517-3524.

CRAIG, E.A., INGOLIA, T.D., and MANSEAU, L.J. (1983). Expression of heat shock cognate genes during heat shock and development. Dev. Biol. **99**, 418-426.

CRAIG, E.A., KRAMER, J., and KOSIC-SMITHERS, J. (1987). *SSC1*, a member of the 70-kDa heat shock protein multigene family of *Saccharomyces cerevisiae*, is essential for growth. Proc. Natl. Acad. Sci. USA **84**, 4156-4160.

DESHAIES, R.J., KOCH, B.D., WERNER-WASHBURNE, M., CRAIG, E.A., and SCHEKMAN, R. (1988). A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. Nature **332**, 800-805.

DRAGON, E.A., SIAS, S.R., KATO, E.A., and GABE, J.D. (1987). The genome of *Trypanosoma cruzi* contains a constitutively expressed, tandemly arranged multicopy gene homologous to a major heat shock protein. Mol. Cell. Biol. **7**, 1271-1275.

HATTORI, M., and SAKAKI, Y. (1986). Dideoxy sequencing method using denatured templates. Anal. Biochem. **152**, 232-238.

HEARING, P., and SHENK, T. (1983). The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. Cell **33**, 695-703.

HENIKOFF, S. (1987). Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155**, 156-165.

HESCHL, M.F.P., and BAILLIE, D.L. (1989). Identification of a heat shock pseudogene from *Caenorhabditis elegans*. Genome (in press).

HUNT, C., and R.I. MORIMOTO. (1985). Conserved features of eukaryotic hsp70 genes revealed by comparison with the nucleotide sequence of human hsp70. Proc. Natl. Acad. Sci. USA **82**, 6455-6459.

INGOLIA, T.D., and CRAIG, E.A. (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.

INGOLIA, T.D., CRAIG, E.A., and McCARTHY, B.J. (1980).

Sequence of three copies of the gene for major *Drosophila* heat shock induced protein and their flanking regions. Cell **21,** 669–679.

ITIKAWA, H., and RYU, J.I. (1979). Isolation and characterization of a temperature-sensitive *dnaK* mutant of *Escherichia coli*. J. Bacteriol. **138,** 339–344.

JONES, D., RUSSNACK, R.H., KAY, R.J., and CANDIDO, E.P.M. (1986). Structure, expression, and evolution of a heat shock gene locus in *Caenorhabditis elegans* that is flanked by repetitive elements. J. Biol. Chem. **261,** 12006–12015.

KARN, J., BRENNER, S., and BARNETT, L. (1983). Protein structural domains in the *Caenorhabditis unc-54* myosin heavy chain gene are not separated by introns. Proc. Natl. Acad. Sci. USA **80,** 4253–4257.

KLASS, M., AMMONS, D., and WARD, S. (1988). Conservation in the 5′ flanking sequences of transcribed members of the *Caenorhabditis elegans* major sperm protein gene family. J. Mol. Biol. **199,** 15–22.

KRAUSE, M., and HIRSH, D. (1987). A *trans*-spliced leader sequence on actin mRNA in *C. elegans*. Cell **49,** 753–761.

LEE, A.S. (1987). Coordinated regulation of a set of genes by glucose and calcium ionophores in mammalian cells. Trends Biochem. Sci. **12,** 20–23.

LIN, A.Y., CHANG, S.C., and LEE, A.S. (1986). A calcium ionophore-inducible cellular promoter is highly active and has enhancerlike properties. Mol. Cell. Biol. **6,** 1235–1243.

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

MANIATIS, T., FRITSCH, E.F., and SAMBROOK, J. (1982). *Molecular Cloning, A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

MILARSKI, K.L., and MORIMOTO, R.I. (1986). Expression of human Hsp70 during the synthetic phase of the cell cycle. Proc. Natl. Acad. Sci. USA **83,** 9517–9521.

MUES, G.I., MUNN, T.Z., and RAESE, J.D. (1986). A human gene family with sequence homology to *Drosophila melanogaster* Hsp70 heat shock gene. J. Biol. Chem. **261,** 874–877.

MUNRO, S., and PELHAM, H.R.B. (1986). An hsp70-like protein in the endoplasmic reticulum identity with the 78 kD glucose-regulated protein and immunoglobulin heavy chain binding in protein. Cell **46,** 291–300.

MUNRO, S., and PELHAM, H.R.B. (1987). A C-terminal signal prevents secretion of luminal ER proteins. Cell **48,** 899–907.

NEIDHARDT, F.C., VANBOGELEN, R.A., and VAUGHN, V. (1984). The genetics and regulation of heat shock proteins. Annu. Rev. Genet. **18,** 295–329.

NORRANDER, J., KEMPE, T., and MESSING, J. (1983). Construction of improved M13-vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26,** 101–106.

PELHAM, H.R.B. (1982). A regulatory upstream promoter element in the *Drosophila* hsp70 heat-shock gene. Cell **30,** 517–528.

PELHAM, H.R.B. (1985). Activation of heat shock genes in eukaryotes. Trends Genet. **1,** 31–35.

RESENDEZ, JR., E., WOODEN, S.K., and LEE, A.S. (1988). Identification of highly conserved regulatory domains and protein binding sites in the promoters of the rat and human gene encoding the stress-inducible 78-kDa glucose-regulated protein. Mol. Cell. Biol. **8,** 4579–4584.

ROSENWEIG, B., LIAO, L.W., and HIRSH, D. (1983). Sequence of the *C. elegans* transposable element Tc1. Nucleic Acids Res. **11,** 4201–4209.

ROBSON, K.J.H., HALL, J.R.S., JENNINGS, M.W., HARRIS, T.J.R., MARSH, K., NEWBOLD, C.I., TATE, V.E., and WEATERALL, D.J. (1988). A highly conserved amino-acid sequence in thrombospondin, properdin and in proteins from sporozoites and blood stages of a human malaria parasite. Nature **335,** 79–82.

RUSSNACK, R.H., and CANDIDO, E.P.M. (1985). Locus encoding a family of small heat shock genes in *Caenorhabditis elegans*: Two genes duplicated to form a 3.8 kilobase inverted repeat. Mol. Cell. Biol. **5,** 1268–1278.

SAITO, H., and UCHIDA, H. (1977). Initiation of the DNA replication of bacteriophage lambda in *Escherichia coli* K12. J. Mol. Biol. **113,** 1–25.

SANGER, F., COULSON, A.R., BARELL, B.G., SMITH, A.J.H., and ROSE, B.A. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143,** 161–178.

SNUTCH, T.P. (1984). "Molecular and genetic analysis of the heat shock response of *Caenorhabditis elegans*." Ph.D. thesis. Simon Fraser University, Burnaby, British Columbia.

SNUTCH, T.P., and BAILLIE, D.L. (1983). Alterations in the pattern of gene expression following heat shock in the nematode *Caenorhabditis elegans*. Can. J. Biochem. Cell Biol. **61,** 480–487.

SNUTCH, T.P., HESCHL, M.F.P., and BAILLIE, D.L. (1988). The *Caenorhabditis elegans* hsp70 gene family: A molecular genetic characterization. Gene **64,** 241–255.

SPIETH, J., DENISON, K., ZUCKER, E., and BLUMENTHAL, T. (1985). The nucleotide sequence of a nematode vitellogenin gene. Nucleic Acids Res. **13,** 7129–7138.

TING, J., and LEE, A.S. (1988). Human gene encoding the 78,000-dalton glucose-regulated protein and its pseudogene: Structure, conservation, and regulation. DNA **7,** 275–286.

VANLOON, A.P.G.M., BRANDLI, A.W., and SCHATZ, G. (1986). The presequences of two imported mitochondrial proteins contain information for intracellular and intramitochondrial sorting. Cell **44,** 801–812.

WEIHER, H., KONIG, M., and GRUSS, P. (1983). Multiple point mutations affecting the simian virus 40 enhancer. Science **217,** 626–631.

WERNER-WASHBURNE, M., STONE, D.E., and CRAIG, E.A. (1987). Complex interactions among members of an essential subfamily of *hsp70* genes in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **7,** 2568–2577.

XIAO, H., and LIS, J.T. (1988). Germline transformation used to define key features of heat shock response elements. Science **239,** 1139–1142.

ZAKERI, Z., and WOLGEMUTH, D. (1987). Developmental stage-specific expression of the hsp70 family during the differentiation of the mammalian germ line. Mol. Cell. Biol. **7,** 1791–1796.

ZAKERI, Z.F., WOLGEMUTH, D.J., and HUNT, C.R. (1988). Identification and sequence analysis of a new member of the mouse hsp70 gene family and characterization of its unique cellular and developmental pattern of expression in the male germ line. Mol. Cell. Biol. **8,** 2925–2932.

Address reprint requests to:
Dr. Mark F.P. Heschl
Department of Physiological Chemistry
University of Wisconsin—Madison
1300 University Avenue
Madison, WI 53706

Received for publication October 4, 1988, and in revised form January 13, 1989.