Biochemical and Biophysical Research Communications **256,** 75–80 (1999)
Article ID bbrc.1999.0283, available online at http://www.idealibrary.com on IDEAL®

Genomic Cloning and Promoter Analysis of the Mouse 105-kDa Heat Shock Protein (HSP105) Gene

Kunihiko Yasuda, Keiichi Ishihara, Kazuo Nakashima, and Takumi Hatayama[1]
Department of Biochemistry, Kyoto Pharmaceutical University, Kyoto 607-8414, Japan

Received January 28, 1999

The 105-kDa heat shock protein (HSP105) is a member of the high-molecular-mass heat shock protein family. We have isolated and characterized the mouse HSP105 gene including about 1.2 kb of the 5′-flanking region. The mouse HSP105 gene spans about 22 kb, consisting of 18 exons separated by 17 introns. Southern blotting analysis revealed the existence of a single copy of HSP105. Primer extension analysis revealed that the transcription initiation site was located 165 bp upstream of the ATG translation initiation codon. The 5′-promoter region of the HSP105 gene contained a TATA box, a CAAT box, an inverted CAAT box, and two GC boxes. Two heat shock element (HSE) sequences were found as four nGAAn repeats at nt −64 and nt −128. Promoter analysis using deletion derivatives revealed that a minimal region which contained the two consensus HSE sequences was active in response to heat shock and also for constitutive expression of the gene. © 1999 Academic Press

Almost all organisms synthesize a group of proteins called heat shock proteins (HSPs) in response to various stresses; i.e., heat shock, heavy metals, ethanol, amino acid analogues, sodium arsenite and oxidative stress (1–3). The 105-kDa heat shock protein (HSP105α), which is expressed constitutively but is also induced by various stresses in mammalian cells, is a member of the high molecular mass HSP family (4, 5). Another 90-kDa HSP (42°C-HSP, HSP105β) that is only synthesized when mammalian cells are heated at 42°C, is similar to HSP105α (6–8). cDNA cloning of mouse and human HSP105α and HSP105β revealed that the former is composed of 858 amino acids and the latter is composed of 814 amino acids lacking 44 amino acids from HSP105α, probably produced by alternative splicing (9, 10). A transient increase in level of HSP105 during mouse embryogenesis is observed in most embryonic tissues, and HSP105 was suggested to play some important roles during mouse embryo development (11). Recently, in addition to HSP105, HSPs belonging to the high molecular mass HSP/HSP70 superfamily were found in organisms from yeast to mammals (12–19). However, among these members of this high molecular mass HSP family, the genomic structure of mammalian high-molecular-mass HSP has not been elucidated.

In this study, to elucidate the genomic structure of mouse HSP105 and to examine the regulation of expression of its gene, we cloned the mouse HSP105 gene and clarified its exon-intron structure as well as the promoter activity of the 5′-flanking region.

MATERIALS AND METHODS

Genomic cloning. Mouse genomic library (BALB/c genomic DNA partially digested with Sau3AI was cloned into EMBL3 phage vector at the *Bam*HI site) was supplied by Japan Cancer Research Resource Bank. The library (approximately 2×10^7 pfu) was screened by plaque hybridization using mouse HSP105α cDNA (9) fragments as probes. Phage DNA was extracted from positive plaques by the liquid lysate method (20), and cloned into pGEM7 (Promega) or pBK-CMV (Stratagene). A DNA fragment from exons 5 to 8 of the mouse HSP105 gene was amplified by PCR using mouse FM3A genomic DNA with a sense primer (18-mer, nt 484 to 501 of HSP105α cDNA) and an antisense primer (18-mer, nt 1059 to 1076 of HSP105α cDNA), then the PCR product was subcloned into pCRII (Invitrogene) and designated as UDR-1.5. DNA was subsequently sequenced by the dideoxy termination method using a Sequenase 7-deaza-dGTP DNA sequencing kit (Amersham). DNA sequence analysis was performed using GeneWorks software (IntelliGenetics) and the computer program FASTA at the EMBL data library (Heidelberg).

Determination of transcription start site. Total RNA was extracted by the guanidinium isothiocyanate method (21) from mouse FM3A cells heat-shocked at 42°C for 3 h, and poly(A)+RNA was further purified with oligotex-dT30 (Nippon Roche) according to the manufacturer's instructions. The poly(A)+RNA (8 μg) prepared from heat-shocked cells was hybridized to 5′-end-labeled primer (18-mer, nt 74–91 of mouse HSP105 cDNA) and extended with AMV reverse transcriptase. Subcloned plasmid DNA containing the genomic HSP105 DNA up to nt −1200 was sequenced with the same primer by the dideoxynucleotide method.

Southern blotting analysis. The genomic DNA (20 μg) extracted from mouse FM3A cells was digested with several different re-

[1] To whom correspondence should be addressed at Department of Biochemistry, Kyoto Pharmaceutical University, 5 Nakauchi-cho, Misasagi, Yamashina-ku, Kyoto 607-8414, Japan. Fax: 075-595-4758. E-mail: hatayama@mb.kyoto-phu.ac.jp.

0006-291X/99 $30.00

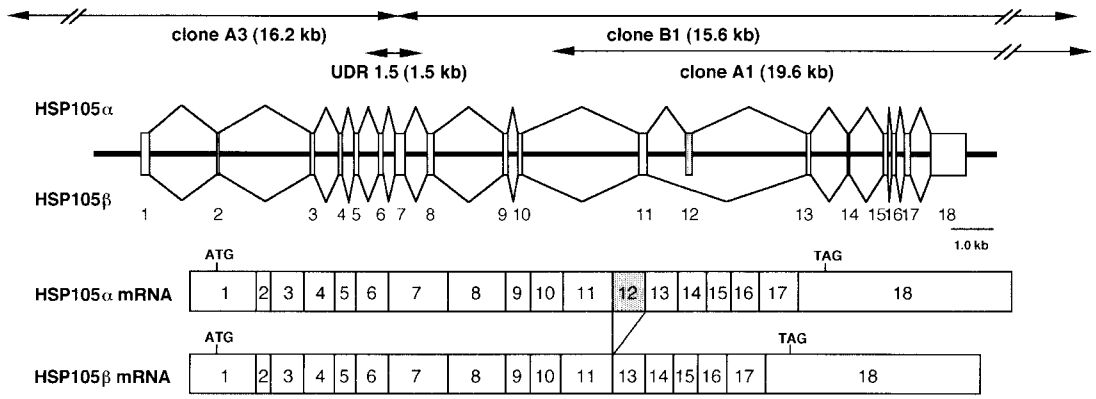

FIG. 1. Genomic organization of the mouse HSP105 gene. The linear map of the exon-intron structure is shown schematically. Exons are represented as numbered boxes. Positions of the three phage clones (clones A3, A1, and B1) and one plasmid clone (UDR-1.5) obtained by PCR are presented by arrows. Two alternative splicing patterns gave rise to HSP105α and HSP105β transcripts. ATG and TAG indicate the positions of initiation and termination codons, respectively. The DNA sequences of these clones have been submitted to the DDBJ/EMBL/GenBank DNA databases with Accession Nos. AB005267–AB005282.

striction enzymes, and the DNA was resolved by electrophoresis through 1% agarose gels, then transferred onto nylon membranes. The blots were hybridized with a probe consisting of the approximately 1.4 kb *Bam*Hl–*Hin*dIII fragment containing intron 17 and exon 18 of the HSP105 gene, according to the standard procedure (20).

Plasmid construction for promoter analysis. To examine the mouse HSP105 gene promoter activity, the plasmid pCAT105, in which a fragment of about 1.2 kb of the 5′-flanking region of the HSP105 gene was inserted into the *Kpn*I and *Sma*I sites of the pCAT3 basic vector (Promega), was constructed and used to construct the test plasmid for 5′-deletion series. The *Kpn*I–*Bgl*II fragment of the resulting construct was replaced with a PCR fragment prepared as described below. A DNA fragment containing the region spanning nt −290 through nt +73 was amplified using specific primers, i.e., forward primer 5′-GGGGTACCACTGGAGTCAACTGGAACGA-3′ (nt −290∼−271 of HSP105 gene) (underline indicates KpnI site), and reverse primer 5′-CTTACTTAGATCGCAGATCTCGAG-3′ (nt 32∼55 of pCAT3 basic vector) (underline indicates *Bgl*II site), using pCAT105 DNA as a template. The PCR product was digested with *Kpn*I and *Bgl*II and introduced into the *Kpn*I–*Bgl*II site of the pCAT3 basic vector to generate construct −290CAT. The constructs −133CAT, −85CAT and −42CAT were generated with the forward primer 5′-GGGGTACCGAATCTTTCTCGAAGGCTCG-3′ (nt −133∼−114 of HSP105 gene), 5′-GGGGTACCAGGCGCCGGTGAGTAAAATG-3′ (nt −85∼−66 of HSP105 gene) and 5′-GGGGTACCTCAGTAGGCTACATAAGGCT-3′ (nt −42∼−23 of HSP105 gene) (underline indicates *Kpn*I site) and the same reverse primer described above. All constructs generated by PCR were verified by sequencing.

TABLE 1

Exon–Intron Junction of the Mouse HSP105 Gene

Exon No.	Position[a]	Exon size: bp	Donor	Intron size: kb	Acceptor
1	1–161	272	CCCCgtaa	2.2	ttagGTCA
2	162–218	58	CCAGgtac	1.7	gcagCAAA
3	219–360	141	AAAGgtaa	1.5	gtagGTCA
4	361–483	123	CTCAgtaa	0.6	gcagGTCC
5	484–583	100	GCTGgtaa	0.25	atagTTGC
6	584–716	134	GAAGgtaa	0.3	ttagGTTC
7	717–962	245	ACAGgtat	0.5	tcagGTCA
8	963–1191	229	GCAGgtgt	1.6	caagTGTG
9	1192–1298	107	AAGGgtat	0.3	ttagTGTG
10	1299–1432	134	ATAGgtaa	2.35	ctagGCCG
11	1433–1641	209	GGATgtaa	1.3	tcagAAAA
12	1642–1773	132	CAAAgtga	2.8	taagGCAA
13	1774–1911	138	AGAGgtaa	0.75	acagGGCA
14	1912–2037	126	GCAGgtac	0.8	aaagGAAC
15	2038–2145	108	GATGgtaa	0.1	ttagAAAA
16	2146–2265	120	CAAGgtgg	0.2	ccagGATG
17	2266–2427	162	GAAGgtga	0.7	ttagGAAT
18	2428–3294	867			
		Consensus	$^{c}/_{A}$AG**gt**$^{a}/_{g}$a		$^{c}/_{t}$**ag**G

Note. For exon and intron sequences, capital and small letters are used, respectively.
[a] Nucleotide position number in HSP105α cDNA.

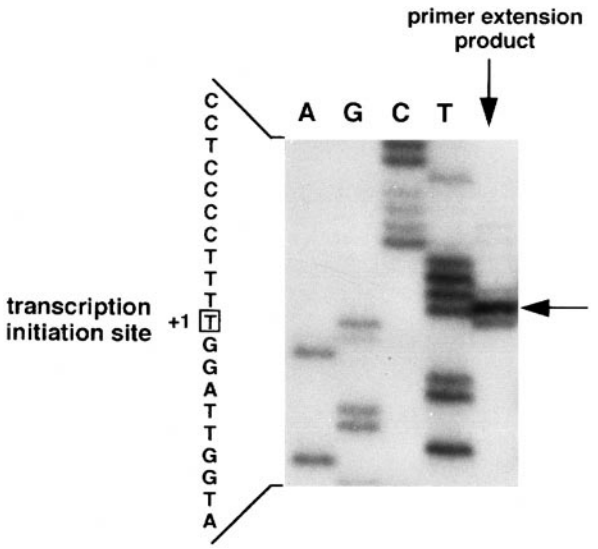

FIG. 2. Primer extension analysis of HSP105 mRNA. The major primer extension product is shown by an arrow. The DNA sequence of the 5′-promoter region of the HSP105 gene is shown on the left, and the nucleotide corresponding to the primer extension product is boxed.

RESULTS AND DISCUSSION

Structure of the HSP105 gene. We screened a mouse BALB/c genomic library using mouse HSP105α cDNA fragments as probes and isolated 3 clones, A1, A3, and B1. These clones included the entire HSP105 gene and spanned over 22 kb of the mouse genomic sequence. The genomic organization of the gene is schematically shown in Fig. 1. Clone A3 contained exons 1-7, clone B1 contained exons 7-18, and clone A1 overlapped with B1 clone and contained exons 11–18. Since no overlapping clone between A3 and B1 was obtained despite extensive screening of the library, we amplified the DNA sequence between exons 5 and 8 using mouse FM3A genomic DNA by PCR. DNA sequence analysis revealed that exon 7 was located over both clones A3 and B1 separated by a Sau3AI site. The sequences at exon–intron boundaries completely conformed to the donor/acceptor splicing rule (24); introns began with GT and ended with AG, and extended nucleotides around the junctions were consistent with consensus motifs (Table 1). Most exons were less than 200 bp in length, consistent with the observation that

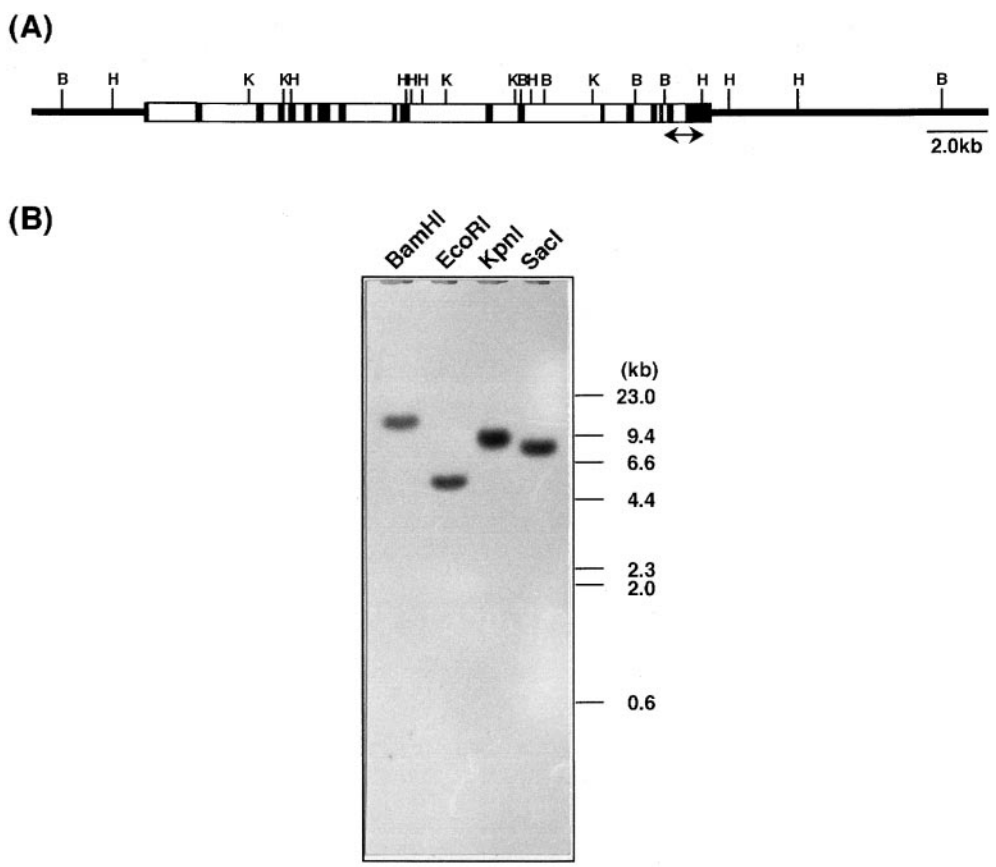

FIG. 3. Southern blotting analysis of the mouse HSP105 gene. (A) The restriction map of the HSP105 gene is shown schematically. Exons and introns are shown as closed and open columns, respectively. The position of the probe on the HSP105 gene used for Southern blotting is shown by an arrow. B, *Bam*HI; K, *Kpn*I; H, *Hin*dIII. (B) After digestion of mouse genomic DNA with *Bam*HI, *Eco*RI, *Kpn*I, or *Sac*I, the DNA was analyzed by Southern blotting using the probe. The numbers on the right indicate sizes (kb) of *Hin*dIII-digested λ DNA markers.

FIG. 4. Nucleotide sequence of the 5′-flanking region of mouse HSP105 gene. The nucleotide sequence of the 5′-upstream region is shown with the transcription initiation site indicated as +1 and the TATA sequence is shown in bold. The putative CAAT box, inverted CAAT box, GC box homology are indicated. The GAA cores of the HSE elements are shown with dots. The nucleotides encoding the cDNA region (exon 1) are shown in uppercase letters, while the 5′-flanking region and intron are in lowercase (GenBank DNA database Accession No. AB005267).

internal exons are usually less than 300 bp, while the last exon (exon 18) was large (867 nt). Introns varied in size from 0.1 to 2.8 kb. Furthermore, exon 12 encoded cDNA sequence between nt 1642 and nt 1773 of mouse HSP105α, which is lacking in the HSP105β cDNA. Thus, exon 12 seemed to be an alternatively spliced exon, and exon 11 can be spliced to exon 13, but not to exon 12, resulting in HSP105β mRNA when cells are heat shocked at 42°C.

Determination of transcription initiation site. The transcription initiation site of the HSP105 gene was determined by primer extension of poly(A)$^+$ RNA with AMV-reverse transcriptase. Figure 2 shows the nucleotide sequence of the DNA near the 5′ end of HSP105 mRNA. Alignment of the nucleotides with the genomic sequence indicated that the majority of transcription originates at or near thymidine residue 165 nt upstream of the ATG initiation codon.

Southern blotting analysis. Genomic DNA extracted from mouse FM3A cells was digested with several different restriction enzymes, i.e., *Bam*HI, *Eco*RI, *Kpn*I, and *Sac*I, and hybridized with the approximately 1.4-kb fragment containing intron 17 and exon 18 of the HSP105 gene as a probe (Fig. 3). Under high stringency conditions, since only one DNA fragment in each digest was hybridized to the probe, one copy of HSP105 was suggested to exist in the mouse genome. The size of the band generated by digestion with *Bam*HI (about 14 kb) was about the same as expected from the restriction map.

Analysis of the promoter sequence. In addition to the exon-intron structure of the HSP105 gene, we determined about 1.2 kb of the 5′-flanking region of the mouse HSP105 gene (Fig. 4). TATA box matching the consensus sequence in 6 of 7 nucleotides (TATA(A/T)A(A/T)) (25) was observed at nt −33. A CAAT box and an inverted CAAT box (26) were found at nt −210 and nt −141, respectively. Two GC boxes, a potential binding site for SP1 transcription factor (27, 28), were also found at nt −248 and nt −166. There were two HSE consensus sequences (29, 30) between nt −99 and nt −128 and between nt −45 and nt −64. Both HSE sequences were composed of four 5 bp (nGAAn or nTTCn) motifs with 11/12 matches to the consensus.

Promoter analysis of the mouse HSP105 gene. To examine promoter activity of the mouse HSP105 gene, constructs containing the chloramphenicol acetyltransferase (CAT) gene under the control of a fragment of the 5′-flanking region were transfected into mouse C3H10T1/2 cells, and the constitutive and heat-inducible promoter activities were determined (Fig. 5). The construct pCAT105 which contained 1.2 kb of the

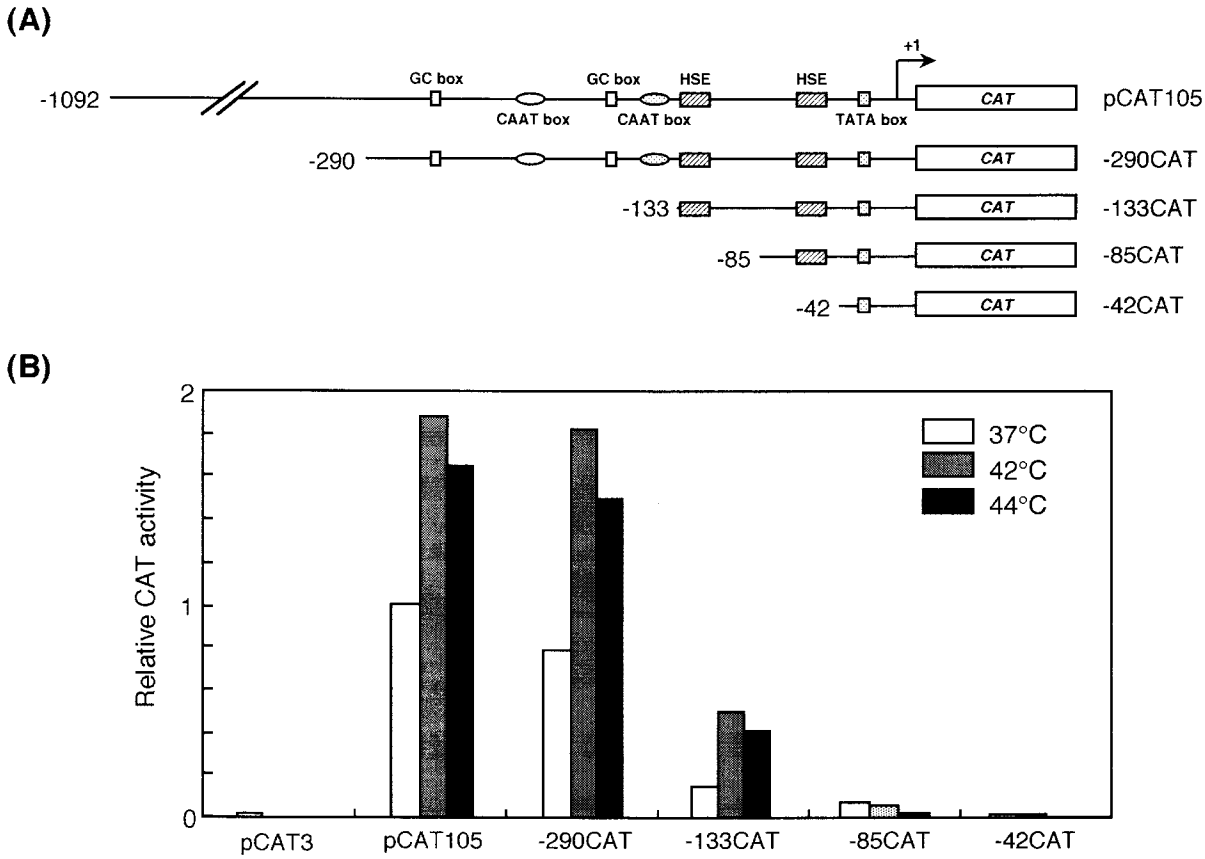

FIG. 5. CAT assay of HSP105 promoter activity. (A) The CAT gene constructs with various 5′-deleted HSP105 gene promoter and the structure of the 5′-promoter region of HSP105 gene are shown schematically. (B) Nine μg of construct pCAT105, −290CAT, −133CAT, −85CAT, −42CAT or empty pCAT3 basic vector with 1 μg of pCDM-Gal from which β-galactosidase was constitutively expressed (22) were transfected into mouse C3H10T1/2 cells (1×10^5 cells) by the calcium phosphate method (23). After incubation at 37°C for 42 h, these cells were incubated at 37°C for 6 h (37°C), at 37°C for 5 h after heat shock at 42°C for 1 h (42°C), or at 37°C for 6 h after heat shock at 44°C for 15 min (44°C). After harvesting cells, CAT activity was assayed, and relative CAT activities are shown as ratios to that of control cells transfected with pCAT105. The transfection efficiency was monitored by determination of the β-galactosidase activity and used for correction of CAT activity. The results of one representative experiment of two are shown.

5′-flanking region exhibited CAT activity in control cells and the activity increased by about 1.6-~2-fold after exposure to heat shock. Since HSP105 is constitutively expressed in non-stressed cells and is also induced to relatively low levels by heat shock (4, 5), the 1.2-kb fragment of the 5′-flanking region seemed to represent the native constitutive and heat-inducible promoter activities of the HSP105 gene. Deletion to nt −290 did not apparently affect the promoter activity in either control or heat-shocked cells (−290CAT). Deletion to nt −133 markedly reduced the expression level in both control and heat-shocked cells (−133CAT). However, the heat-inducibility of −133CAT still remained. These results indicated that the region between nt −290 and nt −133 contained sequences that influence basal expression. Cis-acting elements such as GC and CAAT boxes found in this region may be responsible for constitutive expression. Deletion upstream of nt −85, which resulted in deletion of the first HSE sequence, caused marked reductions in not only heat-inducible expression but also basal expression level (−85CAT). Deletion to nt −42 resulted in no detectable activity (−42CAT). These results suggested that the second HSE sequence could not support heat-inducibility and that downstream of nt −133 was the minimal sequence essential for heat shock response and basal expression of the HSP105 gene.

This is the first report of the unique genomic structure of mammalian HSP105, and these findings provided some important insight into the expression and function of the mammalian high molecular mass heat shock proteins.

ACKNOWLEDGMENT

This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture of Japan.

REFERENCES

1. Welch, W. J. (1992) *Physiol. Rev.* **72,** 1063–1081.
2. Hendric, J. P., and Hartl F-U. (1993) *Annu. Rev. Biochem.* **62,** 349–384.
3. Craig, E. A., Weissman, J. S., and Horwich, A. L. (1994) *Cell* **78,** 365–372.
4. Hatayama, T., Nishiyama, E., and Yasuda, K. (1994) *Biochem. Biophys. Res. Commun.* **200,** 1367–1373.
5. Hatayama, T., Yasuda, K., and Nishiyama, E. (1994) *Biochem. Biophys. Res. Commun.* **204,** 357–365.
6. Hatayama, T., Honda, K., and Yukioka, M. (1986) *Biochem. Biophys. Res. Commun.* **137,** 957–963.
7. Honda, K., Hatayama, T., and Yukioka, M. (1988) *J. Biochem.* **103,** 81–85.
8. Honda, K., Hatayama, T., and Yukioka, M. (1989) *Biochem. Biophys. Res. Commun.* **160,** 60–66.
9. Yasuda, K., Nakai, A., Hatayama, T., and Nagata, K. (1995) *J. Biol. Chem.* **270,** 29718–29723.
10. Ishihara, K., Yasuda, K., and Hatayama, T. (1998) *Biochim. Biophys. Acta,* in press.
11. Hatayama, T., Takigawa, T., Takeuchi, S., and Shiota, K. (1997) *Cell Struct. Funct.* **22,** 517–525.
12. Lee-Yoon, D., Easton, D., Murawski, M., Burd, R., and Subjeck, J. R. (1995) *J. Biol. Chem.* **270,** 15725–15733.
13. Storozhenko, S., De Pauw, P., Kushnir, S., Van Montagu, M., and Inze, D. (1996) *FEBS Lett.* **390,** 113–118.
14. Mukai, H., Kuno, T., Tanaka, H., Hirata, D., Miyakawa, T., and Tanaka, C. (1993) *Gene* **132,** 57–66.
15. Kaneko, Y., Nishiyama, H., Nonoguchi, K., Higashitsuji, H., Kishishita, M., and Fujita, J. (1997) *J. Biol. Chem.* **272,** 2640–2645.
16. Kaneko, Y., Kimura, T., Kishishita, M., Noda, Y., and Fujita, J. (1997) *Gene* **189,** 19–24.
17. Just, M. L., and Lennarz, W. J. (1997) *Dev. Biol.* **184,** 25–30.
18. Chung, K. S., Hoe, K. L., Kim, K. W., and Yoo, H. S. (1998) *Gene* **210,** 143–150.
19. Plesofsky-Vig, N., and Brambl, R. (1998) *J. Biol. Chem.* **273,** 11335–11341.
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
21. Chilgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
22. Tsukada, T., Nakai, Y., Fukushima, M., Usui, T., Imura, H., and Takabe, H. (1994) *DNA Cell Biol.* **13,** 755–762.
23. Graham, F. L., and Van der Eb, A. J. (1973) *Virology* **52,** 456–467.
24. Senapathy, P., Shapiro, M. B., and Harris. M. L. (1990) *Methods Enzymol.* **183,** 252–278.
25. Breathnach, R., and Chambon, P. (1981) *Annu. Rev. Boiochem.* **50,** 349–383.
26. Efstratiadis, A., Posokany, J. W., Maniatis, T., Lawn, R. M., O'Connell, C., Spritz, R. A., DeRiel, J. K., Forget, B. G., Weissman, S. M., Slightom, J. L., Blechl, A. E., Smithies, O., Baralle, F. E., Shoulders, C. C., and Proudfoot, N. J. (1980) *Cell* **21,** 653–668.
27. Dynan, W. S., and Tjian, R. (1985) *Nature* **316,** 774–778.
28. Ishii, S., Kadonaga, J. T., Tjian, R., Brady, J. N., Merlino, G. T., and Pastan I. (1986) *Science* **232,** 1410–1413.
29. Amin, J., Ananthan, J., and Voellmy, R. (1988) *Mol. Cell. Biol.* **8,** 3761–3769.
30. Perisic, O., Xiao, H., and Lis, J. T. (1989) *Cell* **59,** 797–806.