ELSEVIER

Biochimica et Biophysica Acta 1314 (1996) 120–124

BIOCHIMICA ET BIOPHYSICA ACTA
BBA

Short sequence-paper

Isolation and characterization of a human heart cDNA encoding a new member of the small heat shock protein family — HSPL27 [1]

Wai Yip Lam, Stephen Kwok Wing Tsui, Patrick Tik Wan Law, Sharon Chui Wah Luk, Kwok Pui Fung, Cheuk Yu Lee, Mary Miu Yee Waye *

Department of Biochemistry, The Chinese University of Hong Kong, Shatin N.T., Hong Kong

Received 11 June 1996; revised 6 August 1996; accepted 14 August 1996

Abstract

A novel cDNA clone was isolated from a human adult heart cDNA library. This cDNA clone is similar to the small heat shock protein (smhsp) in both DNA and amino acid sequences, especially in the conserved region. Sequence analysis has shown that the putative novel smhsp, named 27 kDa heat-shock-protein-like protein (HSPL27) is a protein of 241 amino acids with a deduced molecular mass of 26.7 kDa and a deduced pI of 8.0. We have expressed the HSPL27 in *E. coli* and the expressed protein was found to be present in the soluble fraction of the bacterial cell lysate. Chromosomal mapping data shows that the HSPL27 gene is located at human chromosome 5q11.2.

Keywords: Small-heat-shock-protein; Heart cDNA; Chromosome 5; (Human)

The molecular weight of heat shock or stress proteins range from 10000 to 110000 Da. These proteins are produced at extremely high levels in the stressed cell and show a remarkably high degree of structural conservation. Recent findings suggested that increased expression of these proteins also occurs when cells are exposed to a number of deleterious agents or treatment, such as heavy metals, amino acid analogues or other kinds of stress [1,2]. Of the six or seven major stress proteins, more attention has been paid to the higher molecular weight stress proteins [3] than the smaller ones in mammalian cells. To date, only a single DNA sequence of human 27 kDa heat shock protein (hsp27) has been reported [4,5]. Three additional members of the human smhsp family have also been identified. They are the alpha A- and alpha B-crystallins [6], and a protein called p20 [7]. We describe here the isolation of a full length cDNA encoding a novel fifth member of human smhsp family called 27 kDa heat-shock-protein-like protein (HSPL27). The cloning, sequencing, expression and chromosomal mapping of this cDNA will also be presented.

Abbreviations: a.a., amino acid; CIAP, chicken inhibitor of actin polymerization; hsp27, 27 kDa heat shock protein; HSPL27, 27 kDa heat-shock-protein-like protein; hsp, heat shock protein; IPTG, isopropylthio-β-D-galactoside; NCBI, National Center of Biotechnology Information; PCR, polymerase chain reaction; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; smhsp, small heat shock proteins

* Corresponding author. Fax: +852 26035123; e-mail: marywaye@cuhk.edu.hk.

[1] The nucleotide sequence data reported in this paper have been submitted to the EMBL/GenBank Databases under the accession number U15590.

PII S0167-4889(96)00121-8

Partial sequencing of cDNA clones (in lambda gt11 vector) isolated from human adult heart were performed as described previously [8–10]. The partial sequences were compared against the EMBL/GenBank Data Libraries using the BLAST electronic mail server [11]. Among the various cDNA clones that we have sequenced, a clone which is similar to hsp27 was identified. Excluding the vector sequences and poly-A sequence, the cDNA obtained is 904 base pairs in length. An open reading frame (ORF) of 723 base pairs was found encoding a protein of 241 amino acid residues (Fig. 1). By aligning the DNA

```
                                    1           12          21          32
                              5'    CAT GCA GCA TCA GCA TCT TAT AGT AGT TGA AC

             41          50          59          68          77          86
     ATG TCG CCA TCC AGG CCG CCC ATC ACA TCC ACA AAC TCT GTG GTC AGC TTA AAG
 1   Met Ser Pro Ser Arg Pro Pro Ile Thr Ser Thr Asn Ser Val Val Ser Leu Lys     18

             95         104         113         122         131         140
     GCT GAC GTC TCA AAG CCC AGA TTT CGG GGT GAG CTG GAG AGG ATG AAG CCA AAG
19   Ala Asp Val Ser Lys Pro Arg Phe Arg Gly Glu Leu Glu Arg Met Lys Pro Lys     36

            149         158         167         176         185         194
     TCG ATG TGG ATG ATG TGG CCT TCT GCG TCC AAA AGG ATA TTC CCA TTG TGT CTG
37   Ser Met Trp Met Met Trp Pro Ser Ala Ser Lys Arg Ile Phe Pro Leu Cys Leu     54

            203         212         221         230         239         248
     AGG AGG GGG AAT AGA GGA GAG GAA GAG GCA GTA GGC AAC TGC AGG GGC TCG CCA
55   Arg Arg Gly Asn Arg Gly Glu Glu Glu Ala Val Gly Asn Cys Arg Gly Ser Pro     72

            257         266         275         284         293         302
     CTG ACT GAA GGC AGT GGA AGG TTG GCA GAA GGA GGC TGT TCA AGG CTG TTT TTG
73   Leu Thr Glu Gly Ser Gly Arg Leu Ala Glu Gly Gly Cys Ser Arg Leu Phe Leu     90

            311         320         329         338         347         356
     CCT TCA CTA TGG CAA AAT CAT TTT GAG GCA CCT CAT AGA GAT TCC AGT GCG TTA
91   Pro Ser Leu Trp Gln Asn His Phe Glu Ala Pro His Arg Asp Ser Ser Ala Leu    108

            365         374         383         392         401         410
     CCA GGA AGA GTT GAA GCT CGA GGT CTA GAA GAC TGC AGG CTG GAT CAT GCT TAT
109  Pro Gly Arg Val Glu Ala Arg Gly Leu Glu Asp Cys Arg Leu Asp His Ala Tyr    126

            419         428         437         446         455         464
     ATG CAC TGC CTG GCC AAC CAT CGT GGA CCT GAG GAA AAC CAG GGC AGC GCA GTC
127  Met His Cys Leu Ala Asn His Arg Gly Pro Glu Glu Asn Gln Gly Ser Ala Val    144

            473         482         491         500         509         518
     TCC TCC AGT GGA CTA AGC GGC AGA GAC GCC ACC CCG AGA AGG CAA ATC CCA CTT
145  Ser Ser Ser Gly Leu Ser Gly Arg Asp Ala Thr Pro Arg Arg Gln Ile Pro Leu    162

            527         536         545         554         563         572
     TCA GAT CCT GTG GAC GTG GTC CAG TTC CTC CCT GAA GAC ATC ATC ATT CAG ACC
163  Ser Asp Pro Val Asp Val Val Gln Phe Leu Pro Glu Asp Ile Ile Ile Gln Thr    180

            581         590         599         608         617         626
     TTC GAA GGC TGG CTA CTG ATA AAA GCA CAA CAC GGA ACC AGA ATG GAT GAG CAC
181  Phe Glu Gly Trp Leu Leu Ile Lys Ala Gln His Gly Thr Arg Met Asp Glu His    198

            635         644         653         662         671         680
     GGT TTT ATC TCA AGA AGC TTC ACC CGA CAG TAC AAA CTA CCA GAT GGC GTG GAA
199  Gly Phe Ile Ser Arg Ser Phe Thr Arg Gln Tyr Lys Leu Pro Asp Gly Val Glu    216

            689         698         707         716         725         734
     ATC AAA GAT TTG TCT GCA GTC CTC TGT CAT GAT GGA ATT TTG GTG GTG GAA GTA
217  Ile Lys Asp Leu Ser Ala Val Leu Cys His Asp Gly Ile Leu Val Val Glu Val    234

            743         752         761         770         779         788
     AAG GAT CCA GTT GGG ACT AAG TGA CAT CGT ATC GGT TCC TGT TCA GAT GAC ATG
235  Lys Asp Pro Val Gly Thr Lys ***                                          241

            797         806         815         824         833         842
     GGG AAG ATG ATG GTT CAT CCA CTG GTA CTA CTA GAA TGT TTG TAT TAC CCA CAT

            851         860         869         878         887         896
     TTG AAA TGC CTT GCT ATG AAT TTT TAT GAA GAA TAA AAA TAT ATA CAC AGT TAA

            905
     AAA AAA AAA AAA AAA AA 3'
```

Fig. 1. The complete DNA and amino acid sequence of human HSPL27. Translational start codon at 33 bp and stop codon at 756 bp. Polyadenylation signal is at 874 bp. Serine-106 is a possible phosphorylation site and the amino acid sequence shows a carboxyl-terminal lysine.

```
HSPL27   150  SGRDATPRRQIPLSDPVDVVQFLPEDIIIQ   179
              :: ..... . . . ..::: .: :......
HSP27     83  SGVSEIRHTADRWRVSLDVNHFAPDELTVK   112
              ::.:::::..::.:.:::::::::.::.::
CIAP      80  SGISEIRQSADSWKVTLDVNHFAPEELVVK   109

HSPL27   180  TFEGWLLIKAQHGTRMDEHGFISRSFTRQY   209
              : .: . :...:..: ::::.:::.::::.:
HSP27    113  TKDGVVEITGKHEERQDEHGYISRCFTRKY   142
              :::..:::::::::.:::::.:::::::::
CIAP     110  TKDNIVEITGKHEEKQDEHGFISRCFTRKY   139

HSPL27   210  KLPDGVEIKDLSAVLCHDGILVVEVKDPV   238
              .:: ::. ...:. :...:::.::.  :
HSP27    143  TLPPGVDPTQVSSSLSPEGILIVEAPMPK   172
              :::::::.:.:.:::::.:.:.::::.::
CIAP     140  TLPPGVEATAVRSSLSPDGMLTVEAPLPK   168
```

Fig. 2. The amino acid sequence alignment of the conserved domain of HSPL27 (HSPL27) with that of the human hsp27 (HSP27) and the chicken inhibitor of actin polymerization (CIAP). The alignment was performed using PROSIS from Hitachi. Amino acids that are identical between the two sequences are marked by ':' while those that are similar were marked by '.'. Those amino acids that are different between the two sequences are left blank.

sequences of the ORF of human hsp27 and HSPL27, we found them to be 50.7% similar (data not shown). When the amino acid sequences of human hsp27 and HSPL27 were aligned, an overall identity of 37.1% and a similarity of 80.4% was shown; while a 44.1% identity and 86.8% similarity within the conserved region (from 167 to 234 amino acid residues with respect to the amino acid sequence of HSPL27) of this family were found (Fig. 2). Since it has been stated that smhsp are the least conserved family [4,5], we do not expect a very high degree of DNA or amino acid sequence similarity, especially in the N-terminals, between human HSPL27 and smhsp of other species, such as drosophila, and mouse. Besides matching with hsp27 from various species, the putative protein sequence of human HSPL27 also matches significantly with that of an inhibitor of actin polymerization in chicken (CIAP) [12], with 56.9% identity in DNA sequences and 39.1% identity, 80.1% similarity in amino acid sequences (Fig. 2). In addition, recent reports have suggested that hsp may mediate protection of ischaemic cells from necrosis by preventing abnormal actin aggregation [13,14]. We speculate that HSPL27 isolated from the heart may have a function in stabilization of cytoskeleton, especially actin [15], in myocytes.

Another indirect evidence in support of the hypothesis that our HSPL27 belongs to the smhsp family is that a type of amine-donor called carboxyl-terminal lysine is present in HSPL27 (residue 241). Recently, it has been shown that such carboxyl-terminal lysine residues are found in B-crystallin and mouse hsp25 and they serve as amine-donor substrates in the transglutaminase-mediated cross-linking reaction of proteins [16]. Although the significance of substrate capacity for transglutaminase is unclear, it has been proposed that transglutaminase activity is implicated in controlled cell death [17], tumor progression [18], and receptor-mediated endocytosis [19]. These are processes which involve changes in intracellular organization. The fact that smhsp are also implicated in cytomorphological changes may provide a clue for the interactions between these proteins and transglutaminase [16]. As reported previously, the function of human heat shock 27 kDa protein (hsp27) may be regulated by phosphorylation [15,20]. Thus, we searched for possible phosphorylation sites of our amino acid sequence which was quite serine rich (10%, 24 out of 241 amino acid residues). Several possible sites were found after analyzing the amino acid sequence by using the computer software MacDNASIS. The most probable site of phosphorylation by protein kinase C or protein kinase A was found to be at ser-106 as it possesses the conserved

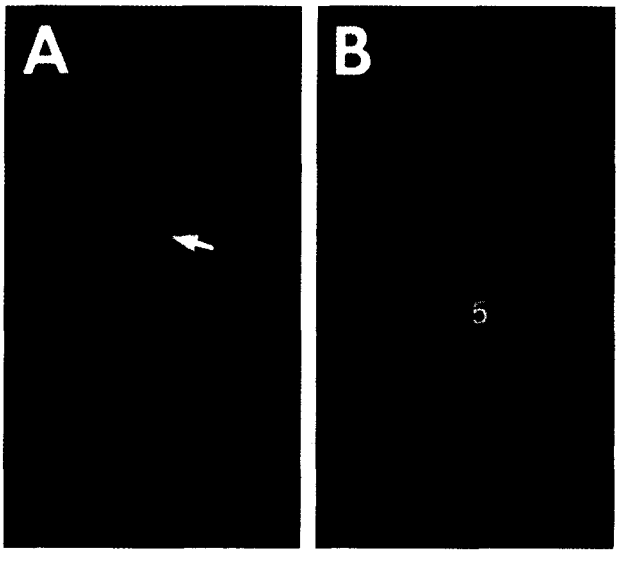

Fig. 3. FISH mapping. Panels A and C showing the FISH signals on chromosomes; panels B and D showing the same mitotic figures stained with DAPI to identify chromosome 5.

kinase recognition amino acid sequence 'Arg-X-X-Ser' [20,21].

In order to express the protein in *E. coli*, the human HSPL27 was cloned into an expression vector pAED4. The recombinant plasmid pAED4-HSPL27 was transformed into the *E. coli* BL21(DE3)pLysS strain. The recombinant protein was expressed according to the protocol previously described [22]. Electrophoresis of crude bacterial extracts was carried out and the recombinant protein showed an apparent mol. wt. of 26.7 kDa (data not shown). For solubility analysis, the bacterial cells were lysed by freeze-thawing twice. After centrifugation, suspension was divided into a water soluble fraction and a pellet. Both fractions were resuspended in SDS/PAGE-sample buffer and then electrophoresed. We found that the recombinant human HSPL27 protein is present in the soluble bacterial cell fraction (data not shown).

The chromosomal mapping of the HSPL27 gene was done by the SeeDNA Biotechology Inc. (Ontario, Canada). The pAED4-HSPL27 plasmid was biotinylated with dATP. The procedure for FISH detection was performed essentially as described previously [23,24]. Based on the results from 10 photos, the HSPL27 gene is located at human chromosome 5q11.2 (Fig. 3). Previous report has shown that the locations for hsp27 gene and its two pseudogenes are at chromosomes 3, 9 and X [25]. Hence, the hsp27 gene and the HSPL27 gene are located at different chromosomes. According to the database of NCBI, dihydrofolate reductase gene [26], gp130 transducer chain gene [27], 5-hydroxytryptamine-1A receptor gene [28], corticotropin releasing hormone-binding protein gene [29] and genes for the NK cell granule serine proteases [30] are also located in this region. Moreover, Klippel-Feil syndrome [31] and Schizophrenia disorder-1 [32] have been mapped to this region.

This work was supported by a grant from Research Grants Council (CUHK 16/93M) and a postdoctoral fellowship to S.K.W. Tsui from the Chinese University of Hong Kong. We would like to thank Dr. Don Doering and Dr. Paul Matsudaira for providing us the pAED4 expression vector. We also thank Dr. H.H.Q. Heng from the SeeDNA Biotechnology Inc. for mapping the HCRP2 by FISH. The helpful advice of Dr. C.C. Liew is greatly appreciated.

References

[1] Parsell, D.A., and Lindquist, S. (1993) Annu. Rev. Genet. 27, 437–496.
[2] Welch, W.J. and Suhan, J.P. (1986) J. Cell Biol. 103, 2035–2052.
[3] Linquist, S. and Craig, E.A. (1988) Annu. Rev. Genet. 22, 631–677.
[4] Hickey, E., Brandon, S.E., Potter, R., Stein, G., Stein, J. and Weber, L.A. (1986) Nucleic Acids Res. 14, 4127–4145.
[5] Hickey, E., Brandon, S.E., Sadis, S., Smale, G. and Weber, L.A. (1986) Gene 43, 147–154.
[6] Quax-Jeuken, Y., Quax, W., van Rens, G., Khan, P.M. and Bloemendal, H. (1985) Proc. Natl. Acad. Sci. USA 82, 5819–5823.
[7] Kato, K., Goto, S., Inaguma, Y., Hasegawa, K., Morishita, R. and Asano, T. (1994) J. Biol. Chem. 269, 15302–15309.
[8] Liew, C.C. (1993) J. Mol. Cell. Cardiol. 25, 891–894.
[9] Liew, C.C., Hwang, D.M., Fung, Y.W., Laurenssen, C., Cukerman, E., Tsui, S. and Lee, C.Y. (1994) Proc. Natl. Acad. Sci. USA 91, 10645–10649.
[10] Tsui, S.K.W., Waye, M.M.Y. and Lee, C.Y. (1995) BioTechniques 19, 577–578.
[11] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.
[12] Miron, T.K., Vancompernolle, K., Vandekerckhove, J., Wilchek, M. and Geiger, B. (1991) J. Cell Biol. 114, 255–261.
[13] Ganote, C. and Armstrong, S. (1993) Cardiovasc. Res. 27, 1387–1403.
[14] Kabakov, A.E. and Gabai, V.L. (1994) Trend. Cell Biol. 4, 193–196.
[15] Lavoie, J.N., Hickey, E., Weber, L.A. and Landry, J. (1993) J. Biol. Chem. 268, 24210–24214.
[16] Merck, K.B. and Groenen, P.J.T.A., Voorter, C.E.M., de Haard Hoekman, W.A., Horwitz, J., Bloemendal, H. and de Jong, W.W. (1993) J. Biol. Chem. 268, 1046–1052.
[17] Fesus, L., Thomazy, V. and Falus, A. (1987) FEBS Lett. 224, 104–108.
[18] Knight, C.R.L., Rees, R.C. and Griffin, M. (1991) Biochim. Biophy. Acta. 1096, 312–318.
[19] Davies, P.J.A., Davies, D.R., Levitzki, A., Maxfield, F.R., Milhaud, P., Willingham, M.C. and Pastan, I.H. (1980) Nature 283, 162–166.
[20] Gaestel, M., Schroder, W., Benndorf, R., Lippmann, C., Buchner, K., Hucho, F., Erdmann, V.A. and Bielka, H. (1991) J. Biol. Chem. 266, 14721–14724.
[21] Landry, J., Lambert, H., Zhou, M., Lovoie, J.N., Hickey, E., Weber, L.A. and Anderon, W. (1992) J. Biol. Chem. 267, 794–803.
[22] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990) Method enzymol. 185, 60–89.
[23] Heng, H.H.Q., Squire, J. and Tsui, L.C. (1992) Proc. Natl. Acad. Sci. USA 89, 9509–9513.
[24] Heng, H.H.Q. and Tsui, L.C. (1993) Chromosoma 103, 325–332.

[25] McGuire, S.E., Fuqua, S.A., Naylor, S.L., Helin-Davis, D.A. and McGuire, W.L. (1989) Somat. Cell Mol. Genet. 15, 167–171.
[26] Funanage, V.L., Myoda, T.T., Moses, P.A. and Cowell, H.R. (1984) Mol. Cell. Biol. 4, 2010–2016.
[27] Rodriguez, C., Grosgeorge, J., Nguyen, V.C., Gaudray, P. and Theillet, C. (1995) Cytogenet. Cell Genet. 70, 64–67.
[28] Kobilka, B.K., Frielle, T., Collins, S., Yang-Feng, T., Kobilka, T.S., Francke, U., Lefkowitz, R.J. and Caron, M.G. (1987) Nature 329, 75–79.
[29] Vamvakopoulos, N.C., Sioutopoulou, T.O., Durkin, S.A., Nierman, W.C., Wasmuth, J.J. and McPherson, J.D. (1995) Genomics 25, 325–327.
[30] Baker, E., Sayers, T.J., Sutherland, G.R. and Smyth, M.J. (1994) Immunogenetics 40, 235–237
[31] Fukushima, Y., Ohashi, H., Wakui, K., Nishimoto, H., Sato, M. and Aihara, T. (1995) Am. J. Med. Genet. 57, 447–449.
[32] Bassett, A.S., McGillivray, B.C., Jones, B.D. and Pantzar, J.T. (1988) Lancet I, 799–801.