Vol. 197, No. 1, 1993
November 30, 1993

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
Pages 235-240

CLONING OF A cDNA FOR HEAT-SHOCK PROTEIN hsp40, A HUMAN HOMOLOGUE OF BACTERIAL DnaJ*

Kenzo Ohtsuka

Laboratory of Experimental Radiology
Aichi Cancer Center Research Institute
Chikusa-ku, Nagoya 464, Japan

Received October 8, 1993

Summary: We here isolated and characterized a cDNA clone encoding 40-kDa heat-shock protein hsp40 from cDNA expression library of human placenta by immunoscreening with anti-hsp40 antibody. N-terminal amino acid sequence (48 residues) deduced from the base sequence was completely identical to that of hsp40 purified from HeLa cells. Deduced amino acid sequence of the cDNA is homologous to bacterial DnaJ heat-shock protein and its homologues in yeast such as SCJ1, YDJ1(MAS5), SIS1, SEC63 and Zuotin. Nucleotide sequence identity between hsp40 and HDJ1 (another human DnaJ homologue) is more than 98%, suggesting that these two proteins are the same gene product.

Several DnaJ homologues have been recently identified in yeast [SCJ1, YDJ1(MAS5), SIS1, SEC63 and Zuotin] [1-6], plant (ANJ1) [7], fruit fly (csp32) [8] and human (HDJ1, HDJ2 and HSJ1) [9-12]. Bacterial DnaJ heat-shock protein is known to function together with DnaK (hsp70) and GrpE as a molecular chaperone, involving them in assembly and disassembly of protein complexes, protein folding, renaturation of denatured proteins, prevention of protein aggregation and protein export [13]. Yeast YDJ1 has been shown to localize in cytosol and ER and to act in concert with hsp70 in protein import into mitochondria and ER [14, 15]. SCJ1 and SEC63 are respectively localized in mitochondria and ER, and are implicated in protein import into each organelle [1, 5, 16]. Also, SIS1 has been reported to be associated with translating ribosomes and is required for initiation of translation [17]. The current notion is that these DnaJ homologues have molecular chaperoning activity together with hsp70 in each cellular compartment [see references 18 and 19 for review].

Recently, we identified a novel heat-shock protein hsp40 in mammalian and avian cells [20] and reported that N-terminal amino acid sequence (48 residues) of purified

---

* The nucleotide sequence data reported in this paper will appear in the DDBJ/GenBank/EMBL Nucleotide Sequence Database with accession number D17749.

0006-291X/93 $4.00

human hsp40 is homologous to DnaJ and its homologues in yeast [21]. Also, we showed that hsp40 is colocalized with hsp70 in the nuclei and nucleoli of heat-shocked HeLa cells [22]. Thus, we recently suggested that hsp70(DnaK)-hsp40(DnaJ) chaperone systems is ubiquitous [22].

Here, we report the isolation and sequence analysis of a cDNA encoding human hsp40 by immunoscreening with anti-hsp40 polyclonal antibody.

MATERIALS AND METHODS

Human placenta cDNA expression library in the lambda ZAPII vector (Stratagene) was screened with polyclonal antibody against human hsp40 [21]. Insert from a positive clone was excised in vivo as subclone in pBluescript plasmid. Insert size was approximately 2 kbp. Double strand plasmid DNA was isolated by QIAGEN miniprep method and sequenced by dye terminator method with a DNA sequencer (Applied Biosystems, Inc., Model 373A). We first use reverse primer, T3 primer and T7 primer that are located near EcoRI insert site for sequencing. Then 20 mer oligonucleotides were synthesized according to the first sequencing and used for further sequencing. The synthetic oligonucleotide primers were 200 ~ 300 bp apart from each other. Sequencing was performed from both orientations. DNA sequence data were analyzed using DNASIS computer program and the nucleotide and amino acid sequence homology search through the GenBank/EMBL was performed by the software program DNASIS.

Total RNA was isolated by AGPC method [23] from non-heat-shocked control and heat-shocked HeLa cells, and subjected to Northern blot analysis with EcoRI fragment insert as probe. ß-actin cDNA probe was used as control.

RESULTS AND DISCUSSION

By screening human placenta cDNA expression library with anti-hsp40 antibody, only one positive clone was identified out of 2 x 10^5 plaques. Insert size was approximately 2 kbp. We sequenced 1537 base pairs of the insert. The insert has an open reading frame of 1020 base pairs that can encode 340 amino acids (Fig. 1). Its molecular weight is 38042.16 dalton and theoretical pI 8.92. These parameters well correspond to that of hsp40 as determined by two-dimensional gel electrophoresis (MW of approximately 40-kDa and pI of 9.2) [20]. Also, N-terminal amino acid sequence (48 residues, double underline in Fig. 1) deduced from the nucleotide sequence is completely identical to that of HeLa cell hsp40 as determined by amino acid sequencer [21]. Northern blot analysis using EcoRI fragment insert as probe showed that the transcription of the gene was drastically induced by heat-shock at 45°C for 15 min (arrowhead in Fig. 2). Thus, we conclude that the isolated cDNA encodes human hsp40.

Comparison of hsp40 to EMBL amino acid data base showed that hsp40 is 34% identical, over 340 amino acid residues, to *E. coli* DnaJ protein [24, 25]. The highest region is at N-terminal 70-amino acid residues, which is common to other DnaJ homologues. Hsp40 is 34% to 40% identical to long stretches of yeast DnaJ homologues SCJ1, YDJ1(MAS5), SIS1, SEC63 and Zuotin, with SIS1 being the most

```
                                   GGGACGGCGACAGCGGGTCGGCGGGCCGCAGGAGGGGGTC    40
ATGGGTAAAGACTACTACCAGACGTTGGGCCTGGCCCGCGGCGCGTCGGACGAGGAGATCAAGCGGGCCTACCGC   115
M  G  K  D  Y  Y  Q  T  L  G  L  A  R  G  A  S  D  E  E  I  K  R  A  Y  R      25

CGCCAGGCGCTGCGCTACCACCCGGACAAGAACAAGGAGCCCGGCGCCGAGGAGAAGTTCAAGGAGATCGCTGAG   190
R  Q  A  L  R  Y  H  P  D  K  N  K  E  P  G  A  E  E  K  F  K  E  I  A  E      50

GCCTACGACGTGCTCAGCGACCCGCGCAAGCGCGAGATCTTCGACCGCTACGGGGAGGAAGGCCTAAAGGGGAGT   265
A  Y  D  V  L  S  D  P  R  K  R  E  I  F  D  R  Y  G  E  E  G  L  K  G  S      75

GGCCCCAGTGGCGGTAGCGGCGGTGGTGCCAATGGTACCTCTTTCAGCTACACATTCCATGGAGACCCTCATGCC   340
G  P  S  G  G  S  G  G  G  A  N  G  T  S  F  S  Y  T  F  H  G  D  P  H  A     100

ATGTTTGCTGAGTTCTTCGGTGGCAGAAATCCCTTTGACACCTTTTTTGGGCAGCGGAACGGGGAGGAAGGCATG   415
M  F  A  E  F  F  G  G  R  N  P  F  D  T  F  F  G  Q  R  N  G  E  E  G  M     125

GACATTGATGACCCATTCTCTGGCTTCCCTATGGGCATGGGTGGCTTCACCAACGTGAACTTTGGCCGCTCCCGC   490
D  I  D  D  P  F  S  G  F  P  M  G  M  G  G  F  T  N  V  N  F  G  R  S  R     150

TCTGCCCAAGAGCCCGCCCGAAAGAAGCAAGATCCCCCAGTCACCCACGACCTTCGAGTCTCCCTTGAAGAGATC   565
S  A  Q  E  P  A  R  K  K  Q  D  P  P  V  T  H  D  L  R  V  S  L  E  E  I     175

TACAGCGGCTGTACCAAGAAGATGAAAATCTCCCACAAGCGGCTAAACCCCGACGGAAAGAGCATTCGAAACGAA   640
Y  S  G  C  T  K  K  M  K  I  S  H  K  R  L  N  P  D  G  K  S  I  R  N  E     200

GACAAAATATTGACCATCGAAGTGAAGAAGGGGTGGAAAGAAGGAACCAAAATCACTTTCCCCAAGGAAGGAGAC   715
D  K  I  L  T  I  E  V  K  K  G  W  K  E  G  T  K  I  T  F  P  K  E  G  D     225

CAGACCTCCAACAACATTCCAGCTGATATCGTCTTTGTTTTAAAGGACAAGCCCCACAATATCTTTAAGAGAGAT   790
Q  T  S  N  N  I  P  A  D  I  V  F  V  L  K  D  K  P  H  N  I  F  K  R  D     250

GGCTCTGATGTCATTTATCCTGCCAGGATCAGCCTCCGGGAGGCTCTGTGTGGCTGCACAGTGAACGTCCCCACT   865
G  S  D  V  I  Y  P  A  R  I  S  L  R  E  A  L  C  G  C  T  V  N  V  P  T     275

CTGGACGGCAGGACGATACCCGTCGTATTCAAAGATGTTATCAGGCCTGGCATGCGGCGAAAAGTTCCTGGAGAA   940
L  D  G  R  T  I  P  V  V  F  K  D  V  I  R  P  G  M  R  R  K  V  P  G  E     300

GGCCTCCCCCTCCCCAAAACACCCGAGAAACGTGGGGACCTCATTATTGAGTTTGAAGTGATCTTCCCCGAAAGG  1015
G  L  P  L  P  K  T  P  E  K  R  G  D  L  I  I  E  F  E  V  I  F  P  E  R     325

ATTCCCCAGACATCAAGAACCGTACTTGAGCAGGTTCTTCCAATATAGCTATCTGAGCTCCCCAAGGACTGACCA  1090
I  P  Q  T  S  R  T  V  L  E  Q  V  L  P  I  *                                340

GGGACCTTTCCAGAGCTCAAGGATTTCTGGACCTTTCTACCAGTTGTGGACCATGAGAGGGTGGGAGGGCCCAGG  1165
GAGGGCTTTCGTACTGCTGAATGTTTTCCAGAGCATATATTACAATCTTTCAAAGTCGCACACTAGACTTCAGTG  1240
GTTTTTCGAGCTATAGGGCATCAGGTGGTGGGAACAGCAGGAAAAGGCATTCCAGTCTGCCCCACTGGGTCTGGC  1315
AGCCCTCCCGGGATGGGCCCACATCCACCTCCAGTCCCTGGCCAGGGGTGAGAGGCAGACCAGCAGATGGACTTG  1390
ATCCCTCTGTGTCTTTGGGCTTCTGGCTGGTAGATAATGTCAACCTGCAGTCTTGATTCCCAGACCCTGTACACT  1465
CCTCCTTTTCTGTTGTGTGATCAGTTTGTGCTTTATTCTGTATTGGTCTCCCATGTCTTGCTCTTCTCCTGG      1537
```

Fig. 1. *Nucleotide sequence and deduced amino acid sequence of the cDNA for* human hsp40. The single-letter symbol for each amino acid is placed below the first nucleotide of its codon. The termination codon (TAG) is labeled with an asterisk. Double-underlined amino acid sequence indicates sequence identical to that of HeLa cell hsp40 determined by amino acid sequencer [21]. Underlined amino acids indicate glycine and phenylalanine residues in G/F domain.

homologous. Hsp40 is 34% and 36% identical to human DnaJ homologues HSJ1 and HDJ2 (HSDJ), respectively. Surprisingly, comparison of hsp40 with HDJ1 [9], another human DnaJ homologue, revealed that while amino acid sequence of both

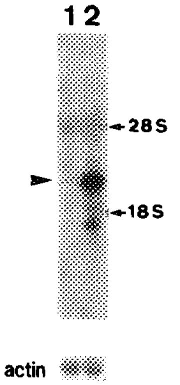

Fig. 2. Northern blot analysis in HeLa cells. Total RNA (8 μg / lane) isolated from non-heat-shocked control cells (lane 1) and cells heated at 45°C for 15 min and allowed to recover for 3 hr (lane 2) were subjected to Northern blot probed with EcoRI fragment insert (upper panel) or with β-actin cDNA (lower panel) as control. Arrowhead indicates the transcript of *hsp40* gene.

TABLE 1
Domain organization and biochemical properties of bacterial DnaJ and its homologues (hsp40 family)

Protein	Domains			MW (kDa)	pI[d]	Induction by heat	Intracellular localization
	J[a]	G/F[b]	C[c]				
DnaJ	J	G/F	C	41.1	7.72	+	membrane fraction
SCJ1	J	G/F	C	39.5	6.78	?[e]	mitochondria
YDJ1(MAS5)	J	G/F	C	44.7	5.91	+	cytosol and ER
SIS1	J	G/F	-	37.6	9.41	+	cytosol and nucleus
SEC63	J	-	-	75.3	4.97	?	ER
Zuotin	J	-	-	49.0	8.67	?	nucleus
ANJ1	J	G/F	C	46.6	6.51	+	(roots and shoots)
csp32	J	-	-	26.9	5.47	?	(brain)
hsp40(HDJ1)	J	G/F	-	38.0	8.92	+	cytosol and nucleus
HDJ2(HSDJ)	J	G/F	C	44.9	6.70	?	?
HSJ1	J	G/F	-	39.0	8.46	?	(brain)

[a] J, J domain highly conserved 70-amino acids N-terminal region.
[b] G/F, glycine- and phenylalanine-rich region distal to the J domain.
[c] C, cysteine-rich repeats in the middle of the protein.
[d] theoretical isoelectric point calculated by the software program DNASIS.
[e] ? indicates not determined.

proteins is 77% identical, their nucleotide sequences are more than 98% identical, including 5' and 3' noncoding region. There are seven mismatches and five frame shifts in coding region between these two cDNAs (data not shown), probably explaining why they do not show identical amino acid sequences. Therefore, it is concluded that hsp40 and HDJ1 are the identical gene products. Since there are at least three DnaJ homologues [hsp40(HDJ1), HSJ1, HDJ2(HSDJ)] in human and five in yeast, we here propose that they represent an hsp40 family in eukaryotes, like the hsp70 family and hsp90 family.

DnaJ protein is known to have three distinct domains [18, 19, 26]. A 70-amino acid sequence at the N-terminus, termed J domain, is highly conserved in all homologues. Distal to the J domain is a sequence rich in glycine and phenylalanine (G/F domain). There is also a cysteine-rich region (C domain) in the middle of the protein. The spacing of cysteine residue in this domain is CxxCxGxG (x is usually charged or polar residue) repeated four times. Amino acid sequence of C terminal half is less homologous to each other. As summarized in TABLE 1, yeast DnaJ homologues appear to have evolved in quite diverse fashion, that is, SCJ1 and YDJ1(MAS5) have all three domains, SIS1 has J and G/F domains, and SEC63 and Zuotin have only J domain.

Our hsp40 has J and G/F domains but not C domain (Fig. 1). In yeast DnaJ homologues, SIS1 has a domain structure and biochemical properties (MW, pI, heat-inducibility and intracellular localization) similar to human hsp40 [20, 21]. Also, amino acid sequence of hsp40 is the most homologous to that of SIS1 among yeast DnaJ homologues. Therefore, it is suggested that human hsp40 is a homologue of yeast SIS1 protein.

Human HDJ2(HSDJ) has been suggested to be a homologue of yeast YDJ1 and plant ANJ1 based on the amino acid sequence, domain organization and biochemical properties [7, 11, 12 and TABLE 1]. Another human homologue, HSJ1, seems not to correspond to any yeast DnaJ homologues. In a forthcoming study, we hope to elucidate the unidentified human DnaJ homologues corresponding to each yeast homologue.

ACKNOWLEDGMENTS

We thank Dr. Nishida for the encouragement and critical reading of the manuscript, and all of our laboratory staffs for cloning hsp40 cDNA. We also thank Mr. Y. Minoura, Biophysics Unit, for synthesizing oligonucleotide probes and Mrs. S. Tokumasu for preparation of the manuscript. This study was supported in part by Grants-in-Aid for Cancer Research (04151035 and 05151039) and for General Scientific Research (04304029 and 03680186) from the Japanese Ministry of Education, Science and Culture, and for Project Research Study from Aichi Cancer Center Research Institute.

REFERENCES

1. Blumberg, H., and Silver, P.A. (1991) Nature 349, 627-630.
2. Caplan, A.J., and Douglas, M.G. (1991) J. Cell Biol. 114, 609-621.
3. Atencio, D.P., and Yaffe, M.P. (1992) Mol. Cell. Biol. 12, 283-291.
4. Luke, M.M., Sutton, A., and Arndt, K.T. (1991), J. Cell Biol. 114, 623-638.
5. Sadler, I., Chiang, A., Kurihara, T., Rothblat, J., Way, J., and Silver, P. (1989) J. Cell Biol. 109, 2665-2675.
6. Zhang, S., Lockshin, C., Herbert, A., Winter, E., and Rich, A. (1992) EMBO J. 11, 3787-3796.
7. Zhu, J.-K., Shi, J., Bressan, R.A., and Hasegawa, P.M. (1993) Plant Cell 5, 341-349.
8. Zinsmaier, K.E., Hofbauer, A., Heimbeck, G., Pflugfelder, G.O., Buchner, S., and Buchner, E. (1990) J. Neurol. 7, 15-29.
9. Raabe, T., and Manley, J.L. (1991) Nucl. Acid Res. 19, 6645.
10. Cheetham, M.E., Brion, J.-P., and Anderton, B.H. (1992) Biochem. J. 284, 469-476.
11. Chellaiah, A., Davis, A., and Mohanakumar, T. (1993) Biochem. Biophys. Acta 1174, 111-113.
12. Oh, S., Iwahori, A., and Kato, S. (1993) Biochem. Biophys. Acta 1174, 114-116.
13. Georgopoulos, C. (1992) Trends Biochem Sci. 17, 295-299.
14. Cyr, D.M., Lu, X., and Douglas, M.G. (1992) J. Biol. Chem. 267, 20927-20931.
15. Caplan, A.J., Cyr, D.M., and Douglas, M.G. (1992) Cell 71, 1143-1155.
16. Deshaies, R.J., Sanders, S.L., Feldheim, D.A., and Schekman, R. (1991) Nature 349, 806-808.
17. Zhong, T., and Arndt, K.T. (1993) Cell 73, 1175-1186.
18. Caplan, A.J., Cyr, D.M., and Douglas, M.G. (1993) Mol. Biol. Cell 4, 555-563.
19. Silver, P.A., and Way, J.C. (1993) Cell 74, 5-6.
20. Ohtsuka, K., Masuda, A., Nakai, A., and Nagata, K. (1990) Biochem. Biophys. Res. Comm. 166, 642-647.

21. Hattori, H., Liu, Y.-C., Tohnai, I., Ueda, M., Kaneda, T., Kobayashi, T., Tanabe, K., and Ohtsuka, K. (1992) Cell Struct. Funct. 17, 77-86.
22. Hattori, H., Kaneda, T., Lokeshwar, B., Laszlo, A., and Ohtsuka, K. (1993) J. Cell Sci. 104, 629-638.
23. Chomczynski, P., and Sacchi, N. (1987) Anal. Biochem.162, 156-159.
24. Ohki, M., Tamura, F., Nishimura, S., and Uchida, H. (1986) J. Biol. Chem. 261, 1778-1781.
25. Bardwell, J.C.A., Tilly, K., Craig, E., King, J., Zylicz, M., and Georgopoulos, C. (1986) J. Biol. Chem. 261, 1782-1785.
26. Bork, P., Sander, C., and Valencia, A. (1992) Trends Biochem. Sci. 17, 129.