Published in final edited form as:

Parasitol Int. 2002 March ; 51(1): 9–16.

Mitochondrial-type hsp70 genes of the amitochondriate protists, *Giardia intestinalis*, *Entamoeba histolytica* and two microsporidians[☆]

Nobuko Arisue^a, Lidya B. Sánchez^b, Louis M. Weiss^{c,d}, Miklós Müller^b, and Tetsuo Hashimoto^{a,b,e,*}

^aDepartment of Biosystems Science, Graduate University for Advanced Studies (Sokendai), Hayama, Kanagawa 240-0193, Japan

^bThe Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

^cDepartment of Medicine, Albert Einstein College of Medicine, Moris Park Avenue, Bronx, NY 10461, USA

^dDepartment of Pathology, Albert Einstein College of Medicine, Moris Park Avenue, Bronx, NY 10461, USA

^eThe Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan

Abstract

Genes encoding putative mitochondrial-type heat shock protein 70 (mit-hsp70) were isolated and sequenced from amitochondriate protists, *Giardia intestinalis*, *Entamoeba histolytica*, and two microsporidians, *Encephalitozoon hellem* and *Glugea plecoglossi*. The deduced mit-hsp70 sequences were analyzed by sequence alignments and phylogenetic reconstructions. The mit-hsp70 sequence of these four amitochondriate protists were divergent from other mit-hsp70 sequences of mitochondriate eukaryotes. However, all of these sequences were clearly located within a eukaryotic mitochondrial clade in the tree including various type hsp70 sequences, supporting the emerging notion that none of these amitochondriate lineages are primitively amitochondrial, but lost their mitochondria secondarily in their evolutionary past.

Keywords

Hsp70; Amitochondriate protists; Phylogeny; *Giardia*; *Entamoeba*; Microsporidia

1. Introduction

Evidence for secondary loss of mitochondrial function in four major amitochondriate protist lineages, diplomonads, parabasalids, entamoebidae and microsporidia has been obtained in recent years by the detection of mitochondrion-related chaperon genes that are coded in the nuclear DNA (for reviews see refs. [1–5]). These include chaperonin (cpn60) in *Entamoeba*

[☆]Nucleotide sequence data reported in this paper have been submitted to the GenBank™, EMBL and DDBJ databases with the accession numbers AB056658, AB056659, AB056660, AB056661 and AB067649.

© 2002 Elsevier Science Ireland Ltd. All rights reserved

*Corresponding author. Tel.: +81-35421-8773; fax: +81-33446-1695. hashimoto@ism.ac.jp (T. Hashimoto)..

After submission of this manuscript, the *Giar-dia* dnaK homolog sequence was published [32] and deposited in GenBank (AF274582). The sequence of its ORF is identical to our genomic DNA sequence (AB056659. with the exception of a synonymous substitution on the third codon position.

histolytica (entamoebidae), *Trichomonas vaginalis* (parabasalids), and *Giardia intestinalis* (diplomonads) and mitochondrial-type heat shock protein 70 (mit-hsp70) in *T. vaginalis* and three microsporidians, *Nosema locustae*, *Vairimorpha necatrix*, and *Encephalitozoon cuniculi*. In addition to these chaperon genes, two other mitochondrion-related genes have also been reported recently. One is encoding ADP–ATP carrier proteins in *T. vaginalis* [6] and the other is *iscS* in *T. vaginalis* and *G. intestinalis* [7].

Phylogenetic reconstruction of each of these proteins demonstrated without exception that sequences from amitochondriate lineages are clearly located within the corresponding eukaryotic mitochondrial clades. Although the functions of mitochondrial-type chaperons in amitochondriate protists still have to be established, cpn60 and mit-hsp70 were shown to be localized to the hydrogenosome in *T. vaginalis* [8], while in *E. histolytica*, cpn60 is localized in a recently identified double membrane bounded organelle, called mitosome [9] or crypton [10]. This evidence suggests a common evolutionary origin of these organelles with the mitochondrion [11–13], although this view is not universally accepted [14,15].

Since until very recently mit-hsp70 sequences were reported only from two amitochondriate lineages, parabasalids and microsporidia, we isolated mit-hsp70 genes from the other amitochondriate lineages, diplomonads and entamoebidae, in order to analyze their evolutionary relationships. Here we report complete sequences of the genes putatively encoding mit-hsp70 from *G. intestinalis*, *E. histolytica*, and a microsporidian, *Encephalitozoon hellem*, and a partial one containing most of the mit-hsp70 open reading frame (ORF) from another microsporidian, *Glugea plecoglossi*. While this work was under way, a mit-hsp70 sequence became available from *E. histolytica* [16]. Our phylogenetic analysis demonstrated that all mit-hsp70 sequences from these organisms are located within a eukaryotic mitochondrial clade, supporting the emerging notion that none of these amitochondriate lineages are primitively amitochondrial, but lost their mitochondria secondarily in their evolutionary past.

2. Materials and methods

2.1. Cloning and sequence analysis

Giardia intestinalis, strain WB, clone 6 (ATCC 30957), *Entamoeba histolytica*, strain HM-1:IMSS (ATCC 30459), *Encephalitozoon hellem* (ATCC 50451) isolated from the urine of an adult patient with AIDS (gift of Dr E.S. Didier, LA, NY), and *Glugea plecoglossi* collected from infected Ayu fishes in the Biwako Lake in Japan were used. Genomic DNA was extracted from *G. intestinalis* and *E. histolytica* with the use of a blood and culture DNA kit (QIAGEN, Chatsworth, CA, USA) according to the manufacturer's protocol. Extraction of the two microsporidian genomic DNA was described previously; *E. hellem* [17] and *G. plecoglossi* [18]. Using degenerate primers synthesized based on the conserved amino acid residues in the alignment of hsp70 sequences and unique primers synthesized based on the partial sequencing results, major parts of the mit-hsp70 genes were amplified by polymerase chain reaction (PCR), and the corresponding products were cloned, and sequenced. The amino acid and the corresponding nucleotide sequences used for synthesizing the degenerate primers are: *G. intestinalis*, GIDLGTN (forward) (5'-GGNATHGAYYTNGGNACNACNAA-3'), and F(D/E)(I/V)DANG (reverse) (5'-CCRTTNGCRTCNAAYNTCRAA-3'); *E. histolytica*, GIGLGTT (forward) (5'-GGNATHGGNYTNGGNACNAC-3'), V(F/Y)DLGGGT (forward) (5'-GTNTWYGAYYTNGGNGGNAC-3'), MTRMPLVQ (reverse) (5'-TGNACNARNGGCATNCKNGTCAT-3'), and EVTFDIDA (reverse) (5'-GCRTCDAATRCRAANGTNACYTC-3'); *E. hellem*, V(F/Y)DLGGGT (forward) and MTRMPLVQ (reverse); *G. plecoglossi*, GIGLGTT (forward), V(F/Y)DLGGGT (forward), MTRMPLVQ (reverse), and FEA(A/S)DANG(M/I) (reverse) (5'-

ATNCCRTTNGCRTCNSHNGCYTCRAA-3'). An uneven PCR method was used to obtain N- or C-terminal portion of each gene [19]. Specific primers used for the uneven PCR were based on the sequence information already obtained, while arbitrary primers were synthesized based on the sequences described in Chen and Wu [19]. Plasmid vectors, pT7 (Novagen, Madison, WI, USA) and pCR2.1 (Invitrogen, Carlsbad, CA, USA), were used for cloning the PCR amplified bands. A ZAP express genomic DNA library of *E. hellem* [17] was also screened with a PCR based fragment as a probe to isolate the C-terminal portion of the gene. Nucleotide sequences were determined on both strands with the use of an automated DNA sequencer (ABI model 310 Genetic Analyzer, Perkin-Elmer Cetus, Norwalk, CT, USA).

In addition, *G. intestinalis* cDNA encoding mit-hsp70 was cloned and sequenced. Total RNA of *G. intestinalis* was extracted with the use of Tri Reagent LS-RNA/DNA/Protein isolation reagent (Molecular Research Center, Cincinnati, OH, USA), according to the manufacturer's instructions. Total RNA was treated with Super-Script™ II RNase H⁻ Reverse Transcriptase (RT) (GIBCO BRL, Life Technologies, Rockville, MD), and the resultant RNA:cDNA duplex was used as template for RT-PCR analysis. A set of primers that covers an entire part of the ORF was synthesized based on the sequence information of the genomic DNA. The amplified 2.0 kb fragment was cloned and sequenced as described above.

2.2. Sequence alignment and phylogenetic reconstruction

The deduced mit-hsp70 sequences of the four amitochondriate protists were aligned with all hsp70 homologs currently available in the database by manual editing. Based on the alignment, 375 unambiguously aligned positions were used for phylogenetic analyses. Selected positions are (*G. intestinalis* numbering): 17–22, 26–27, 29–47, 57–58, 68–83, 117–146, 150–189, 208–222, 226–254, 277–280, 282–286, 322–352, 356–408, 413–418, 420–525 and 527–537. The maximum likelihood method of protein phylogeny was applied for the data set by using PROTML program in MOLPHY version 2.3 [20], since the method was shown to be the most robust against the violation of rate constancy [21,22], and since evolutionary rates for different lineages in the present data set showed extreme variation. The JTT-F model was used for amino acid replacement process. Since the number of operational taxonomic units to be analyzed was large, topology search for obtaining the best tree was performed by the use of the local rearrangement (–R) and the quick stepwise OTU addition procedure (–q –n2000) options of the PROTML program. A neighbor-joining tree was used as an initial tree for the –R option analysis. Alternative tree topologies produced by the search step were used for the RELL bootstrap analysis [23].

3. Results and discussion

3.1. Mit-hsp70 sequences of four amitochondriate protists

Putative mit-hsp70 ORFs were 1923 bp (640 aa), 1797 bp (598 aa), and 1782 bp (593 aa), respectively for *G. intestinalis*, *E. histolytica*, and *E. hellem*. A truncated reading frame of 1542 bp (513 aa) probably lacking only a few N-terminal residues, was obtained for *G. plecoglossi*. All of these sequences were not interrupted by intronlike sequences. A cDNA sequence encoding *G. intestinalis* mit-hsp70 ORF was the same as the genomic DNA sequence with the exception of a synonymous substitution on the third codon position. No introns were seen in the ORF, being congruent with the finding that none of the reported *G. intestinalis* sequences have introns [24]. A BLAST search [25] revealed high similarity of the deduced amino acid sequences to mit-hsp70 and bacterial-type hsp70 (dnaK) sequences, identifying these as mit-hsp70 homologs.

The sequence of the putative ORF of the *E. histolytica* (strain HM-1:IMSS. mit-hsp70 was 98.5% identical to a mit-hsp70 ORF of the same strain, recently reported [16]. The two

sequences differed at 10 substitutions and three insertions, resulting in the putative translations in eight replacements and an insertion. The marked difference indicated the presence of different copies. A BLAST search of the genome sequencing project database of the same *E. histolytica* isolate (<http://www.tigr.org/tdb/edb2/enta/htmls/>) demonstrated that at least two different mit-hsp70 sequences exist in the database and that both the above two ORFs are likely to be present in the database. However, these two sequences may represent allelic heterogeneity of the same locus, as already mentioned by Bakatselou et al. [16], because 5'- and 3'-non-coding regions of these two sequences were very similar with each other, and the level of conservation seems to be unusual for different copy genes. Approximately 80 bp upstream of the initiation codon of the *E. hellem* mit-hsp70 gene, a putative tryptophanyl-tRNA synthetase was coded on the opposite strand. A comparable gene organization is seen in another microsporidian, *Encephalitozoon cuniculi* [26].

3.2. Comparison of hsp70 sequences

Deduced mit-hsp70 amino acid sequences obtained in this study were aligned with other mithsp70 sequences and bacterial-type, chloroplast-type, cytosolic-type and endoplasmic reticulum (ER)-type hsp70 sequences (data not shown, the alignment is available from T.H. upon request). On the N-terminal portion, short extensions of 9, 12 and 16 residues, were found in *G. intestinalis*, *E. histolytica*, and *E. hellem* sequences, respectively, just 5' of the highly conserved motif GIDLGT. No similarity was detected among these three extension sequences. However, the *E. hellem* extension was very similar to that seen in *E. cuniculi* [26], and the *E. histolytica* extension was similar to the N-terminal sequences of some *T. aginalis* hydrogenosomal proteins, as also noted by Bakatselou et al. [16]. Two mit-hsp70 signatures recognized at first by Germot et al. [27] were present in all the sequences reported here (Fig. 1). The first motif (GDAWV) is shared by mitochondrial and proteobacterial hsp70s, and the second motif (YSPSQI) is generally conserved in mit- and α -proteobacterial-hsp70s. Several conservative amino acid replacements were found in the corresponding motifs of the sequences reported here, and the first motif of *G. plecoglossi* and both motifs of *G. intestinalis* were quite divergent from the consensus.

The mit-hsp70 sequences of *G. intestinalis* and *G. plecoglossi* are most divergent of all the hsp70 homologs currently available. In addition, the *G. intestinalis* sequence contained several short insertions that are not found in other mit-hsp70s and other type hsp70 sequences, while the *G. plecoglossi* sequence had a long deletion on the C-terminal portion. The *E. histolytica* mit-hsp70 sequence was also divergent, but no large insertion/deletion was found in the sequence. The *E. hellem* mit-hsp70 sequence showed high similarity with the *E. cuniculi* homolog, and neither the *Encephalitozoon* homolog was as divergent as other mit-hsp70 sequences of microsporidia.

After elimination of the positions that could not be aligned reliably, 375 positions were used for the phylogenetic analysis. A difference matrix of these positions for 11 sequences is shown in Table 1. The pairwise differences revealed that mit-hsp70 sequence of amitochondriate protists, *G. intestinalis*, *E. histolytica* and microsporidia are highly divergent, in agreement with the alignment features described above. The differences between *Homo sapiens* and amitochondriate protists excluding *T. vaginalis* ranged from 34.1 to 50.7%. These are larger than the differences between *H. sapiens* and proteobacteria.

3.3. Phylogenetic reconstruction analysis of mithsp70 sequences

Preliminary phylogenetic analysis including 92 hsp70 homologs of various type clearly demonstrated that the mit-hsp70 sequences of amitochondriate protists reported here are located within the clade comprising of mit-hsp70 sequences of other eukaryotes and that the sister group of the mit-hsp70 clade is shared by α -proteobacteria (data not shown). Based on

the preliminary analysis, we selected 24 hsp70 sequences for a more detailed phylogenetic analysis. This data set included most of the mit-hsp70 sequences except for very similar ones and six proteobacterial hsp70s.

Fig. 2 demonstrates the best tree of the maximum likelihood analysis using the PROTML program [20]. The tree clearly demonstrates that all the mit-hsp70 sequences from amitochondriate protists are monophyletic with other eukaryotic mit-hsp70 sequences with a high bootstrap proportion value (97%) confirming the mitochondrial-type features of these sequences. Within the mit-hsp70 clade, monophyly of animalia, fungi, microsporidia, plantae, and of euglenozoa are clearly reconstructed. The monophyletic origin of and the branching order within microsporidia are also clearly resolved, but the close relationship of *N. locustae* with *G. plecoglossi* is incongruent with the small subunit (SSU) rRNA phylogeny, in which *Nosema* and *Varimorpha* are the closest relatives [28,29]. However, since no SSU rRNA data is available from *N. locustae* and since monophyly of the genus *Nosema* has yet been established in the SSU rRNA phylogeny [28,29], we cannot simply compare the two trees.

The monophyly of three amitochondriate protists lineages, *G. intestinalis*, *E. histolytica* and microsporidia, is clearly supported by the high bootstrap proportion value (96%) in disagreement with the accepted eukaryotic phylogeny [30]. Since branch lengths of the tree leading to these amitochondriate protists are very long compared to those leading to the other eukaryotes, the positions of these three lineages are possibly affected by a long branch attraction of the phylogenetic reconstruction analysis [31]. Especially for *G. plecoglossi* and *G. intestinalis*, the lengths are extremely long, being comparable with the highly divergent sequence features of these organisms. The long branch lengths of these amitochondriate lineages suggest that the functional constraints have been relaxed. The function of mit-hsp70 might have been changed during the evolution of these organisms.

Acknowledgments

We thank Dr H.S. Zhang for extracting genomic DNA of *Encephalitozoon hellem* and for preparing its genomic DNA library, Dr G. Wu, Dr M. Hasegawa and Dr T. Yano for helpful discussions, and A. Deguchi and S. Kikuchi for technical assistance. This work was carried out under the ISM Co-operative Research Program (00ISM•CRP-2042, and 01ISM•CRP-2038) and was supported by grants from the Japanese Society for the Promotion of Science (10044219 and 13640709) to T.H. Work carried out at the Rockefeller University in New York was supported by the USPHS National Institutes of Health Grant (AI11942) to M.M. Work carried out at the Albert Einstein College of Medicine in New York was supported by the USPHS National Institutes of Health Grant (AI31788) to L.M.W. Visits of T.H. to the New York laboratories and of L.B.S. to the Hayama laboratory were supported by the US-Japan Co-operative Research Project by National Science Foundation and Japanese Society for the Promotion of Science (INT-9726707). Preliminary sequence data for *E. histolytica* is deposited regularly into the GSS division of Gen-Bank. The sequencing effort is part of the International *Entamoeba* Genome Sequencing Project and is supported by an award from the National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Abbreviations

Mit-hsp70	mitochondrial-type heat shock protein 70
ORF	open reading frame
PCR	polymerase chain reaction
RT	reverse transcriptase
SSU	small subunit

References

- [1]. Roger AJ. Reconstructing early events in eukaryotic evolution. *Am Nat.* 1999; 154:S146–63. [PubMed: 10527924]
- [2]. Müller M. Evolutionary origin of hydrogenosomes. *Parasitol Today.* 1997; 13:166–7. [PubMed: 15275085]
- [3]. Müller M. What are the microsporidia? *Parasitol Today.* 1997; 13:455–6. [PubMed: 15275131]
- [4]. Müller M. A mitochondrion in *Entamoeba histolytica*? *Parasitol Today.* 2000; 16:368–9. [PubMed: 10951594]
- [5]. Embley TM, Hirt RP. Early branching eukaryotes? *Curr Opin Genet Dev.* 1998; 8:624–9. [PubMed: 9914207]
- [6]. Dyall SD, Koehler CM, Delgadillo-Correa MG, et al. Presence of a member of the mitochondrial carrier family in hydrogenosomes: conservation of membrane-targeting pathways between hydrogenosomes and mitochondria. *Mol Cell Biol.* 2000; 20:2488–97.
- [7]. Tachezy J, Sánchez LB, Müller M. Mitochondrial type iron-sulfur cluster assembly in the amitochondriate eukaryotes *Trichomonas vaginalis* and *Giardia intestinalis*, as indicated by the phylogeny of IscS. *Mol Biol Evol.* 2001; 18:1919–28. [PubMed: 11557797]
- [8]. Bozner P. Immunological detection and subcellular localization of Hsp70 and Hsp60 homologs in *Trichomonas vaginalis*. *J Parasitol.* 1997; 83:224–9. [PubMed: 9105301]
- [9]. Tovar J, Fischer A, Clark CG. The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol Microbiol.* 1999; 32:1013–21. [PubMed: 10361303]
- [10]. Mai Z, Ghosh S, Frisardi M, Rosenthal B, Rogers R, Samuelson J. Hsp60 is targeted to a cryptic mitochondrion-derived organelle ('Crypton'). in the microaerophilic protozoan parasite *Entamoeba histolytica*. *Mol Cell Biol.* 1999; 19:2198–205. [PubMed: 10022906]
- [11]. Martin W, Müller M. The hydrogen hypothesis for the first eukaryote. *Nature.* 1998; 392:37–41. [PubMed: 9510246]
- [12]. Dyall SD, Johnson PJ. Origins of hydrogenosomes and mitochondria: evolution and organelle biogenesis. *Curr Opin Microbiol.* 2000; 3:404–11. [PubMed: 10972502]
- [13]. Rotte C, Henze K, Müller M, Martin W. Origins of hydrogenosomes and mitochondria. *Curr Opin Microbiol.* 2000; 3:481–6. [PubMed: 11050446]
- [14]. Sogin ML. Organelle origins: energy-producing symbionts in early eukaryotes? *Curr Biol.* 1997; 7:R315–7. [PubMed: 9115381]
- [15]. Kurland CG, Andersson SGE. Origin and evolution of the mitochondrial proteome. *Microbiol Mol Biol Rev.* 2000; 64:786–820. [PubMed: 11104819]
- [16]. Bakatselou C, Kidgell C, Clark CG. A mitochondrial-type hsp70 gene of *Entamoeba histolytica*. *Mol Biochem Parasitol.* 2000; 110:177–82. [PubMed: 10989156]
- [17]. Keohane EM, Orr GA, Zhang HS, et al. The molecular characterization of the major polar tube gene from *Encephalitozoon hellem*, a microsporidian parasite of humans. *Mol Biochem Parasitol.* 1998; 94:227–36. [PubMed: 9747973]
- [18]. Kamaishi T, Hashimoto T, Nakamura Y, et al. Protein phylogeny of translation elongation factor EF-1 α suggests microsporidians are extremely ancient eukaryotes. *J Mol Evol.* 1996; 42:257–63. [PubMed: 8919877]
- [19]. Chen X, Wu R. Direct amplification of unknown genes and fragments by uneven polymerase chain reaction. *Gene.* 1997; 185:195–99. [PubMed: 9055815]
- [20]. Adachi J, Hasegawa M. MOLPHY version 2.3: programs for molecular phylogenetics based on maximum likelihood. *Comput Sci Monogr.* 1996; 28:1–50.
- [21]. Hasegawa M, Kishino H, Saitou N. On the maximum likelihood method in molecular phylogenetics. *J Mol Evol.* 1991; 32:443–5. [PubMed: 1904100]
- [22]. Hasegawa M, Fujiwara M. Relative efficiencies of the maximum likelihood, maximum parsimony, and neighbor-joining methods for estimating protein phylogeny. *Mol Phylog Evol.* 1993; 2:1–5.

- [23]. Hasegawa M, Kishino H. Accuracies of the simple methods for estimating the bootstrap probability of a maximum-likelihood tree. *Mol Biol Evol.* 1994; 11:142–5.
- [24]. Adam RD. The *Giardia lamblia* genome. *Int J Parasitol.* 2000; 30:475–84. [PubMed: 10731570]
- [25]. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997; 25:3389–402. [PubMed: 9254694]
- [26]. Peyretilade E, Broussolle V, Peyret P, Méténier G, Gouy M, Vivarés CP. Microsporidia, amitochondriate protists, possess a 70-kDa heat shock protein gene of mitochondrial evolutionary origin. *Mol Biol Evol.* 1998; 15:683–9. [PubMed: 9615449]
- [27]. Germot A, Philippe H, Le Guyader H. Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. *Proc Natl Acad Sci USA.* 1996; 93:14614–17. [PubMed: 8962101]
- [28]. Weiss LM, Vossbrinck CR. Microsporidiosis: molecular and diagnostic aspects. *Adv Parasitol.* 1998; 40:351–95. [PubMed: 9554079]
- [29]. Weiss, LM.; Vossbrinck, CR. Molecular biology, molecular phylogeny, and molecular diagnostic approaches to the microsporidia. In: Wittner, M.; Weiss, LM., editors. *Microsporidiosis and the Microsporidia*. American Society of Microbiology; Norwalk: 1999. p. 129-71.
- [30]. Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF. A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science.* 2000; 290:972–7. [PubMed: 11062127]
- [31]. Philippe H, Laurent J. How good are deep phylogenetic trees? *Curr Opin Genet Dev.* 1998; 8:616–23. [PubMed: 9914208]
- [32]. Morrison HG, Roger AJ, Nystul TG, et al. *Giardia lamblia* expresses a proteobacterial-like DnaK homolog. *Mol Biol Evol.* 2001; 18:530–41. [PubMed: 11264404]

		130	140	150	160	170
	<i>Homo sapiens</i> (Mit)	GRRYDDPEVQKDIKNVPPKIVRA--	SNGDAMV --	EAHGKL	YSPSQI	GAFVL
	<i>Saccharomyces cerevisiae</i> (Mit)	GRRFEDAEVQRDINQVPPKIVKH--	SNGDAMV --	EARNRT	YSPAQI	GGFVL
	<i>Arabidopsis thaliana</i> (Mit)	GRRFDDPQTQKEMKMPYKIVKA--	PNGDAMV --	EANGKF	YSPSQI	GAFVL
	<i>Trypanosoma cruzi</i> (Mit)	GRRFEDSNIQHDIKNVPPYKIGRS--	SNGDAMVQ --	DANGKF	YSPSQI	GAFVL
Eukaryota	<i>Trichomonas vaginalis</i> (Hyd)	GCSFDDMTKKCREMVPYQIVKA--	KSGDAMVK --	DEKGNK	YSPSQI	GAFVL
	<i>Glugea plecoglossi</i>	GRSY--AEVQDYIKKLPPDVVK--	EKEDVKI --	KVQGNK	YSPSKI	SSFIL
	<i>Encephalitozoon hellem</i>	GRKLEDPEVKKYVKGLPFDTS--	CNGDVWI --	KVDGKK	YSPAQI	GGFVL
	<i>Entamoeba histolytica</i>	GCQY--KNVEQQ--RKFYKIVEG--	RNGEGWI --	YINGKT	YSPSEI	SSFIL
	<i>Giardia intestinalis</i>	GRRFTDAEVKRAAKHVSVDIVQG--	PNGEAMIN	VNLHKS	YSPSEI	SSFIL
Proteobacteria	α <i>Rickettsia prowazekii</i>	GRNFTDPMVRKDQGLVPYNIVKA--	DNGDAMV --	EADNHK	YSPSQI	SAFIL
	β <i>Burkholderia cepacia</i>	GRRFEEQVQRDQGLMPYSIIKA--	DNGDAMVE --	GHGKLM	APPQVSAEAV	
	γ <i>Escherichia coli</i>	GRRFQDEEVQRDVSIMPFKIIAA--	DNGDAMV --	EVKGQK	MAPPQISAEVL	
Eukaryota	<i>Pisum sativum</i> (Chl)	GRKM--SEVDEESKQVSRYRIRDDNGNVKLC--	PAIGSF	AAEIS	SAQVL	
Cyanobacteria	<i>Synechocystis</i> sp.	GRRW--EDTEAERSRVTYTCVPG--	KDDTVNV --	TIRDRFCT	PQEISAMVL	
Gram+ bacteria	<i>Bacillus subtilis</i>	GTDY-----	KV --	EIEGKDY	TPQEVSAILL	
	<i>Mycobacterium tuberculosis</i>	GSDW-----	SI --	EIDGKKY	TAPEISARIL	
Mycoplasmataceae	<i>Mycoplasma capricolum</i>	GTTS-----	KV --	NLEGKDY	SPEQISAEIL	
Archaeobacteria	<i>Methanobacterium thermoautotrophicum</i>	GTDR-----	KV --	KVHGKEY	TPQEISAFIL	

Fig. 1.

Two signature motifs found in mit-hsp70 sequences of eukaryotes including amitochondriate protists and in homologous proteobacterial sequences. Amino acid residues within the originally defined motifs[27], GDAWV and YSPSQI, are shown in bold face characters. Position numbers for the *H. sapiens* mit-hsp70 sequence were shown above the alignment. Localization within cell compartments was shown in parentheses: Mit, mitochondria; Hyd, hydrogenosome; and Chl, chloroplasts.

Fig. 2. Phylogenetic tree of mit-hsp70 homologs using proteobacterial hsp70s as outgroups. The best tree obtained by the maximum likelihood analysis of protein phylogeny using the PROTML program is shown. The horizontal length of each branch is proportional to the estimated number of substitutions. Bootstrap proportions are attached to the internal branches; * denotes amitochondriate protists. Mit, mitochondria; and Hyd, hydrogenosome. Accession numbers of the sequences are shown in parentheses.

Table 1

Percentage pairwise amino acid difference between mitochondrial-type and proteobacterial-type hsp70s of various species (375 shared residues)

Species	Percentage difference									
	H.s.	N.l.	G.p.	V.n.	E.eu.	E.he.	E.hi.	G.i.	T.v.	R.p.
<i>Homo sapiens</i>										
Microsporidia										
<i>Nosema locustae</i>	40.3									
<i>Glugea plecoglossi</i>	46.7	42.4								
<i>Vairimorpha necatrix</i>	38.1	32.5	43.5							
<i>Encephalitozoon cuniculi</i>	34.7	32.0	43.7	26.9						
<i>Encephalitozoon hellem</i>	34.1	43.7	43.5	24.3	10.7					
<i>Entamoeba histolytica</i>	38.9	53.1	48.5	43.7	42.4	42.7				
<i>Giardia intestinalis</i>	50.7	41.3	53.6	53.3	48.8	47.7	54.4			
<i>Trichomonas vaginalis</i>	27.5	39.7	47.7	42.4	38.1	38.4	40.0	50.1		
Proteobacteria										
<i>Rickettsia prowazekii</i>	23.7	42.7	46.4	37.9	35.5	35.7	40.0	51.2	29.1	
<i>Escherichia coli</i>	30.4	44.0	52.0	41.9	38.1	37.6	40.0	51.7	35.2	26.1