Characterization of chaperonin 10 (Cpn10) from the intestinal human pathogen *Entamoeba histolytica*

Mark van der Giezen,† Gloria León-Avila‡ and Jorge Tovar

*Entamoeba histolytica* is the causative agent of amoebiasis, a poverty-related disease that kills an estimated 100 000 people each year. *E. histolytica* does not contain ‘standard mitochondria’, but harbours mitochondrial remnant organelles called mitosomes. These organelles are characterized by the presence of mitochondrial chaperonin Cpn60, but little else is known about the functions and molecular composition of mitosomes. In this study, a gene encoding molecular chaperonin Cpn10 – the functional partner of Cpn60 – was cloned, and its structure and expression were characterized, as well as the cellular localization of its encoded protein. The 5′ untranslated region of the gene contains all of the structural promoter elements required for transcription in this organism. The amoebic Cpn10, like Cpn60, is not significantly upregulated upon heat-shock treatment. Computer-assisted protein modelling, and specific antibodies against Cpn10 and Cpn60, suggest that both proteins interact with each other, and that they function in the same intracellular compartment. Thus, *E. histolytica* appears to have retained at least two of the key molecular components required for the refolding of imported mitosomal proteins.

INTRODUCTION

The human parasite *Entamoeba histolytica* was once considered as an example of a ‘primitive’ eukaryote since it lacks typical features of ‘higher’ eukaryotes, such as mitochondria, peroxisomes, Golgi apparatus or rough endoplasmic reticulum (Meza, 1992). *Entamoeba*, along with all other amitochondriate eukaryotes, was thought to have diverged from the main eukaryotic trunk before the mitochondrial endosymbiont became established. Their putative primitive nature appeared to be supported by rRNA phylogenies that placed the microsporidia (e.g. *Vairimorpha*, *Trachipleistophora*, *Encephalitozoon*), diplomonads (e.g. *Hexamita*, *Sporonucleus*, *Giardia*) and parabasalids (e.g. *Tririchomonas*, *Trichomonas*) as the deepest, most ancestral branches of the evolutionary tree (Sogin, 1989). However, the proposed pre-mitochondrial nature of *Entamoeba* did not receive support from rRNA phylogenetic reconstructions, as it branched after well-established mitochondrial groups (Sogin, 1991). This placement suggested the possibility that *Entamoeba* might have mitochondrial ancestors.

Over the past few years, several genes of mitochondrial ancestry have been identified in the genomes of ‘amitochondrial’ eukaryotes (reviewed by Embly & Hirt, 1998), and the reliability of rRNA phylogenies to resolve early evolutionary events has been brought into question (Gribaldo & Philippe, 2002). The discovery of mitochondrial genes in amitochondrial protists led to the subsequent identification of mitosomes, intracellular compartments housing mitochondrial proteins that are thought to represent mitochondrial remnants. Originally discovered in *E. histolytica* (Mai et al., 1999; Tovar et al., 1999), mitochondrial remnant organelles have also been identified in *Trachipleistophora hominis* (Williams et al., 2002), *Giardia intestinalis* (Tovar et al., 2003) and *Cryptosporidium parvum* (Riordan et al., 2003), and it is becoming apparent that their distribution within the protist kingdom is widespread. Thus, it is clear that the absence of classic mitochondria in these organisms is a secondarily derived condition rather than a primitive trait (van der Giezen et al., 2005).

Because of their recent discovery, little is known about the functions and metabolic composition of mitosomes. Only three genes encoding mitochondrial proteins have been described in *E. histolytica*, i.e. pyridine nucleotide transhydrogenase (PNT), molecular chaperonin 60 (Cpn60) and mitochondrial heat-shock protein 70 (mHsp70) (Bakatselou et al., 2000; Clark & Roger, 1995; Tovar et al., 1999). All these proteins seem to contain a mitochondrial-like targeting presequence at their amino terminus. The
putative targeting signal on the Cpn60 has been shown to be required for targeting the protein into the *E. histolytica* mitosome. When the targeting sequence was deleted, the protein accumulated in the cytosol, a mutant phenotype that could be reversed by the addition of a functional mitochondrial targeting signal from the *Trypanosoma cruzi* mHsp70 protein (Tovar et al., 1999).

The availability of sequencing data from the recently completed *E. histolytica* genome project (Loftus et al., 2005) allows the search for additional remnant mitochondrial genes/proteins, and may provide additional information as to the metabolic capacity and protein composition of the mitosome. Here, we report on the identification and structural characterization of a gene encoding the molecular chaperonin Cpn10 – the functional partner of Cpn60 (Hartl, 1996). We find that the upstream region of the gene contains the promoter elements required for transcription in *E. histolytica*, and that – similar to its functional partner – transcription of *cpn10* is not significantly upregulated by heat shock. Amoebal Cpn10 is compartmentalized, and, like many other Cpn10 proteins, it is targeted into the organelle via a mechanism that does not require an amino-terminal targeting presequence.

**METHODS**

**Micro-organism cultivation, and DNA manipulations.** *E. histolytica* HM-1 : IMSS clone 9 was maintained axenically by subculture in YI-S medium with 15% adult bovine serum, as described by Clark & Diamond (2002). *E. histolytica* genomic DNA was isolated using cetyltrimethylammonium bromide (CTAB), according to Clark (1992). *Escherichia coli* strain DH5α was grown in Luria-Bertani (LB) medium (with 100 μg ampicillin ml⁻¹) at 37°C. Standard recombinant DNA techniques were used, as described elsewhere (Sambrook et al., 1989).

**Cloning and sequencing of the *E. histolytica* cpn10 gene.** The human mitochondrial Cpn10 protein sequence (GenBank accession no. Q04984) was used to search the *E. histolytica* genome sequence database at http://www.sanger.ac.uk/Projects/E_histolytica/. Clones Ent141h02_p1c, Ent066d09_q1c, Ent066d09_p1c and Ent091d08_q1c contained sequences similar to the human mitochondrial *cpn10* gene sequence. Based on an alignment of these clones, oligonucleotide primers complementary to the 5’ and 3’ untranslated regions of the putative *E. histolytica* cpn10 homologue (Ehcpn10_280F, 5’-CCA CCC GAA ACC ATT TCA ACA CG-3’; Ehcpn10_604R, 5’-AAA GAG GAA TAA ACG AAT TTT ATG-3’) were synthesized and used for PCR amplification; *E. histolytica* genomic DNA was used as a template. Amplified DNA products were purified, cloned into the pGEM-T-Easy vector (Promega), and sequenced using a BigDye Terminator v3.1 Cycle Sequencing Kit on an Applied Biosystems 377-96 DNA Sequencer.

**Construction of the expression vector harbouring *E. histolytica* cpn10.** The putative *E. histolytica* cpn10 gene was cloned directionally into the KpnI and EcoRV restriction sites of the *Entamoeba* expression vector pJT1 (Fig. 1a). pJT1 was obtained by removing the *cpn60* gene from construct A (Tovar et al., 1999) – a derivative of expression vector pEHneo/CAT (Hamann et al., 1995). KpnI and EcoRV restriction sites were added at the termini of Ehcpn10 by means of PCR using the primers Ehcpn10_KpnF (5’-aga aga GGT ACC CCC CCA CCA ATT TCA ACA CG-3’) and Ehcpn10_Eco RV-R (5’-ctc tct GAT ATC CAT TTT TGC AAA AAT GTC-3’) (stuffer regions are indicated in lower case, restriction sites in italics, and perfect matches in roman capitals). The resulting construct pJT1-Ehcpn10 (Fig. 1a) was sequenced as described above.

**Parasite transfection.** *E. histolytica* trophozoites were detached from culture tubes by chilling on ice, and collected by centrifugation at 300 g at 4°C for 10 min. The cell pellet was washed twice with ice-cold PBS, and once in cytomix buffer consisting of 10 mM K₂HPO₄/KH₂PO₄ (pH 7.6), 120 mM KCl, 0.5 mM CaCl₂, 25 mM HEPES, 2 mM EDTA and 5 mM MgCl₂ (Nickel & Tannich, 1994). Trophozoites (1 × 10⁷) were resuspended in 0.8 ml cytomix, and supplemented with 4 mM ATP, 10 mM glutathione and 100 μg pJT1-Ehcpn10 DNA in an electroporation cuvette (4 mm gap). Cells were electroporated twice using a Bio-Rad Gene Pulser II with pulse controller and capacitance extender, using 3000 V cm⁻¹.

![Fig. 1. Structure of recombinant constructs and of the 5’-flanking region of the *E. histolytica* cpn10 gene. (a) pJT1 contains a c-myc epitope tag (filled box) in-frame with the *Eco*RV restriction site. The actin and lectin transcriptional regulatory elements present in the plasmid are indicated (see Methods; Hamann et al., 1995). Cloning of the *E. histolytica* cpn10 gene into pJT1 resulted in an in-frame fusion construct tagged at its carboxy terminus (pJT1-Ehcpn10) that was used to generate transgenic parasite lines. (b) The three typical upstream regulatory elements are depicted as shown by Purdy et al. (1996): the putative initiator element, double underlined; the ‘GAAC’ element, grey box; and the putative TATA element, boxed.](image-url)
and 25 μl. Electroporated cells were incubated on ice for 10 min, after which 13 ml fresh culture medium was added. After 48 h incubation at 37 °C, G418 (6 μg ml⁻¹) was added to select for transformants.

**RT-PCR studies on heat-shocked and control cells.** To test for expression, primers Ehcpn10_startF (5'-ATG GCA AAA ATT AAA CCA ACT GGA GAC ATG GTT TTA GTC G-3') and Ehcpn10_stopR (5'-TTA TTC GAT TTT TGC AAA AAT GTC ATG TTG TTT TAA TAA AG-3') were used for cpn10 amplification. In addition, cpn60- and actin-specific primers were used as controls; (Eh_partCpn60281F, 5'- ATG GGA CAA CAA CAG CAA CA -3', and Eh_partCpn60544R, 5'- CAA CAG CAC CAT CTC TTC CA -3'), and Eh_acRF, 5'- ATG GGA GAC GAA GAA GGT CA -3', and Eh_acRR, 5'- AAG CAT TTT CTG TGG ACA AT -3'). Total RNA was isolated using TriPure (Roche). DNase-treated total RNA was used as a template; the RNA was tested for DNA contamination by a standard PCR reaction using the same primers. For heat-use as a template; the RNA was tested for DNA contamination by a GeneFlash gel documentation system (SynGene), and band intensities were quantified by densitometry using ImageQuant of the IQ Solutions software package (Amerham Biosciences).

**Fractionation and Western blotting.** *E. histolytica* trophozoites were fractionated as described previously for *G. intestinalis* extracts were fractionated as described previously (Tovar et al., 2003), followed by SDS-PAGE and Western blotting using a semidry electroblotter. Blots were stained using heterologous (Peitsch, 1997), and manually improved based on an independent Protein database searches using the conceptually translated protein entry EAL51277 (Loftus et al., 2005). A BLAST search of the entire published genome indicated that there is only a single copy of cpn10 present. The 5′ untranscribed region of the amoebal cpn10 gene contains distinct putative promoter elements reported to be typical for *E. histolytica* (Purdy et al., 1996). The initiator region, and the putative TATA and GAAC boxes, are all present within the first 30 bases upstream of the start codon (Fig. 1b), suggesting that the gene is functional. The *E. histolytica* cpn10 gene encodes a protein of 87 aa, with a predicted molecular mass of 9.7 kDa, and a theoretical isoelectric point of 9.3. The G+C content is 31 mol% for the coding region, 21 mol% for the 5′ untranslated region (UTR), and 25 mol% for the 3′ UTR. The mean G+C content of *E. histolytica* genes is 32 mol% based on 241 genes analysed (Nakamura et al., 2000). No introns were present in the ORF. The *E. histolytica* chaperonin contains Pfam domain PF00166, which is typical for this protein family. To test whether the protein assumes a normal three-dimensional conformation, its amino acid sequence was modelled on the solved Cpn10 protein structure from *M. tuberculosis* (Taneja & Mande, 2002). The identity between the two proteins is 26 %, indicating that they belong to the same fold (the GroES fold; Taneja & Mande, 1999). The overall topology of both the monomeric (Fig. 2a) and the heptameric (Fig. 2b) models are quite similar to the solved *M. tuberculosis* structure, suggesting that the *E. histolytica* protein is indeed involved in a chaperonin function. The alignment (Fig. 2c) indicates that deletions in the *E. histolytica* protein occur in connecting loops only, and not in the β strands of this β-barrel protein, further supporting the validity of the model.

**Effect of heat shock on expression of *E. histolytica* Cpn10**

To check if the *E. histolytica* cpn10 gene is expressed, we performed RT-PCR experiments as described in Methods. Using total RNA isolated from *E. histolytica* trophozoites, a band of 0.3 kb was obtained, in agreement with the size

**RESULTS**

**Identification, isolation and characterization of the *E. histolytica* gene encoding Cpn10**

Searches of preliminary data generated by the *E. histolytica* genome project at the Sanger Institute revealed several clones with sequence similarity to the human Cpn10 protein sequence. PCR primers based on these clones, when used in PCR reactions on *E. histolytica* genomic DNA as template, amplified a 324 bp fragment, which was purified, cloned and sequenced. The *E. histolytica* cpn10 nucleotide sequence has been submitted to GenBank under accession number AF513821, and is identical to the nucleotide coding sequence for the conceptually translated protein entry EAL51277 (Loftus et al., 2005). A BLAST search of the entire published genome indicated that there is only a single copy of cpn10 present. The 5′ untranscribed region of the amoebal cpn10 gene contains distinct putative promoter elements reported to be typical for *E. histolytica* (Purdy et al., 1996). The initiator region, and the putative TATA and GAAC boxes, are all present within the first 30 bases upstream of the start codon (Fig. 1b), suggesting that the gene is functional. The *E. histolytica* cpn10 gene encodes a protein of 87 aa, with a predicted molecular mass of 9.7 kDa, and a theoretical isoelectric point of 9.3. The G+C content is 31 mol% for the coding region, 21 mol% for the 5′ untranslated region (UTR), and 25 mol% for the 3′ UTR. The mean G+C content of *E. histolytica* genes is 32 mol% based on 241 genes analysed (Nakamura et al., 2000). No introns were present in the ORF. The *E. histolytica* chaperonin contains Pfam domain PF00166, which is typical for this protein family. To test whether the protein assumes a normal three-dimensional conformation, its amino acid sequence was modelled on the solved Cpn10 protein structure from *M. tuberculosis* (Taneja & Mande, 2002). The identity between the two proteins is 26 %, indicating that they belong to the same fold (the GroES fold; Taneja & Mande, 1999). The overall topologies of both the monomeric (Fig. 2a) and the heptameric (Fig. 2b) models are quite similar to the solved *M. tuberculosis* structure, suggesting that the *E. histolytica* protein is indeed involved in a chaperonin function. The alignment (Fig. 2c) indicates that deletions in the *E. histolytica* protein occur in connecting loops only, and not in the β strands of this β-barrel protein, further supporting the validity of the model.
predicted from the gene sequence (Fig. 3, lane 1). Since, in several other organisms, heat-shock proteins can be induced upon incubation at elevated temperatures, a semi-quantitative comparison of amplified DNAs from control and heat-shocked parasites (incubated at 42 °C for 1 h; Mai et al., 1999) was carried out. No major increase of amplified cpn10 (118 %, n = 3) and cpn60 (116 %, n = 3) DNA was observed upon incubation at 42 °C using this method (Fig. 3), suggesting that these genes are not subjected to heat-shock regulation in E. histolytica. This observation is in contrast to the reported heat-shock inducibility of the E. histolytica cpn60 gene at 42 °C (Mai et al., 1999), but in agreement with our previous findings using Northern and Western dot blotting (Tovar et al., 1999).

Cellular localization of E. histolytica Cpn10

Proteins destined for the mitochondrial lumen usually contain characteristic amino-terminal targeting presequences that are proteolytically removed upon import into the organelle (Pfanner & Truscott, 2002). E. histolytica Cpn60 contains a functional targeting presequence that helps to target the protein into the mitochondrial remnant organelles. Given that Cpn10 is the functional partner of Cpn60, we hypothesized that E. histolytica Cpn10 might also be targeted into mitosomes. Alignment of amoebal Cpn10 with homologues from bacteria and other eukaryotes failed to identify a putative targeting presequence (Fig. 4), suggesting that, if targeted into mitosomes, the amoebal protein may be targeted into mitosomes.

Fig. 2. (a, b) Three-dimensional models of E. histolytica Cpn10, drawn using MOLMOL (Koradi et al., 1996). The E. histolytica Cpn10 putative three-dimensional monomeric (a) and heptameric (b) structures were deduced using the conceptually translated Ehcpn10 sequence. The previously solved crystal structure of M. tuberculosis Cpn10 (PDB accession no. 1HX5; Taneja & Mande, 2002) was used as a template. (c) The E. histolytica Cpn10 sequence was aligned with the M. tuberculosis Cpn10 sequence using DeepView v3.7 sp8 (http://www.expasy.org/spdbv/; Guex & Peitsch, 1997), and manually improved based on an independent CLUSTAL W alignment (Thompson et al., 1994). Predicted β strands, and the Cpn60-binding hydrophobic tripeptide, are indicated.

Fig. 3. RT-PCR analysis of E. histolytica cpn10, using RNA from cells maintained at 37 °C (lane 1) and RNA from cells heat-shocked for 1 h at 42 °C (lane 2). Lanes 3–6, templates as before, but using either cpn60- or actin-specific primers.
imported into the organelle by a targeting mechanism that does not require amino-terminal extensions. This would not be unprecedented, as several other mitochondrial Cpn10 proteins that lack an amino-terminal extension have been reported (Hartman et al., 1992).

To investigate the cellular localization of *E. histolytica* Cpn10, cellular fractions were prepared, blotted onto a solid support, and probed with a heterologous polyclonal antibody raised against the human Cpn10 homologue (Santa Cruz Biotechnology). Using this antibody at a low dilution (1:50), a very faint band of the predicted size was observed in the mixed membrane fraction (not shown). To boost the signal, transgenic *E. histolytica* lines harbouring pJT1-EhCpn10 (Fig. 1a) were generated. These parasites overexpress the amoebal Cpn10 antigen tagged with a c-myc epitope. Immunoblotting of the overexpressing parasite cellular fractions showed cross-reactivity of the antibody with a protein enriched in the mixed-membrane fraction (Fig. 5). The apparent molecular mass of 11 kDa for Cpn10 is identical to the calculated molecular mass of the predicted epitope-tagged gene product. When the antisera raised against the *E. histolytica* Cpn60 (Tovar et al., 1999) was used, Cpn60 was observed in the cell-free extract and in the mitosomal fraction, but not in the cytosolic fraction (Fig. 5). Detection of Cpn10 in the cytosolic fraction is likely to be the result of overexpression and limited organelle import capacity. Despite repeated attempts, the use of a monoclonal anti-tag antibody failed to detect the overexpressed protein in Western blotting and immunomicroscopy experiments. One possible explanation for this observation would be that the tag is not sufficiently exposed on the surface of the heptameric structure for efficient interaction with the antibody; alternatively, codon usage may prevent efficient biosynthesis of the tag in *E. histolytica*. Either way, the high level of Cpn10 overexpression, and its observed distribution in transgenic parasites (Fig. 5), makes it unsuitable for the localization of the native antigen by laser scanning confocal microscopy (León-Avila & Tovar, 2004).

**Phylogenetic analyses of the *E. histolytica* Cpn10**

Bayesian and ML analyses of *E. histolytica* Cpn10 revealed that phylogenetic relations are poorly resolved (Fig. 6). Although major eukaryotic clades receive proper support, their branching order is unresolved. The placement of *E. histolytica* at the base of eukaryotes is probably a reflection of limited sampling, since no ‘deep-branching’ eukaryotes other than *Trichomonas vaginalis* are present. Nonetheless, the *E. histolytica* Cpn10 sequence is part of the large eukaryotic clade, consistent with the Cpn60 phylogeny (Roger et al., 1998; Horner & Embley, 2001). The clustering of *T. vaginalis* with bacteria does not receive statistical support, and is therefore unresolved. The low phylogenetic signal of small proteins such as Cpn10 has been suggested previously (Fast et al., 2002).

**DISCUSSION**

Until recently, the origin of eukaryotes was explained by a simple model that assumed the gradual acquisition of typical eukaryotic traits that led from a simple bacterial-like cellular architecture to a more complex eukaryotic one (Margulis, 1970). This view was based on the existence of contemporary microbial eukaryotes that contain a nucleus, but lack double-membrane-bounded organelles, such as mitochondria and chloroplasts. The identification of genes encoding putative mitochondrial proteins (e.g. Cpn60, mtHsp70) in the genomes of several amitochondrial eukaryotes (Bui et al., 1996; Clark & Roger, 1995; GERMOT et al., 1996; Hashimoto et al., 1998; Horner et al., 1996; Roger et al., 1996, 1998), and the cellular localization of their encoded proteins by immunomicroscopy, led to the discovery of organelles of mitochondrial ancestry – collectively known as mitosomes (MAI et al., 1999; RIORDAN et al., 2003; TOVAR et al., 1999, 2003; WILLIAMS et al., 2002) – and to the demonstration that trichomonad hydrogenosomes – organelles discovered in the early 1970s – shared a common ancestry with mitochondria (reviewed by van der Giezen et al., 2005).
The exact protein complement of mitochondrion-related organelles remains to be determined. In *E. histolytica*, candidates to make up the mitosome are still limited: the molecular chaperonins Cpn60 (Clark & Roger, 1995; Tovar et al., 1999) and Hsp70 (Bakatselou et al., 2000), pyridine nucleotide transhydrogenase (PNT) (Clark & Roger, 1995), and possibly IscU and IscS (Ali et al., 2004; van der Giezen et al., 2004). Only Cpn60 has actually been shown to reside inside the mitosome, while evidence for the other proteins is circumstantial. Screening of the *E. histolytica* genome (http://www.sanger.ac.uk/Projects/E_histolytica) identified a gene encoding a putative homologue of the Cpn60 co-chaperonin Cpn10. Typical *E. histolytica* promoter elements seem to be present in the 5' UTR. *E. histolytica* is quite unique in having three core promoter elements: apart from the standard TATA-like sequence and a putative initiator, there is a third GAAC-element of unknown function that is also present in the upstream region of *cpn10*. Heat shock did not induce significantly higher expression levels of either *cpn10* or *cpn60* (Fig. 3). This is in contrast to earlier observations by Mai et al. (1999), who reported increased production of Cpn60 in parasites heat-shocked at 42 °C, but it is in agreement with findings by Tovar et al. (1999), who did not detect a significant increase in Cpn60 expression upon heat shock using Northern and Western blotting.

Co-localization of Cpn10 with Cpn60 is suggested by immunolabelling of cellular fractions, a distribution consistent with their functional partnership in protein folding (Hartl, 1996). Cpn60 and Cpn10 form large multimeric protein complexes consisting of two stacked heptameric rings of Cpn60, and a smaller single heptameric ring of Cpn10 subunits (Hartl, 1996). A hydrophobic tripeptide demonstrated to be involved in the interaction between Cpn10 and Cpn60 (Richardson et al., 2001) is conserved in...
the *Entamoeba histolytica* Cpn10 (Fig. 2c). The large Cpn60-barrel provides an enclosed folding compartment, but the association of Cpn10 with this Cpn60 structure induces conformational changes that actually allow unfolded proteins to bind. The presence of the interacting partner of the *E. histolytica* Cpn60 (which is known to be localized inside the mitosomes) suggests that chaperonin-assisted protein folding does occur in this organelle. Since chaperonin-assisted protein folding is an ATP-dependent process, our data suggest that ATP is either imported into the mitosomes or generated inside the organelle. Current evidence suggests that the former process, and not the latter, operates in *E. histolytica* mitosomes. An unusual member of the ADP/ATP mitochondrial carrier family of transporters has been recently identified and characterized in this organism (Chan et al., 2005). In contrast to typical ADP/ATP carriers, the *E. histolytica* carrier is not reliant on a membrane potential for function, nor is it inhibited by classic inhibitors such as carboxyatractyloside and bongkrekic acid. Although most eukaryotes have many mitochondrial carriers, *E. histolytica* contains only a single mitochondrial carrier in its genome (Chan et al., 2005; Loftus et al., 2005). There is no evidence to suggest that energy metabolism might be compartmentalized in *E. histolytica* (Müller, 2000).

Although the evolutionary origins of Cpn10 could not be unequivocally established due to insufficient phylogenetic signal, the mitochondrial ancestry of Cpn60 is well documented (Horner & Embley, 2001; Viale & Arakaki, 1994; Clark & Roger, 1995). Since Cpn10 and Cpn60 are interacting partners, one might assume that both have a similar, if not identical, ancestry. However, in order to reliably trace this past, more phylogenetic signal than these 69 aa can provide is needed, as noted by Fast et al. (2002). Major clades are supported by reasonably high bootstrap support, but ‘deeper’ relationships remain without support.

Most mitochondrial matrix proteins contain amino-terminal presequences that are cleaved upon import into the mitochondrion (Pfanner & Truscott, 2002). In contrast, a number of mitochondrial proteins lack such a cleavable presequence, but are nevertheless sorted to the mitochondrial matrix. The targeting information is thought to reside within the amino-terminal part of the protein. No common denominator seems to connect these proteins, since they are involved in a variety of metabolic functions. Known mitochondrial targeted proteins without presequences are bovine rhodanese, rat 3-oxoacyl-CoA thiolase, the β subunit of human electron transfer flavoprotein, and yeast mitochondrial ribosomal protein Ynl8 (Jarvis et al., 1995, and references therein). Interestingly, several eukaryotic Cpn10 homologues have been identified (rat, bovine, human, yeast and potato), and the absence of any amino-terminal targeting information seems to be a common phenomenon (Legname et al., 1995). Our data suggest that the *E. histolytica* Cpn10 is another addition to this list, and indicate that there must be some cryptic signal residing in this protein as well, in order to sort it into the mitosomes. It is becoming apparent that the diversity of mitochondrial remnants in various eukaryotic lineages is reflected in the heterogeneity of their individual protein complements. *E. histolytica* mitosomes, like *T. vaginalis* hydrogenosomes, contain a typical mitochondrial protein refolding system, i.e. Cpn60, Cpn10 and mtHsp70 (this work; Bakatselou et al., 2000; Bui et al., 1996; Tovar et al., 1999). *G. intestinalis* contains genes encoding homologues of chaperonin Cpn60 and mtHsp70 proteins (Arisue et al., 2002; Horner & Embley, 2001; Morrison et al., 2001; Roger et al., 1998; Tovar et al., 2003), but standard screening of the *Giardia* genome database (McArthur et al., 2000; www.mbl.edu/Giardia) provides little evidence for the presence of chaperonin Cpn10. In contrast, the genome of the microsporidian *Encephalitozoon cuniculi* does not appear to have genes encoding Cpn60 or Cpn10, but has retained a mtHsp70 (Katinka et al., 2001). So, while these molecular chaperones have been widely used as reliable tracers of mitochondria, they have not necessarily been retained in all mitochondrial-related organelles. It is possible that, like many other proteins found in the original mitochondrial endosymbiont, molecular chaperones may have been retargeted to a different cell compartment, or lost during the course of evolution.

Information from the first published genome sequence of any amitochondrial eukaryote – that of the microsporidian *Encephalitozoon cuniculi* (Katinka et al., 2001) – suggested that the principal function of the then hypothetical microsporidian mitosome might be iron–sulphur cluster (Isc) assembly, an essential function of mitochondria (Lill & Muhlenhoff, 2005). Homologues of *isc* genes have now been identified in the genomes of many other ‘amitochondrial’ eukaryotes, including *Giardia*, *Cryptosporidium*, *Trichomonas* and *Entamoeba* (Ali et al., 2004; LaGier et al., 2003; Lill & Muhlenhoff, 2005; Tachezy et al., 2001; Tovar et al., 2003; van der Giezen et al., 2004). Whilst the cellular localization of *isc* proteins in *Cryptosporidium* and *Entamoeba* remains to be unequivocally demonstrated, in *Giardia* and *Trichomonas*, these proteins have been localized to mitosomes and hydrogenosomes, respectively (Sutak et al., 2004; Tovar et al., 2003). The apparent ubiquity of *isc* proteins prompted the suggestion that *isc* assembly might have played an important role in the original endosymbiont that gave rise to mitochondria (Embley et al., 2003; Tovar et al., 2003; van der Giezen et al., 2005). Identifying the full metabolic capacity and protein complement of mitosomes may lead to the identification of distinctive features that could be exploited as antiparasitic drug targets – as has been elegantly demonstrated for the endosymbiosis-derived malarial apicoplast (Ralph et al., 2004).

**ACKNOWLEDGEMENTS**

The authors wish to thank Professor Richard Pickersgill for advice, and Neil Sommerville for maintaining parasite cultures. This work was supported by grants from the BBSRC (111/C13820) and the Wellcome Trust (059845).


van der Giezen, M., Cox, S. & Tovar, J. (2004). The iron-sulfur cluster assembly genes *iscS* and *iscU* of *Entamoeba histolytica* were acquired by horizontal gene transfer. *BMC Evol Biol* 4, 7.

