

The Essentials of Protein Import in the Degenerate Mitochondrion of *Entamoeba histolytica*

Pavel Dolezal^{1,2,3*}, Michael J. Dagley¹, Maya Kono⁴, Peter Wolyne², Vladimir A. Likić², Jung Hock Foo¹, Miroslava Sedinová³, Jan Tachezy³, Anna Bachmann⁴, Iris Bruchhaus⁴, Trevor Lithgow¹

1 Department of Biochemistry and Molecular Biology, Monash University, Clayton Campus, Melbourne, Victoria, Australia, **2** Bio21 Institute, University of Melbourne, Parkville, Victoria, Australia, **3** Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic, **4** Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany

Abstract

Several essential biochemical processes are situated in mitochondria. The metabolic transformation of mitochondria in distinct lineages of eukaryotes created proteomes ranging from thousands of proteins to what appear to be a much simpler scenario. In the case of *Entamoeba histolytica*, tiny mitochondria known as mitosomes have undergone extreme reduction. Only recently a single complete metabolic pathway of sulfate activation has been identified in these organelles. The *E. histolytica* mitosomes do not produce ATP needed for the sulfate activation pathway and for three molecular chaperones, Cpn60, Cpn10 and mtHsp70. The already characterized ADP/ATP carrier would thus be essential to provide cytosolic ATP for these processes, but how the equilibrium of inorganic phosphate could be maintained was unknown. Finally, how the mitochondrial proteins are translocated to the mitosomes had remained unclear. We used a hidden Markov model (HMM) based search of the *E. histolytica* genome sequence to discover candidate (i) mitochondrial phosphate carrier complementing the activity of the ADP/ATP carrier and (ii) membrane-located components of the protein import machinery that includes the outer membrane translocation channel Tom40 and membrane assembly protein Sam50. Using *in vitro* and *in vivo* systems we show that *E. histolytica* contains a minimalist set up of the core import components in order to accommodate a handful of mitochondrial proteins. The anaerobic and parasitic lifestyle of *E. histolytica* has produced one of the simplest known mitochondrial compartments of all eukaryotes. Comparisons with mitochondria of another amoeba, *Dictyostelium discoideum*, emphasize just how dramatic the reduction of the protein import apparatus was after the loss of archetypal mitochondrial functions in the mitosomes of *E. histolytica*.

Citation: Dolezal P, Dagley MJ, Kono M, Wolyne P, Likić VA, et al. (2010) The Essentials of Protein Import in the Degenerate Mitochondrion of *Entamoeba histolytica*. PLoS Pathog 6(3): e1000812. doi:10.1371/journal.ppat.1000812

Editor: Dominique Soldati-Favre, University of Geneva, Switzerland

Received: October 15, 2009; **Accepted:** February 8, 2010; **Published:** March 19, 2010

Copyright: © 2010 Dolezal et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: P.D. is the recipient of Marie Curie Outgoing International Fellowship and M.J.D. is the recipient of an Australian Postgraduate Reward. Supported by the German Science Foundation (DFG; BR 1744/7-1) (I.B.), grants from the Australian Research Council (to T.L. and V.A.L.) and the Ministry of Education of the Czech Republic (MSM 0021620858 and LC 07032 to J.T.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: paveldolezal@yahoo.com

‡ Current address: Department of Parasitology, Faculty of Science, Charles University, Prague, Czech Republic

Introduction

Mitosomes and hydrogenosomes are metabolically-specialized forms of mitochondria, found in some of the unicellular pathogens which inhabit oxygen poor environments [1]. A lack of a recognizable mitochondrial compartment had led to the proposal of a group of primitive, primarily amitochondriate, eukaryotes [2]. However, recent evidence has shown the organelles referred to as hydrogenosomes and mitosomes in the ‘amitochondriate eukaryotes’ to be highly evolved mitochondria, having reduced their metabolic pathways as a response to their anaerobic and partly parasitic lifestyles in diverse eukaryotic lineages [1,3–6]. There is no eukaryote known to be primarily amitochondriate, and even secondary loss of mitochondria has not been found.

The biogenesis of mitochondria is the defining aspect of the organelle and depends on the import of proteins from the cytosol, driven by a set of characteristic protein translocases installed in the outer and inner mitochondrial membranes. Mitochondrial precursor proteins are translated on ribosomes in the cytosol,

and then recognized by a protein translocase in the outer mitochondrial membrane (the TOM complex). This TOM complex imports precursor proteins through a channel formed by the essential subunit Tom40. Subsequently, imported proteins are transferred to the sorting and assembly machinery (SAM complex) for assembly into the outer membrane, or one of two translocases in the inner mitochondrial membrane: the TIM22 complex for assembly into the inner membrane, or the TIM23 complex for translocation through the membrane and into the mitochondrial matrix [7,8]. The presence of components of the TOM, TIM and SAM complexes in hydrogenosomes and mitosomes shows these organelles to be mitochondria, despite the impressive metabolic simplification that have taken place in these specialized compartments [9–13].

Mitosomes are the simplest form of mitochondria: they have lost their capacity for ATP synthesis, lost all vestiges of a mitochondrial genome and so far only limited set of proteins have been localized into these tiny double membrane-bound vesicles. This secondary reduction of function has occurred independently in distinct

Author Summary

All eukaryotic organisms have mitochondria, organelles cordoned by a double membrane, which are descendants of an ancestral bacterial endosymbiont. Nowadays, mitochondria are fully integrated into the context of diverse cellular processes and serve in providing energy, iron-containing prosthetic groups and some of the cellular building blocks like lipids and amino acids. In multi-cellular organisms, mitochondria play an additional vital role in cell signaling pathways and programmed cell death. In some unicellular eukaryotes which inhabit oxygen poor environments, intriguing mitochondrial adaptations have taken place resulting in the creation of specialized compartments known as mitosomes and hydrogenosomes. Several important human pathogens like *Entamoeba histolytica*, *Giardia intestinalis*, *Trichomonas vaginalis* and microsporidia contain these organelles and in many cases the function and biogenesis of these organelles remain unknown. In this paper, we investigated the protein import pathways into the mitosomes of *E. histolytica*, which represent one of the simplest mitochondria-related compartment discovered yet. In accordance with the limited organellar proteome, we show that only core components of mitochondria-related protein import machines are present in *E. histolytica* to serve for the import of a small set of substrate proteins.

lineages of eukaryotes, being well characterized in the diplomonad *Giardia intestinalis* [14], the microsporidians (such as *Encephalitozoon cuniculi* [15], *Antonospora locustae* [12], *Trachipleistophora hominis* [16]) and the amoebozoan *Entamoeba histolytica* [17,18]. The majority of known proteins found in the mitosomes of *G. intestinalis* and microsporidia are functional counterparts of mitochondrial proteins found in other organisms, and a unifying feature of all these organelles is their role in the synthesis of iron-sulfur clusters [14,19,20]. So far it is the sole metabolic process known to occur in mitosomes of *G. intestinalis* and microsporidia and conflicting data exist on the presence of the iron-sulfur clusters biosynthesis in *E. histolytica* mitosomes [21,22]. In addition to being widespread in hydrogenosomes and mitosomes, the biogenesis of iron-sulfur centers is the only essential metabolic role of mitochondria in the model organism *Saccharomyces cerevisiae* [23].

Entamoeba histolytica, the causative agent of invasive amoebiasis in humans, seems to have taken further steps towards the extreme reduction of the mitochondrial compartment [24]. It represents the only known eukaryote in which the synthesis of iron-sulfur clusters is mediated by an NIF (nitrogen fixation) system acquired by horizontal gene transfer from an ϵ -proteobacterium [25]. According to prediction algorithms this biosynthetic pathway is predicted to be present in the cytosol instead of the mitochondrial compartment [26]. Consistently, Mi-Ichi et al, did not find either of Nif proteins in their mitosomal proteomic analysis [22] and also no iron-sulfur cluster containing protein is known to be present in the organelles as a candidate substrate for the NIF system [21]. However, Maralikova et al, presented data arguing for the dual localization of both Nif proteins with their specific enrichment in the mitosomes [21].

So far, the mitosomes of *E. histolytica* represent one of the simplest mitochondria known. With the presence of sulfate activation pathway in the mitosome the need for ATP in addition to the molecular chaperones within the organelle is obvious [17,27–29]. Although the ADP/ATP carrier in the mitosomal membrane provides for ATP import [22,27], two questions are left open. Firstly, how do mitosomes recycle inorganic phosphate (P_i)

arising from ATP hydrolysis? Secondly, how are all the mitosomal proteins transported across the membranes of the organelle?

To address these questions, we performed hidden Markov model (HMM) searches of the *E. histolytica* genome and discovered candidate sequences for (i) a P_i carrier to complement the activity of the ATP/ADP carrier, (ii) Tom40, a channel for substrate protein transport across the outer mitosomal membrane, and (iii) Sam50, an assembly machine for Tom40 in the outer membrane. The mitosomal protein import pathway in *E. histolytica* is mitochondrial in nature. Analysis of the mitochondrial protein import machinery of the related amoeba *Dictyostelium discoideum*, suggests that *E. histolytica* has extensively stripped the mitochondrial protein import machines to their essentials. This remarkable degeneration of protein import is in keeping with the apparent paucity of proteins imported into mitosomes in *E. histolytica*.

Results

Substrate transport by mitochondrial carrier proteins in *E. histolytica*

The family of mitochondrial carrier proteins is well characterized in terms of primary sequence motives and crystal structure. According to the structure of the bovine ADP/ATP carrier, six transmembrane segments form an α -helical bundle, consisting of three homologous modules [30]. Each of these modules contains a signature motif in the odd-numbered helices [P]-x-[D/E]-x-x-[K/R]. The proline residue in the motif introduces a kink into three of the six transmembrane helices. In order to obtain a sensitive tool for analysis of *E. histolytica* genome, a hidden Markov model was built that describes the defining features of the mitochondrial carrier protein family.

Only two protein sequences were identified in the HMM search of the *E. histolytica* genome: 269.m00084 (E-value of 2.8×10^{-94}) and 13.m00296 (E-value of 4.7×10^{-5}) (Accession numbers XP_649800 and XP_656350, respectively). The first sequence corresponds to the ADP/ATP carrier [27] and the second to the novel, unannotated protein sequence. Additional analysis of the sequence length and predicted topology support the possibility that 13.m00296 could encode a mitochondrial carrier protein (Figure 1A), one of the most divergent of the carrier protein family. Based on the functional analyses that follow, we labeled the gene as *ehpic* and its translation product as *EhPiC*.

The 31.7 kDa protein is predicted to have six transmembrane helices, although with the lack of signature motives in the odd-number helices. Only the third helix has a proline residue in the conserved position (Figure 1A). As is the case in fungi, animals and plants, *D. discoideum* has a diverse set of mitochondrial carrier proteins that can be compared in a phylogenetic analyses [31,32]. When added to the dataset from *D. discoideum*, the *E. histolytica* ADP/ATP carrier clusters together with transporters of adenine nucleotides and coenzyme A, while *EhPiC* clusters with the P_i carrier McfN (Figure 1B). The overall sequence divergence of *EhPiC* is reflected by a long branch formed in the tree and also by low statistical support. However, a similar tree topology was obtained using different reconstruction methods (Supplementary information Figure S1), which indicated the affinity of *EhPiC* to McfN. The amino acid residues at three contact sites in the central pore are believed to determine the specificity of the transport channel and their composition can therefore serve for the estimation of the substrate nature [32]. Perhaps due to the high degree of sequence divergence, none of the three sites of *EhPiC* provided any leads towards the putative pore specificity (Figure 1A). The proposed function of a phosphate carrier in the mitosomes is summarized diagrammatically in Figure 1C.

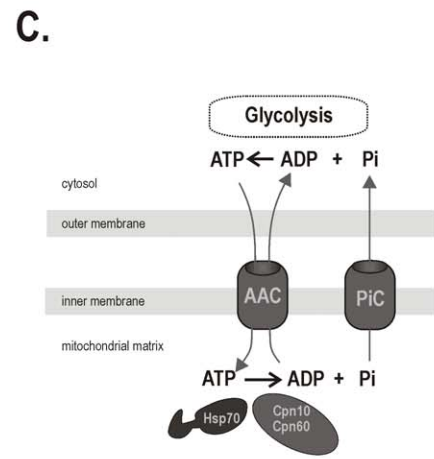
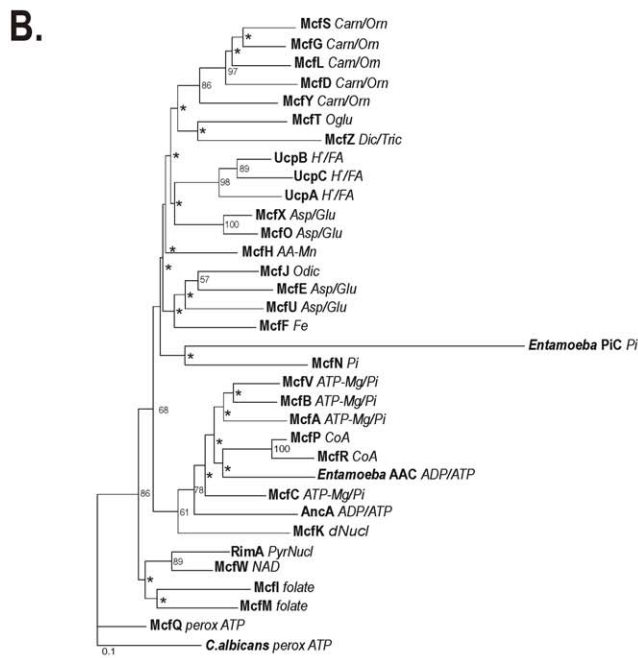
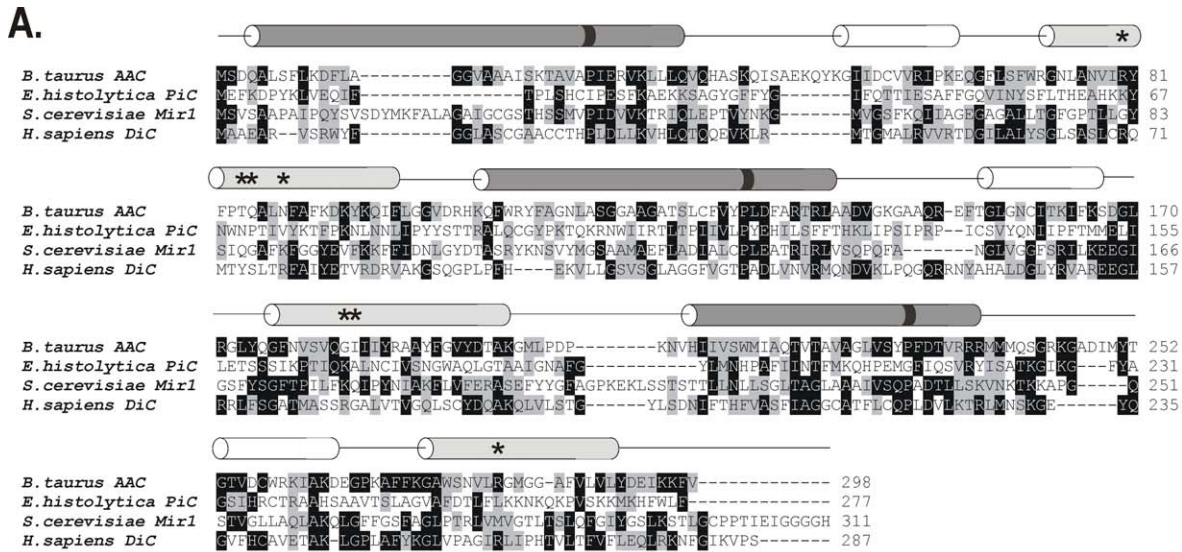


Figure 1. Mitochondrial phosphate carrier in *E. histolytica*. (A) Protein sequence alignment of *EhPiC* with *Bos taurus* ADP/ATP carrier, *S. cerevisiae* phosphate carrier and human dicarboxylate carrier. The secondary structure is schematically depicted by colored cylinders above the alignment according to [30]. The odd- and even-numbered transmembrane α -helices are shown in dark or light grey, respectively. The short helices exposed to the mitochondrial matrix are shown in white. The black stripes on the helices depict the presence/absence of conserved proline residues in the mitochondrial carriers signature motif. Asterisks indicate substrate contact sites. (B) Protein maximum likelihood tree of 34 mitochondrial carrier proteins from *D. discoideum* together with the ADP/ATP carrier and PiC carrier from *E. histolytica* constructed by PhymI [80]. *C. albicans* peroxisomal ADP/ATP carrier was used as the outgroup. The putative substrates of the carriers in *D. discoideum* are inferred from sequence similarity to *S. cerevisiae* and human carrier proteins [31] and indicated as follows: *Carn/Orn* – carnitine or ornithine, *Asp/Glu* – aspartate/glutamate, *Oglu* – 2-oxoglutarate, *Dic/Tric* – dicarboxylate/tricarboxylate, *Pi* – phosphate, *PyrNucl* – pyrimidine NTP/NMP, *perox ATP* – peroxisomal ATP carrier, *H⁺ FA* – H⁺ fatty acid, *CoA* – coenzyme A, *dNucl* – deoxynucleotide, *AA-Mn* – amino acid (Mn²⁺), *Fe* – iron (mitoferrin), asterisks denote nodes with bootstrap value lower than 50. (C) A schematic representation of *EhPiC* and *EhADP/ATP* carrier in the mitosomes, depicting the role of *EhPiC* in recycling phosphate, which is released during ATP hydrolysis by the action of molecular chaperones. *EhADP/ATP* carrier mediates the exchange of cytosolic ATP for mitosomal ADP.

doi:10.1371/journal.ppat.1000812.g001

To determine whether *EhPiC* could function as a phosphate carrier protein, we made use of *S. cerevisiae* as a cellular assay system. Fluorescence microscopy showed that *EhPiC* contains the targeting sequences necessary for the import into mitochondria *in vivo* (Figure 2A). The *EhPiC* protein can be translated *in vitro* and is

imported by mitochondria isolated from *S. cerevisiae* as readily as the *EhAAC* (Figure 2B). Figure 2C shows that *EhPiC* behaves as a typical integral membrane protein, exclusively distributed in the pellet fraction after resisting sodium carbonate extraction of mitochondrial membranes. Soluble and peripheral membrane

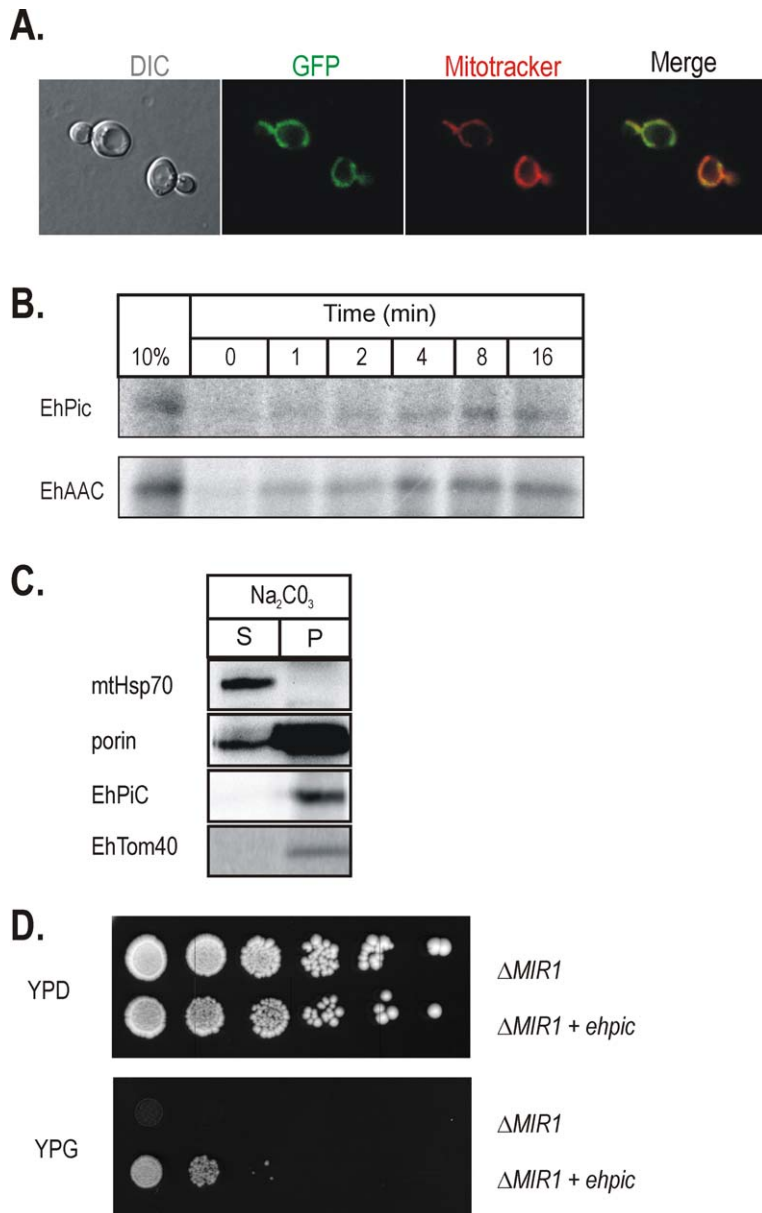


Figure 2. EhPiC is a mitochondrial-type phosphate carrier. (A) The C-terminal GFP fusion of EhPiC expressed in yeast cells (green) co-stained with the mitochondria-specific stain Mitotracker red (red). The merged image demonstrates the co-localization of EhPiC with the mitochondrial compartment. DIC – differential contrast (Nomarski). (B) ³⁵S-labelled EhPiC and EhAAC were incubated with yeast mitochondria, reactions were stopped on ice and the organelles subjected to treatment with proteinase K. 10%–10% of precursor amount used for the import reaction. (C) Yeast mitochondria with ³⁵S-labelled precursors (after 15 minutes of import) were subjected to treatment with proteinase K and then extracted with sodium carbonate. The immunodecoration of mitochondrial Hsp70 and porin shows the distribution of the typical soluble and the integral membrane proteins, respectively. S-soluble fraction, P-pellet. (D) The $\Delta mir1$ deletion strain [33] was transformed by a plasmid carrying *ehpic* ($\Delta mir1+ehpic$) or with a plasmid without *ehpic* ($\Delta mir1$). Serial dilutions of transformed cells plated on fermentable (YPD) and non-fermentable (YPG) carbon source demonstrate the capability of EhPiC to restore the growth defect of the mutant strain. doi:10.1371/journal.ppat.1000812.g002

proteins, like mtHsp70, are extracted from membranes by sodium carbonate.

To test if EhPiC functions to transport phosphate, the coding sequence was cloned into a yeast expression vector and transformed into *S.cerevisiae* strain lacking the dominant phosphate carrier (*Amir1*) [33,34]. The *Amir1* mutants grow on glucose containing media but, due to failure to maintain phosphate transport to sustain oxidative phosphorylation, *Amir1* cells fail to grow on the non-fermentable carbon source glycerol (Figure 2D).

EhPiC expression complements the growth defect of these cells, demonstrating that it functions as a phosphate carrier.

Protein import machinery encoded in the genome of *E. histolytica*

Both *E. histolytica* and the social amoeba *D. discoideum* are amoebozoans, one of the six super-groups of eukaryotes [35,36]. As a sister clade to the animal and fungal lineages, it was not surprising that the core components of the TOM and TIM

machinery, characterized in fungi and animals, have been identified in *D. discoideum* [37–40].

Given this conservation in protein import pathways between amoebozoans, fungi and animals, it was highly surprising when the only homologs of the mitochondrial protein import machinery identified in the genome of *E. histolytica* [41] were the chaperones mtHsp70, Cpn60 and Cpn10.

Hidden Markov models were built, representing search tools for 33 mitochondrial components known to participate during the import process either in fungal/animal or plant mitochondria. Each model was built from sequences clearly homologous to the functionally characterized import component. The models were assembled into a library and used to search the *E. histolytica* genome (see Materials and Methods). In parallel we analyzed the genome of *D. discoideum* (Figure 3), which is the most extensively studied species of Amoebozoa super-group thus providing a referential dataset for our analysis of the *E. histolytica* genome.

Two sequences 38.m00236 and 137.m00093 (accession numbers XP_655014 and XP_651988) were identified by the Tom40- and Sam50-specific HMMs, respectively. Accordingly, these two open reading frames were named *ehtom40* and *ehsam50* and their putative protein products as *EhTom40* and *EhSam50*. The third open reading frame identified in *E. histolytica* corresponds to the previously identified gene encoding for mitochondrial-type Hsp70 [42].

A core component of the TOM complex: *EhTom40*

The gene *ehtom40* encodes a protein of 305 amino acids with a theoretical molecular weight of 34.4 kDa, similar in size to the Tom40 from *Encephalitozoon cuniculi* [43]. *EhTom40* is expressed in *E. histolytica* with the *ehtom40* mRNA detected in extracts from amoebae by reverse transcription coupled with specific PCR amplification (Supplementary information Figure S2). The reverse BLASTP search with *EhTom40* as a query provided hits to a ‘porin family protein’ from *Arabidopsis thaliana* (NP_175457.1), the Tom40 sequence from *Trimastix pyriformis* (ABW76113.1) and a

hypothetical protein from *Leishmania infantum* (XP_001463193.1) representing an unidentified protein predicted to be a β -barrel. Such a structure is typical for Tom40 proteins. The mitochondrial porin 3 superfamily of proteins (Pfam01459) encompasses both the eukaryotic voltage-dependent anion channel (VDAC) proteins and Tom40s [44,45]. Other mitochondrial β -barrel proteins like Sam50 do not fall within the porin 3 superfamily. Given the similarity of *EhTom40* to β -barrel proteins predicted to be Tom40 (ABW76113.1) and VDAC (NP_175457.1), we performed a CLANS analysis using a sequence set containing 137 VDAC and 79 Tom40 sequences (including *EhTom40*). Membrane proteins are often difficult to resolve on phylogenetic trees but CLANS, that graphically depicts homology in large datasets of proteins, has been previously found to be very efficient in the classification of the β -barrel proteins from bacteria [46] and also in the characterization of a Tom40 homologue in *G. intestinalis* [47]. This approach utilizes all-against-all pairwise BLAST, clustering ‘vertices’ (i.e. individual protein sequences) in three-dimensional space using an algorithm which applies a weak repulsive force to each vertex and an attractive force between each pair of vertices. The attractive and repulsive forces are proportional to the BLAST high scoring segment pair score. Figure 4A shows that cluster analysis performed on this superset of porin3 family proteins using a P-value cutoff of 10^{-82} results in the sequences clustering into two distinct groups. The clusters are either exclusively VDAC homologues or Tom40 homologues, corresponding to their independent phylogenies across eukaryotes. The *EhTom40* is positioned on the periphery of Tom40 cluster, clearly separated from the compact cluster of VDAC sequences.

Mitochondrial β -barrel proteins contain a specific mitochondrial targeting signal (the β -signal, PxGxxHxH, where P stands for polar, G for glycine and H for hydrophobic residue) that sits in the last β -strand of the protein [48]. The β -signal is recognized and bound by the SAM complex, which then mediates folding and membrane insertion of the β -barrel. *EhTom40* has a β -signal that strictly follows this rule (Figure 4B). When *EhTom40* is expressed

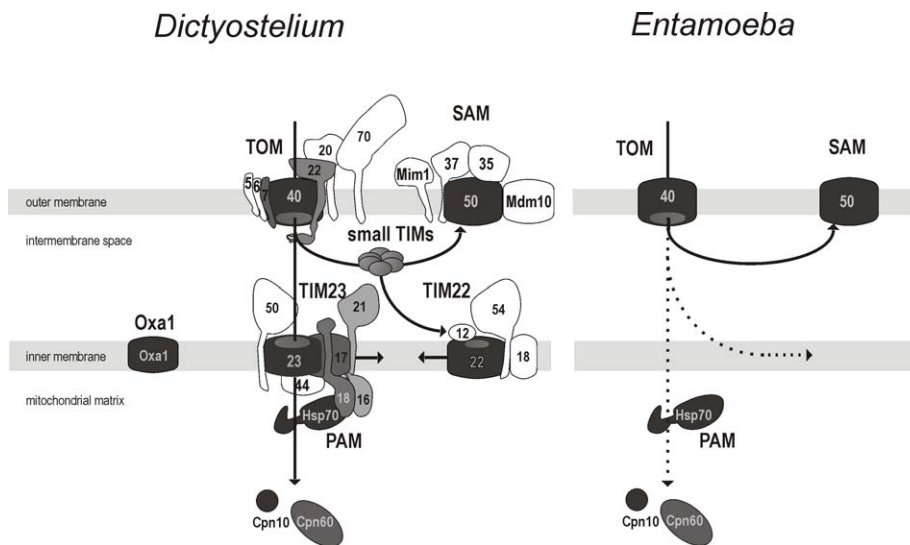
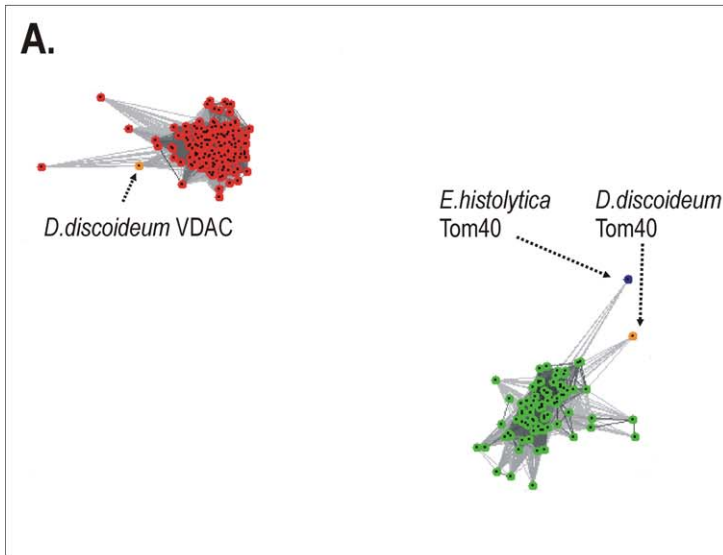


Figure 3. The minimalism of the protein import and assembly pathways in the mitosomes of *E. histolytica*. The schematic representation depicts the complex molecular apparatus in the amoebozoan mitochondria of *D. discoideum* and the mitosome of *E. histolytica* based on hidden Markov model analysis. Of the components present in animal and fungal mitochondria, those shown shaded were found in *D. discoideum*. The panel on the right shows the five components identified in *E. histolytica*. Some of the proteins found in the animals and fungi appear to be missing in Amoebozoa (shown in white). The dashed arrows highlight the unknown translocation pathway across or into the inner mitosomal membrane in *E. histolytica*.

doi:10.1371/journal.ppat.1000812.g003

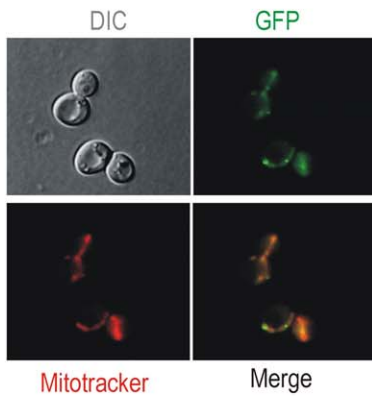


B.

<i>E. histolytica</i>	GSYSWGASIQIF	304
<i>S. cerevisiae</i>	NDTKIGCGLQFE	360
<i>H. sapiens</i>	NKFQCGFGLTIG	361
<i>D. discoideum</i>	QDYKFGLGLTFQ	313
<i>A. thaliana</i>	KDYKFGFGLTVG	309

xxxPxGxxHxHx

C.



D.

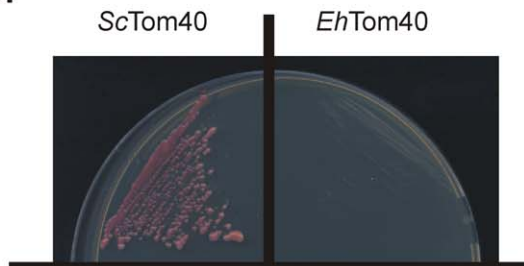


Figure 4. A mitochondrial Tom40 in *Entamoeba histolytica*. (A) A representative figure from the CLANS analysis of a data set containing 79 Tom40 sequences (green dots), 137 VDAC (red dots) sequences and the *EhTom40* sequence (blue dot). A P-value threshold of 10^{-82} was used to give complete separation of the VDAC and Tom40 clusters. Tom40 and VDAC sequences from *D. discoideum*, as the closest relative in the dataset, are highlighted (orange). (B) The alignment shows the C-terminal β -strand from various Tom40 homologues. The motif for the general import signal for mitochondrial β -barrel proteins is highlighted, P-polar residue, G-glycine residue, H-hydrophobic residue [48]. (C) The C-terminal GFP fusion of *EhTom40* in yeast cells (green), the yeast mitochondria stained with Mitotracker red (red). The merged image demonstrates the co-localization of *EhTom40* with the mitochondrial compartment. DIC – differential contrast (Nomarski). (D) *S. cerevisiae* $\Delta tom40$ mutants, carrying a counterselectable *TOM40* gene and expressing either *S. cerevisiae* Tom40 (ScTom40) or *EhTom40* were plated onto 5-FOA media to select against the covering plasmid. Cells were incubated for 4 days at 30°C and only strain expressing ScTom40 was viable.
doi:10.1371/journal.ppat.1000812.g004

in *S. cerevisiae*, it is specifically targeted to mitochondria, as judged by fluorescence microscopy (Figure 4C). In order to test whether *EhTom40* can functionally substitute for *S. cerevisiae* Tom40 we have transformed a haploid $\Delta tom40$ strain that carries the *S. cerevisiae* TOM40 on a counterselectable plasmid. Upon plating on 5-FOA media, to select against the covering plasmid, no viable transformants were obtained (Figure 4D), while the strain transformed with the homologous *TOM40* remained viable. Apparently the sequence divergence between the fungal and amoebic Tom40 interferes with the correct docking and interaction with other TOM complex subunits and thus does not allow for functional complementation. Therefore we subsequently tested whether *EhTom40* integrates into the native *S. cerevisiae* TOM complex using the *in vitro* import assays. Although the protein accumulated as a high molecular weight species in the mitochondrial membranes (Supplementary information Figure S3), the specific pull-down assay using the antibodies against TOM complex subunits did not recover any of *EhTom40* protein (data not shown).

Given that TOM complex requires numerous binding sites on several distinct subunits to assemble properly, it is not surprising that these experiments failed. It was previously reported that even the dual point mutations in Tom40 of *S. cerevisiae* can result in the disassembly of the native TOM complex [49].

A core component of the SAM complex: *EhSam50*

The cDNA sequence of Sam50 homologue found in *E. histolytica* is deposited at NCBI under accession number XM_646896. It encodes for a protein of 388 amino acids with a theoretical molecular weight of 45.3 kDa. However, when aligned with the corresponding genomic sequence, the presence of an intervening sequence was revealed. The insertion of 72 bp starting at position 802 is limited by GAATGATT and TAG at the 5'- and 3'-ends. These are conserved intron donor and acceptor splice sites identified in *E. histolytica* [50], suggesting that *ehsam50* gene is interrupted by a single intron. According to the molecular weight of the protein when episomally expressed in *E. histolytica* and *S. cerevisiae* the mRNA is likely processed *in vivo* in both cellular systems and provides for *EhSam50* translation.

As shown on phylogenetic reconstruction the protein sequence clusters with other mitochondrial Sam50 sequences [51,52] with α -proteobacterial Omp85 sequences in a sister clade (Figure 5B) [53], implying that *EhSam50* is of direct mitochondrial origin.

Omp85 proteins contain a C-terminal β -barrel domain (the 'surface antigen domain'), which is integrated into the outer membrane. A characteristic N-terminal extension consists of five POTRA (polypeptide-transport-associated) domains [54]. In contrast, mitochondrial Sam50 proteins, including *EhSam50*, contain a single POTRA domain along a β -barrel domain (Figure 5A). Together with obtained phylogenetic data this provides additional support that *EhSam50* is a genuine mitochondrial homologue, not a bacterial Omp85 acquired by a horizontal gene transfer. While the structural domains are well conserved in *EhSam50*, we could not obtain a viable *S. cerevisiae* $\Delta sam50$ strain

expressing the *E. histolytica* protein (Figure 5C). As in the case of Tom40, the primary sequence was likely too divergent to support functional integration of *EhSam50* into SAM complex of *S. cerevisiae*. So far, there is no record in the literature of successful heterologous replacement of SAM complex components, which suggests that tight protein-protein or protein-lipid interactions are involved in the SAM function.

Localization of proteins in *E. histolytica*

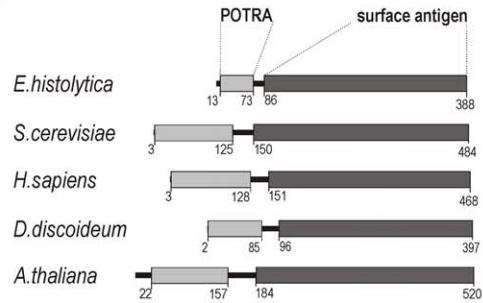
In order to test *EhSam50 in vivo*, the expression of *EhSam50* (and *EhTom40* and *EhPiC*) in *E. histolytica* was confirmed by RT-PCR (Supplementary information Figure S2). As an experimental system, *E. histolytica* is challenging. In order to detect whether *EhSam50*, *EhTom40* and *EhPiC* are localized to a subcellular compartment, we set about both raising antisera to the three proteins and creating HA-tagged versions of the proteins for expression in *E. histolytica*. Neither strategy allowed localization of *EhPiC* or *EhTom40* in *E. histolytica* (recent data of Mi-Ichi et al. support the mitochondrial distribution of *EhTom40* [22]). However, an α -*EhSam50* specific serum was raised and an *E. histolytica* strain expressing *EhSam50* with a C-terminal HA-tag was cultivated. Fixed *E. histolytica* cells were probed with α -HA and with either α -*EhSam50*, α -mtHsp70 or α -Atg8 antibody and immunofluorescent labeling determined. Atg8 is a marker of autophagocytic vesicles, which might be of similar morphology to mitosomes [55]. *EhSam50* was found in vesicles with a size and overall pattern similar to the cpn60-specific labeling by León-Avila and Tovar [56]. The co-distribution of HA- and *EhSam50*- specific signal on the merged image shows that the antibodies specifically label the same cellular compartment (Figure 6A), distinct from the Atg8-specific labeling of autophagosomes [55]. More importantly, using a α -mtHsp70 raised against mitochondrial Hsp70 from *Neocallimastix patriciarum* [57] the specific co-localization with *EhSam50* was found. This data strongly supports the mitochondrial distribution of *EhSam50* together with mitochondrial Hsp70 [42].

To further test the mitochondrial distribution of the protein import machinery in *E. histolytica*, the cell fractions were probed for the mitochondrial and the cytosolic marker proteins (Figure 6B). While our data suggested the cytosolic distribution of NifU and NifS proteins as found in the analysis of Mi-Ichi et al. [22], *EhSam50* and mitochondrial Cpn60 were distributed in the high-speed pellet fraction containing mitosomes.

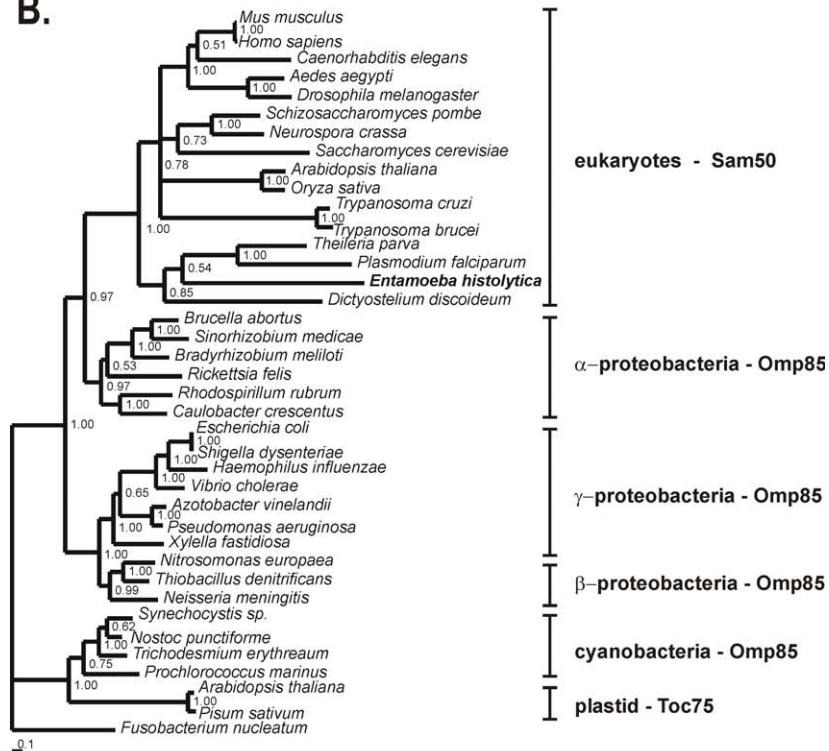
Discussion

Hidden Markov models can be used for a highly sensitive search of sequence databases [58], and discovered three new mitochondrial proteins in *E. histolytica*. Tom40 provides a channel for the import of mitochondrial proteins including Sam50 and Tom40 itself, whereas Sam50 is necessary for assembly of the Tom40 channel (and Sam50 itself). The characterization of homologs of Sam50 and Tom40 in *E. histolytica* suggests that mitosomes have a functional outer membrane as also recently shown by Maralikova et al. [21]. It is now clear, even with so few mitochondrial proteins identified in *E. histolytica*, that

A.



B.



C.

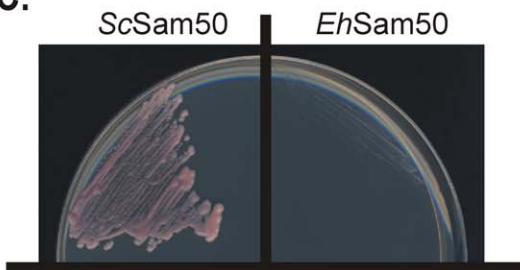


Figure 5. Sam50 in *Entamoeba histolytica*. (A) Signature domains present in Sam50: A C-terminal surface antigen domain (dark grey) which corresponds to the membrane embedded β -barrel, and the POTRA domain (light grey) which is exposed to the intermembrane space. Both domains are present in *EhSam50* as determined by Pfam search [45] and HHpred analyses [81]. (B) Protein maximum-likelihood phylogenetic tree was derived from a dataset of 407 aligned amino acids from 39 sequences using MrBayes [70]. Numbers at the individual nodes represent MrBayes posterior probabilities. (C) *S. cerevisiae* Δ *sam50* mutants, carrying a counterselectable *SAM50* gene and expressing either *S. cerevisiae* Sam50 (ScSam50) or *EhSam50* were plated onto 5-FOA media to select against the covering plasmid. Cells were incubated for 4 days at 30°C and only strain expressing ScSam50 was viable.
doi:10.1371/journal.ppat.1000812.g005

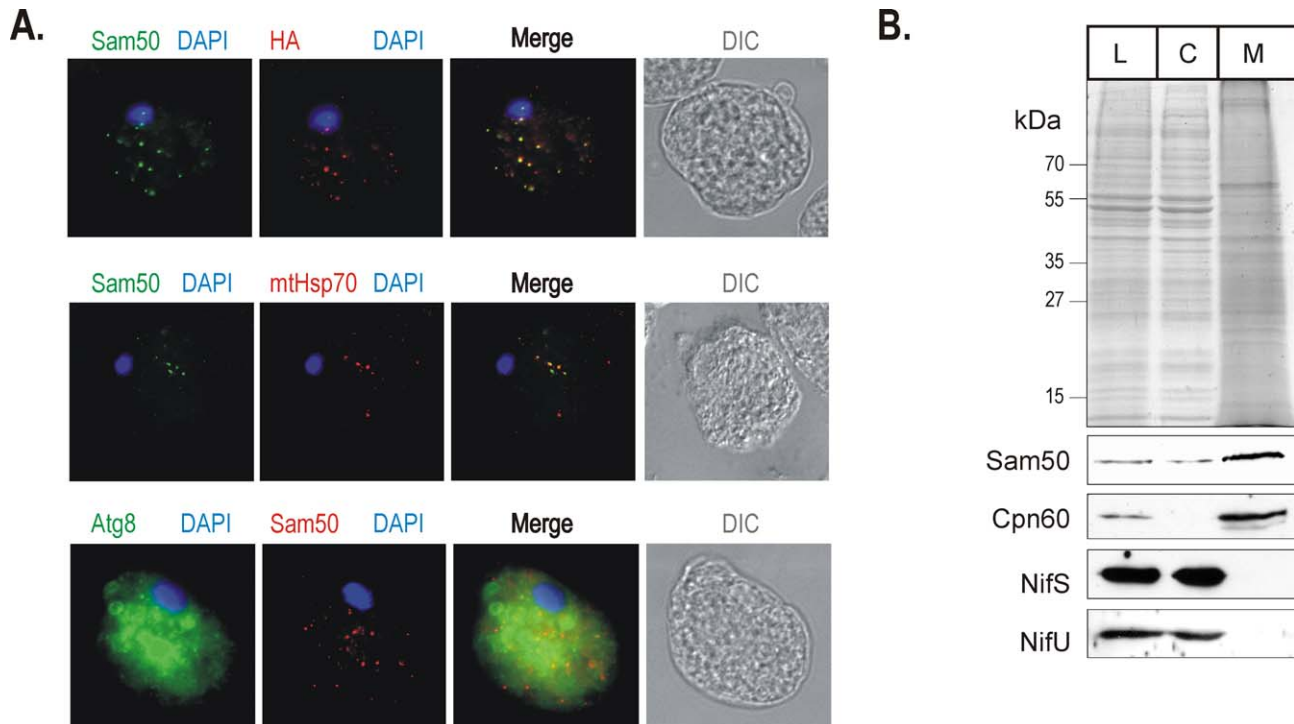


Figure 6. Sam50 is confined to the mitosomal vesicles uniformly distributed through the cytoplasm. (A) The upper panel shows fixed *E. histolytica* cells expressing the HA-tagged version of *EhSam50* were probed with α -HA (green) and α -*rEhSam50* (red) antibody for immunofluorescent labeling. DIC – differential contrast (Nomarski). In the middle panel HA-tagged *EhSam50* was labeled by α -HA (green) and mtHsp70 with polyclonal α -mtHsp70 (red) raised against *Nyctotherus ovalis* mtHsp70 [57]. The lower panel shows fixed *E. histolytica* cells expressing the HA-tagged version of *EhSam50* were probed with α -Atg8, an autophagosome marker (green) and α -*EhSam50* (red) antibody. (B) Cultured *E. histolytica* cells were disrupted and the cell lysate (L), cytosol (C) and mitosome-enriched fraction (M) and probed for NifS, NifU using homologous polyclonal antibodies [41] and Cpn60 using heterologous polyclonal antibody raised against *Giardia intestinalis* Cpn60. *EhSam50* was decorated with α -HA monoclonal antibody.
doi:10.1371/journal.ppat.1000812.g006

we have in hand marker proteins for the outer membrane (Sam50, Tom40), the inner membrane (ADP/ATP carrier, PiC, sodium/sulfate symporter) and the mitosomal matrix (e.g. molecular chaperones, rubrerythrin, ATP sulfurylase).

The inner membrane of all mitochondria known so far is populated by the proteins from Tim17/22/23 family. These proteins contain four membrane spanning helices, by which they build up the core channels of the translocase of the inner membrane (TIM). In animals fungi, plants and diverse protists groups these proteins are present in three different forms as Tim17, Tim22 and Tim23, which play distinct roles in the import of matrix and/or inner membrane proteins. Some eukaryotes, notably trypanosomes and microsporidians, appear to make use of just a single Tim17/22/23 protein [13,59], which might functionally substitute for the three distinct forms as a “jack of all trades”. Despite the high degree of similarity among the Tims, from these various eukaryotes including *D. discoideum*, the HMM models of the Tim17/22/23 family of proteins did not reveal any homologous sequences in *E. histolytica*. This could either mean that the putative *E. histolytica* Tims have diverged in sequence to such an extent as to be beyond the sensitivity of our HMM models, or that *E. histolytica* lacks any Tim17/22/23-related protein due to the overall mitochondrial reduction. Failure to detect a Tim17 family protein in mitosomes of *G. intestinalis* argues for such a common reductive step in mitochondrial adaptation [10,60]. An interesting question is raised about the mechanism of carrier protein assembly in the mitosomal membranes of *E. histolytica* as these proteins normally require functional TIM22 translocase.

Owing to its extreme lifestyle, *E. histolytica* is believed to depend solely on cytosolic production of ATP through glycolysis [61]. The ADP/ATP carrier in *E. histolytica* mitosomes provides for retrograde transport of ATP into the mitosomal matrix, enabling the ATP-dependent activity of Hsp70 and Cpn60 as well as ATP sulfurylase and APS kinase [17,22,28,29]. Our finding of the phosphate carrier *EhPiC* completes this metabolic shunt.

Targeting of the proteins into *E. histolytica* mitosomes

Most of soluble mitochondrial proteins contain cleavable N-terminal targeting sequence that usually forms an amphipathic helix, which is recognized by the receptor on the outer mitochondrial membrane and upon the translocation into mitochondrial matrix cleaved off by a matrix processing peptidase (MPP) [62]. The robust bioinformatic search for more proteins with putative mitosomal targeting sequences in *E. histolytica* is hindered by very poor training dataset available to date (Supplementary information Figure S4) [22]. However, in our additional HMM search for a putative MPP homologue we identified single candidate protein from M16 peptidase family and we are currently characterizing this protein. Although MPP is a heterodimer of two different but evolutionary related subunits, a single subunit minimalist version of the enzyme was reported recently in the mitosomes of *G. intestinalis* [60]. Given that mitosomes of *E. histolytica* represent even more reduced mitochondria one can anticipate analogous reduction occurring herein.

Protein import into amoebic mitochondria - a window into mitochondrial evolution

Phylogenetic reconstructions and the presence of flagellate cells in the life-cycles of some Amoebozoans suggested an affiliation of this supergroup with animals and fungi [36], so that studies on fundamental pathways in the Amoebozoa provide important additional information on the cell biology of the earliest animals and fungi.

Genome sequence analysis of the highly-studied amoebozoan, *D. discoideum*, revealed the presence of 14 different components taking part in the mitochondrial protein import in animals and fungi (see Figure 3). This analysis suggests that all four major membrane complexes: SAM, TOM, TIM22 and TIM23 are present in *D. dictyostelium*, as is the intermembrane space-located small TIM chaperones that shuttle substrates to the SAM or TIM22 complexes. The presence of these various complexes in *D. discoideum* is consistent with them having been installed in the earliest stages of eukaryote evolution. The components that are missing in *D. discoideum* are those that appear to be specific for fungi and animals, as has been previously discussed for receptor subunits Tom20 [63] and Tom70 [64]. That so many TOM and TIM components are missing from *E. histolytica*, is strong evidence for a secondary gene loss having occurred, as part of the reductive evolution impacting on the mitochondrial organelle. The anaerobic lifestyle, combined with parasitism, of *E. histolytica* likely selected for a minimalist mitochondrial set up and the enormous reduction of the import machinery. This likely reflects the even more dramatic reduction of the overall mitochondrial metabolism. The anaerobic, free-living amoeba *Mastigamoeba balamuthi* looks to have traveled part way the same direction: 21 putative mitochondrial proteins have been identified in the limited EST dataset, while no TOM, TIM and SAM components were found so far [65].

The recent data on the function of *E. histolytica* mitosome have revealed remarkable divergence of the processes occurring within this mysterious mitochondrion. It now seems that while these organelles have lost vast majority of typical mitochondrial functions, they have accommodated several new unexpected roles not seen in other mitochondria [21,22]. In this paper, we uncovered the essentials of the protein import into these organelles and it is very likely that the further research on the function and the biogenesis of *E. histolytica* mitosome will bring more of the unexpected.

Materials and Methods

Sequence search and analysis

The Hidden Markov models describing 33 mitochondrial components were constructed and compiled into a HMM library from manually prepared families of sequences (Oxa1, Tim10, Tim44, Tom20opistho, Tom20plants, Tom70, Pam16, Tim13, Tim50, Tom22opistho, Tom22plants, Hsp70, Pam18, Tim17, Tim54, Sam35, Tim18, Tim8, Tom40, Mdm10, Sam37, Tim21, Tim9, Tom5, Metaxin1, Sam50, Tim22, Tim9+10, Tom6, Metaxin2, Tim23, and Tom7). The HMM library was used to scan the genomes of *E. histolytica* (<http://www.tigr.org/tdb/e2k1/cha1/>) and *D. discoideum* (NCBI). In addition, a hidden Markov model based on 34 manually compiled MCP sequences was built and used to scan the two genomes. The program HMMER 2.3.2 was used in all calculations [66], and the search results were extracted with the programs prepared in-house.

The homology modeling of the mitochondrial carrier protein was performed with SwissModel at <http://swissmodel.expasy.org//SWISS-MODEL.html> [67]. The structure of bovine ATP/ADP carrier (PDB ID 2C3E) was used as a template [30]. The

sequences were aligned using ClustalX [68] and edited manually in BioEdit (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>).

For the neighbor joining analysis of mitochondrial carrier proteins amino acid sequences were aligned and the resulting alignment edited manually into a dataset of 33 sequences of 190 amino acid residues. SplitsTree4 [69] software was used to calculate the bootstrapping of 500 runs and to combine the results into a Neighbor-Net.

Protein maximum-likelihood phylogenetic tree of Omp85/Toc75/Sam50 proteins was derived from a dataset of 407 aligned amino acids from 39 sequences. The tree was obtained using the program MrBayes under the JTT substitution matrix with amino acid frequencies estimated from the dataset [70]. Site rate variation was modeled under a discrete approximation to the Γ distribution (one invariable and four variable rate categories). Four Monte Carlo Markov Chains, each with 2,000,000 generations, were performed with trees sampled every 100 generations. For compilation of Bayesian consensus topologies, a “burn-in” of 500 trees was used.

CLANS version 2 October 9, 2006 was obtained from <http://bioinfoserwer.rsbs.anu.edu.au/programs/clans/>. The *Eh*Tom40 sequence was added to a sequence set of 137 VDAC and 79 Tom40 sequences; the sequence set was derived from both the results of BLAST and HMM searches using VDAC and Tom40 sequences. Cluster analysis was performed with a P-value cutoff (10^{-82}) sufficient to observe complete separation of the VDAC and Tom40 sequence clusters. As the method is non-deterministic, the analysis was run until stable clusters formed (in excess of 200 iterations). Multiple runs were performed to ensure that the observed clusters formed using different starting positions for the sequences.

Entamoeba histolytica culture and preparation of RNA

Trophozoites of the *E. histolytica* isolate HM-1:IMSS were cultured axenically in TYI-S-33 medium in plastic tissue culture flasks [71]. For further experiments 1×10^6 trophozoites were cultivated for 24 h in 75 ml culture flasks. The trophozoites were then harvested after being chilled on ice for 5 min and sedimented at $430 \times g$ for 5 min. The resulting pellet was washed once in phosphate-buffered saline pH 7.2 and once in 20 mM MOPS pH 7.2. The cell pellet was resuspended in 2 ml of 20 mM MOPS pH 7.2 with protease inhibitors and passed 6 times through a 25 G needle until most cell were broken. The cell lysate was diluted with 25 ml of 250 mM Sucrose, 20 mM MOPS pH 7.2 and spun down twice at $650 \times g$ for 10 min, resulting supernatant was spun down at $2,850 \times g$ for 10 min. The final high-speed pellet representing the mitochondrial fraction was obtained after centrifugation at $100,000 \times g$ for 30 min. The high-speed supernatant corresponded to the cytosolic fraction.

For total RNA isolation 1×10^6 *E. histolytica* trophozoites were cultivated in 75 ml culture flasks for 24 h. The cells were harvested via chilling on ice for 5 min and sedimented at $200 \times g$ for 5 min at 4°C . The cell pellet was washed twice with PBS. The trophozoites were treated with TRIZOL reagent (Invitrogen) following the manufacturer's instructions. Extracted RNA was purified using the RNeasy mini kit (Qiagen) without β -mercaptoethanol and DNA was digested with DNase (Qiagen). cDNA synthesis was accomplished with SuperScriptIII Reverse Transcriptase (Invitrogen). In a final volume of 20 μl , 1 μg of RNase-free and DNase-treated total RNA was mixed with $5 \times$ First-Strand buffer, 500 μM dNTPs, 500 nM OdT-T71 (5'-GAG AGA GGA TCC AAG TAC TAA TAC GAC TCA CTA TAG GGA GAT24), 2 mM DTT, 40 U RnaseOut (Invitrogen) and SuperScriptIII (200 U/ μl). cDNA was synthesised for 1 h at 42°C .

Yeast culture and cell fractionation

Saccharomyces cerevisiae strain W303a was grown in rich medium or selective medium as previously described [72]. The $\Delta mir1$ strain was a gift from Dr. Geneviève Dujardin (Centre de Génétique Moléculaire, CNRS, Paris, France) [33]. For the preparation of mitochondria for the *in vitro* study *S. cerevisiae* strain W303a was grown in lactate media at 25°C. The mitochondria were isolated by differential centrifugation as described previously [73]. For the growth assays the cells were grown to a mid-logarithmic phase in a complete media, diluted into OD₆₀₀ = 0.2, spotted in the series of fivefold dilutions on the plates and incubated at 30°C for 3–6 days.

Cloning and expression of *ehpic*, *ehtom40* and *ehsam50*

For GFP-tagging, the open reading frames were amplified by PCR using *E. histolytica* genomic DNA as template and primers containing 5' *EcoRI* and 3' *BamHI* or *BglII* sites (see Table S1) and cloned into p416MET25 vector [74]. To create the C-terminal HA-tag fusions the ORFs were cloned into a pYX143 vector with the use of 5' *EcoRI* and 3' *MluI* restriction sites. For the expression of *ehsam50* in *E. histolytica*, the ORF was amplified from pYX143 with the C-terminal HA-tag using the 5' *KpnI* and 3' *BglII* restriction sites. The plasmid (pNC) used for transfection is a derivative of the expression vector EhNEO/CAT. The plasmid contains the neomycin phosphotransferase-coding sequence flanked by 480 bp of the 5'-untranslated sequence and 600 bp of the 3'-untranslated sequence of an *E. histolytica* actin gene. Transfections were performed by electroporation as described previously [75]. Drug selection started 48 h after transfection, using 10 µg/ml of the neomycin analogue G418. Two weeks later, the G418 concentration was increased to 50 µg/ml.

For the production of EhSam50-derived antigen the first 300 bp of *ehsam50* were amplified by PCR and ligated into pET23a vector (Novagen) using 5' *NdeI* and 3' *XhoI* restriction sites. *E. coli* strain BL21 (DE3) was used to produce the recombinant protein with C-terminal six histidine tag. The protein was purified to homogeneity under denaturing condition (8 M urea) on NTA-nickel column (Qjagen). For generation of polyclonal antibodies 100 µg recombinant EhSam50 (rEhSam50) was injected into a mouse, followed by two further injections.

In vitro protein import

Mitochondria were prepared according to the method of Daum et al [73]. For *in vitro* transcription the genes were amplified by PCR with the forward primers containing the SP6 promoter followed by a Kozak's sequence (*ehpic*, *ehtom40*, *ehsam50*) or cloned into pSP73 vector (*ehAAC*), which was linearized at a unique site downstream of the gene. [76,77]. *In vitro* imports were assayed according to Gabriel and Pfanner [78].

Immunofluorescence microscopy and western blot analysis

E. histolytica cells were analyzed with α -EhSam50 antiserum, with an antiserum raised against the *E. histolytica* autophagosome marker Atg8 (autophagy related gene 8, a gift from Dr. Tomoyoshi Nozaki, National Institute of Infectious Diseases, Tokyo, Japan) or with a mouse monoclonal anti-HA antibody. Cells were fixed at room temperature for 30 min in PBS containing 3% paraformaldehyde and subsequently permeabilized with 0.05% saponin (PBSS). Samples were incubated at room temperature for 1 h with antisera against EhSam50 (1:250 dilution), against Atg8 (1:500 dilution) or against the HA-tag (1:200, Roche). Secondary antibodies were Alexa-594 coupled α -mouse, Alexa-488 coupled α -rabbit, and Alexa-594 coupled α -

mouse antibodies. Subsequently, cells were mounted on glass slides and examined under 6300× magnification. For deconvolution microscopy images of selected cells were captured with a 63× oil immersion lens in a UV equipped Leica DM RB microscope with 0.2-µm-diameter step Z-sections. Deconvoluted Z sections were examined for colocalisation of Atg8 and EhSam50 staining with the Openlab 4.0.4 program. Adobe Photoshop 7.0.5 was used for additional processing of the images. *S. cerevisiae* were analyzed as previously described (Beilharz et al. 2003). ImageJ software was used for additional image processing (<http://rsbweb.nih.gov/ij/>). In the western blot analysis cell fractions were probed with the mouse monoclonal anti-HA antibody (1:500), rabbit polyclonal antibodies raised against *Giardia intestinalis* Cpn60 (1:1000) (a kind gift from Dr. Robert Hirt, Newcastle University, UK), *E. histolytica* NifS and NifU (both 1:1000) (a kind gift from Dr. Tomoyoshi Nozaki [26]).

The distribution of EhPic, EhTom40 and EhSam50 in *S. cerevisiae* mitochondria extracted in fresh 100mM Na₂CO₃ was done as previously described [79].

Supporting Information

Figure S1 Neighbor joining tree of 31 mitochondrial carrier proteins from *D. discoideum* with ADP/ATP carrier and PiC carrier from *E. histolytica* constructed by SplitsTree4 [86]. The carriers cluster according to their substrate specificity. The putative substrates of *D. discoideum* carriers [87] are indicated as follows: *Carn/Orn* - carnitine or ornithine, *Asp/Glu* - aspartate/glutamate, *Oglu* - 2-oxoglutarate, *Dic/Tric* - dicarboxylate/tricarboxylate, *Pi* - phosphate, *Pyr/Nucl* - pyrimidine NTP/NMP, *perox ATP* - peroxisomal ATP carrier, *H⁺ FA* - H⁺ fatty acid, *CoA* - coenzyme A, *dNucl* - deoxynucleotide, *AA-Mn* - amino acid (Mn²⁺), *Fe* - iron (mitoferrin).

Found at: doi:10.1371/journal.ppat.1000812.s001 (0.12 MB PDF)

Figure S2 The products of specific RT are shown: 1 - *ehpic*, 2 - *ehtom40*, 3 - *ehsam50*. The marker is in base pairs.

Found at: doi:10.1371/journal.ppat.1000812.s002 (0.05 MB PDF)

Figure S3 ³⁵S-labelled *EhTom40* was incubated with yeast mitochondria, solubilized with 1% digitonin and the samples resolved on BN-PAGE. The time-dependent formation of high-molecular weight complexes in the mitochondrial membranes is demonstrated. The large white arrow points to disappearance of the monomeric form of *EhTom40* in the reaction, small black arrows highlight the formations of the of high-molecular weight complexes.

Found at: doi:10.1371/journal.ppat.1000812.s003 (0.05 MB PDF)

Figure S4 The protein sequence alignments of all three known soluble mitochondrial matrix proteins Cpn10, Cpn60, Hsp70 with their bacterial and eukaryotic counterparts. The presence of extremely short N-terminal targeting sequence does not allow for creation of a prediction algorithm.

Found at: doi:10.1371/journal.ppat.1000812.s004 (0.07 MB PDF)

Table S1 Primers used in the study.

Found at: doi:10.1371/journal.ppat.1000812.s005 (0.03 MB PDF)

Acknowledgments

We thank Graham Clark (London School of Hygiene and Tropical Medicine, UK) for *E. histolytica* genomic DNA, Geneviève Dujardin (Centre de Génétique Moléculaire, CNRS, Paris, France) for a $\Delta mir1$ *S. cerevisiae* strain, Tomoyoshi Nozaki (National Institute of Infectious Diseases, Tokyo, Japan) for α -Atg8, α -NifS and α -NifU antibodies, and Robert Hirt (Newcastle University, UK) for α -Cpn60 antibody.

Author Contributions

Conceived and designed the experiments: PD MJD MK IB TL. Performed the experiments: PD MJD MK VAL JHF MS AB IB. Analyzed the data:

PD MJD MK PW VAL TL. Contributed reagents/materials/analysis tools: PD JT. Wrote the paper: PD TL.

References

- Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. *Nature* 440: 623–630.
- Cavalier-Smith T (1989) Molecular phylogeny. *Archaea* and *Archezoa*. *Nature* 339: 100–101.
- Akhmanova A, Voncken F, van Alen T, van HA, Boxma B, et al. (1998) A hydrogenosome with a genome. *Nature* 396: 527–528.
- Hrdý I, Hirt RP, Dolezal P, Bardonová L, Foster PG, et al. (2004) *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. *Nature* 432: 618–622.
- Dyall SD, Yan W, Delgadillo-Correa MG, Lunceford A, Loo JA, et al. (2004) Non-mitochondrial complex I proteins in a hydrogenosomal oxidoreductase complex. *Nature* 431: 1103–1107.
- van derGiezen (2009) Hydrogenosomes and mitosomes: conservation and evolution of functions. *J Eukaryot Microbiol* 56: 221–231.
- Mokranjac D, Neupert W (2005) Protein import into mitochondria. *Biochem Soc Trans* 33: 1019–1023.
- Kutik S, Guiard B, Meyer HE, Wiedemann N, Pfanner N (2007) Cooperation of translocate complexes in mitochondrial protein import. *J Cell Biol* 179: 585–591.
- Dolezal P, Likić V, Tachezy J, Lithgow T (2006) Evolution of the molecular machines for protein import into mitochondria. *Science* 313: 314–318.
- Dolezal P, Smid O, Rada P, Zubacová Z, Bursac D, et al. (2005) *Giardia* mitosomes and trichomonad hydrogenosomes share a common mode of protein targeting. *Proc Natl Acad Sci U S A* 102: 10924–10929.
- Regoes A, Zourmanou D, León-Avila G, van derGiezen, Tovar J, et al. (2005) Protein import, replication, and inheritance of a vestigial mitochondrion. *J Biol Chem* 280: 30557–30563.
- Burri L, Williams BA, Bursac D, Lithgow T, Keeling PJ (2006) Microsporidian mitosomes retain elements of the general mitochondrial targeting system. *Proc Natl Acad Sci U S A* 103: 15916–15920.
- Burri L, Keeling PJ (2007) Protein targeting in parasites with cryptic mitochondria. *Int J Parasitol* 37: 265–272.
- Tovar J, León-Avila G, Sanchez LB, Sutak R, Tachezy J, et al. (2003) Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature* 426: 172–176.
- Tsaousis AD, Kunji ER, Goldberg AV, Lucocq JM, Hirt RP, et al. (2008) A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*. *Nature* 453: 553–556.
- Williams BA, Hirt RP, Lucocq JM, Embley TM (2002) A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. *Nature* 418: 865–869.
- Tovar J, Fischer A, Clark CG (1999) The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol Microbiol* 32: 1013–1021.
- Mai Z, Ghosh S, Frisardi M, Rosenthal B, Rogers R, et al. (1999) Hsp60 is targeted to a cryptic mitochondrion-derived organelle (“crypton”) in the microaerophilic protozoan parasite *Entamoeba histolytica*. *Mol Cell Biol* 19: 2198–2205.
- Goldberg AV, Molik S, Tsaousis AD, Neumann K, Kuhnke G, et al. (2008) Localization and functionality of microsporidian iron-sulphur cluster assembly proteins. *Nature* 452: 624–628.
- Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, et al. (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* 414: 450–453.
- Maralikova B, Ali V, Nakada-Tsukui K, Nozaki T, van derGiezen, et al. (2009) Bacterial-type oxygen detoxification and iron-sulfur cluster assembly in amoebal relict mitochondria. *Cell Microbiol*.
- Mi-Ichi F, Yousuf MA, Nakada-Tsukui K, Nozaki T (2009) Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. *Proc Natl Acad Sci U S A*.
- Lill R, Mühlenhoff U (2006) Iron-sulfur protein biogenesis in eukaryotes: components and mechanisms. *Annu Rev Cell Dev Biol* 22: 457–486.
- Clark CG (2000) The evolution of *Entamoeba*, a cautionary tale. *Res Microbiol* 151: 599–603.
- 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.
- Ali V, Shigeta Y, Tokumoto U, Takahashi Y, Nozaki T (2004) An intestinal parasitic protist, *Entamoeba histolytica*, possesses a non-redundant nitrogen fixation-like system for iron-sulfur cluster assembly under anaerobic conditions. *J Biol Chem* 279: 16863–16874.
- Chan KW, Slotboom DJ, Cox S, Embley TM, Fabre O, et al. (2005) A novel ADP/ATP transporter in the mitosome of the microaerophilic human parasite *Entamoeba histolytica*. *Curr Biol* 15: 737–742.
- Tovar J, Cox SS, van der Giezen M (2007) A mitosome purification protocol based on percoll density gradients and its use in validating the mitochondrial nature of *Entamoeba histolytica* mitochondrial Hsp70. *Methods Mol Biol* 390: 167–177.
- van der Giezen M, León-Avila G, Tovar J (2005) Characterization of chaperonin 10 (Cpn10) from the intestinal human pathogen *Entamoeba histolytica*. *Microbiology* 151: 3107–3115.
- Pebay-Peyroula E, Dahout-Gonzalez C, Kahn R, Trézéguet V, Lauquin GJ, et al. (2003) Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractylsulfide. *Nature* 426: 39–44.
- Satre M, Mattei S, Aubry L, Gaudet P, Pelosi L, et al. (2007) Mitochondrial carrier family: repertoire and peculiarities of the cellular slime mould *Dictyostelium discoideum*. *Biochimie* 89: 1058–1069.
- Kunji ER, Robinson AJ (2006) The conserved substrate binding site of mitochondrial carriers. *Biochim Biophys Acta* 1757: 1237–1248.
- Hamel P, Saint-Georges Y, de Pinto B, Lachacinski N, Altamura N, et al. (2004) Redundancy in the function of mitochondrial phosphate transport in *Saccharomyces cerevisiae* and *Arabidopsis thaliana*. *Mol Microbiol* 51: 307–317.
- Murakami H, Blobel G, Pain D (1990) Isolation and characterization of the gene for a yeast mitochondrial import receptor. *Nature* 347: 488–491.
- Simpson AG, Roger AJ (2002) Eukaryotic evolution: getting to the root of the problem. *Curr Biol* 12: R691–R693.
- Nikolaev SI, Berner C, Petrov NB, Mylnikov AP, Fahrni JF, et al. (2006) Phylogenetic position of *Multicilia marina* and the evolution of Amoebozoa. *Int J Syst Evol Microbiol* 56: 1449–1458.
- Eichinger L, Pachebat JA, Glockner G, Rajandream MA, Sugang R, et al. (2005) The genome of the social amoeba *Dictyostelium discoideum*. *Nature* 435: 43–57.
- Macasev D, Whelan J, Newbigin E, Silva-Filho MC, Mulhern TD, et al. (2004) Tom22', an 8-kDa trans-site receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes. *Mol Biol Evol* 21: 1557–1564.
- Barth C, Le P, Fisher PR (2007) Mitochondrial biology and disease in *Dictyostelium*. *Int Rev Cytol* 263: 207–252.
- Nagayama K, Itono S, Yoshida T, Ishiguro S, Ochiai H, et al. (2008) Antisense RNA inhibition of the beta subunit of the *Dictyostelium discoideum* mitochondrial processing peptidase induces the expression of mitochondrial proteins. *Biosci Biotechnol Biochem* 72: 1836–1846.
- Loftus B, Anderson I, Davies R, Alsmark UC, Samuelson J, et al. (2005) The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433: 865–868.
- Bakatselou C, Kidgell C, Graham CC (2000) A mitochondrial-type hsp70 gene of *Entamoeba histolytica*. *Mol Biochem Parasitol* 110: 177–182.
- Waller RF, Jabbour C, Chan NC, Celik N, Likić VA, et al. (2009) Evidence of a reduced and modified mitochondrial protein import apparatus in microsporidian mitosomes. *Eukaryot Cell* 8: 19–26.
- Pusnik M, Charriere F, Maser P, Waller RF, Dagley MJ, et al. (2009) The single mitochondrial porin of *Trypanosoma brucei* is the main metabolite transporter in the outer mitochondrial membrane. *Mol Biol Evol* 26: 671–680.
- Finn RD, Tate J, Misty J, Coghill PC, Sammut SJ, et al. (2008) The Pfam protein families database. *Nucleic Acids Res* 36: D281–D288.
- Frickey T, Lupas A (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* 20: 3702–3704.
- Dagley MJ, Dolezal P, Likić VA, Smid O, Purcell AW, et al. (2009) The protein import channel in the outer mitochondrial membrane of *Giardia intestinalis*. *Mol Biol Evol* 26: 1941–1947.
- Kutik S, Stojanovski D, Becker L, Becker T, Meinecke M, et al. (2008) Dissecting membrane insertion of mitochondrial beta-barrel proteins. *Cell* 132: 1011–1024.
- Gabriel K, Egan B, Lithgow T (2003) Tom40, the import channel of the mitochondrial outer membrane, plays an active role in sorting imported proteins. *EMBO J* 22: 2380–2386.
- Wilihoef U, Campos-Gongora E, Touzni S, Bruchhaus I, Tannich E (2001) Introns of *Entamoeba histolytica* and *Entamoeba dispar*. *Protist* 152: 149–156.
- Gentle IE, Burri L, Lithgow T (2005) Molecular architecture and function of the Omp85 family of proteins. *Mol Microbiol* 58: 1216–1225.
- Voulhoux R, Tommassen J (2004) Omp85, an evolutionarily conserved bacterial protein involved in outer-membrane-protein assembly. *Res Microbiol* 155: 129–135.
- Gentle I, Gabriel K, Beech P, Waller R, Lithgow T (2004) The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. *J Cell Biol* 164: 19–24.
- Sánchez-Pulido L, Devos D, Genevros S, Vicente M, Valencia A (2003) POTRA: a conserved domain in the FtsQ family and a class of beta-barrel outer membrane proteins. *Trends Biochem Sci* 28: 523–526.
- Picazarri K, Nakada-Tsukui K, Nozaki T (2008) Autophagy during proliferation and encystation in the protozoan parasite *Entamoeba invadens*. *Infect Immun* 76: 278–288.
- León-Avila G, Tovar J (2004) Mitosomes of *Entamoeba histolytica* are abundant mitochondrion-related remnant organelles that lack a detectable organellar genome. *Microbiology* 150: 1245–1250.

57. van derGiezen, Birdsey GM, Horner DS, Lucocq J, Dyal PL, et al. (2003) Fungal hydrogenosomes contain mitochondrial heat-shock proteins. *Mol Biol Evol* 20: 1051–1061.
58. Eddy SR (1996) Hidden Markov models. *Curr Opin Struct Biol* 6: 361–365.
59. Schneider A, Bursac D, Lithgow T (2008) The direct route: a simplified pathway for protein import into the mitochondrion of trypanosomes. *Trends Cell Biol* 18: 12–18.
60. Smid O, Matuskova A, Harris SR, Kucera T, Novotny M, et al. (2008) Reductive evolution of the mitochondrial processing peptidases of the unicellular parasites *Trichomonas vaginalis* and *Giardia intestinalis*. *PLoS Pathog* 4: e1000243. doi:10.1371/journal.ppat.1000243.
61. Saavedra E, Encalada R, Pineda E, Jasso-Chavez R, Moreno-Sánchez R (2005) Glycolysis in *Entamoeba histolytica*. Biochemical characterization of recombinant glycolytic enzymes and flux control analysis. *FEBS J* 272: 1767–1783.
62. Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N (2009) Importing mitochondrial proteins: machineries and mechanisms. *Cell* 138: 628–644.
63. Likić VA, Perry A, Hulett J, Derby M, Traven A, et al. (2005) Patterns that define the four domains conserved in known and novel isoforms of the protein import receptor Tom20. *J Mol Biol* 347: 81–93.
64. Chan NC, Likić VA, Waller RF, Mulhern TD, Lithgow T (2006) The C-terminal TPR domain of Tom70 defines a family of mitochondrial protein import receptors found only in animals and fungi. *J Mol Biol* 358: 1010–1022.
65. Gill EE, az-Trivino S, Barbera MJ, Silberman JD, Stechmann A, et al. (2007) Novel mitochondrion-related organelles in the anaerobic amoeba *Mastigamoeba balamuthi*. *Mol Microbiol* 66: 1306–1320.
66. Eddy SR (1998) Profile hidden Markov models. *Bioinformatics* 14: 755–763.
67. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22: 195–201.
68. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25: 4876–4882.
69. Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23: 254–267.
70. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
71. Diamond LS, Harlow DR, Cunnick CC (1978) A new medium for the axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. *Trans R Soc Trop Med Hyg* 72: 431–432.
72. Lithgow T, Junne T, Wachter C, Schatz G (1994) Yeast mitochondria lacking the two import receptors Mas20p and Mas70p can efficiently and specifically import precursor proteins. *J Biol Chem* 269: 15325–15330.
73. Daum G, Bohni PC, Schatz G (1982) Import of proteins into mitochondria. Cytochrome b2 and cytochrome c peroxidase are located in the intermembrane space of yeast mitochondria. *J Biol Chem* 257: 13028–13033.
74. Beilharz T, Egan B, Silver PA, Hofmann K, Lithgow T (2003) Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in *Saccharomyces cerevisiae*. *J Biol Chem* 278: 8219–8223.
75. Hamann L, Nickel R, Tannich E (1995) Transfection and continuous expression of heterologous genes in the protozoan parasite *Entamoeba histolytica*. *Proc Natl Acad Sci U S A* 92: 8975–8979.
76. Nijtmans LG, Henderson NS, Holt IJ (2002) Blue Native electrophoresis to study mitochondrial and other protein complexes. *Methods* 26: 327–334.
77. Wittig I, Braun HP, Schagger H (2006) Blue native PAGE. *Nat Protoc* 1: 418–428.
78. Gabriel K, Pfanner N (2007) The mitochondrial machinery for import of precursor proteins. *Methods Mol Biol* 390: 99–117.
79. Youker RT, Walsh P, Beilharz T, Lithgow T, Brodsky JL (2004) Distinct roles for the Hsp40 and Hsp90 molecular chaperones during cystic fibrosis transmembrane conductance regulator degradation in yeast. *Mol Biol Cell* 15: 4787–4797.
80. Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52: 696–704.
81. Soding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33: W244–W248.