Evidence of a Reduced and Modified Mitochondrial Protein Import Apparatus in Microsporidian Mitosomes

Ross F. Waller,1* Carole Jabbour,1,2 Nickie C. Chan,2 Nermin Çelik,2 Vladimir A. Likić,2 Terrence D. Mulhern,2 and Trevor Lithgow2†

School of Botany, University of Melbourne, Victoria 3010, Australia;1 and Department of Biochemistry and Molecular Biology, Bio21 Institute of Biotechnology and Molecular Science, University of Melbourne, Victoria 3010, Australia2

Received 15 September 2008/Accepted 8 November 2008

Microsporidia are a group of highly adapted obligate intracellular parasites that are now recognized as close relatives of fungi. Their adaptation to parasitism has resulted in broad and severe reduction at (i) a genomic level by extensive gene loss, gene compaction, and gene shortening; (ii) a biochemical level with the loss of much basic metabolism; and (iii) a cellular level, resulting in lost or cryptic organelles. Consistent with this trend, the mitochondrion is severely reduced, lacking ATP synthesis and other typical functions and apparently containing only a fraction of the proteins of canonical mitochondria. We have investigated the mitochondrial protein import apparatus of this reduced organelle in the microsporidian Encephalitozoon cuniculi and find evidence of reduced and modified machinery. Notably, a putative outer membrane receptor, Tom70, is reduced in length but maintains a conserved structure chiefly consisting of tetratricopeptide repeats. When expressed in Saccharomyces cerevisiae, EcTom70 inserts with the correct topology into the outer membrane of mitochondria but is unable to complement the growth defects of Tom70-deficient yeast. We have scanned genomic data using hidden Markov models for other homologues of import machinery proteins and find evidence of reduced and modified machinery. Notably, a putative outer membrane receptor, Tom70, is reduced in length but maintains a conserved structure chiefly consisting of tetratricopeptide repeats. When expressed in Saccharomyces cerevisiae, EcTom70 inserts with the correct topology into the outer membrane of mitochondria but is unable to complement the growth defects of Tom70-deficient yeast. We have scanned genomic data using hidden Markov models for other homologues of import machinery proteins and find evidence of reduced and modified machinery.

Microsporidia are a group of highly adapted obligate intracellular parasites (31, 50). They infect a diverse range of vertebrate and invertebrate animal hosts. In humans they are the cause of a number of diseases (e.g., gastroenteritis, encephalitis, and hepatitis), having their greatest impact on immunocompromised individuals, notably in children with human immunodeficiency virus (14, 31). Microsporidia are most closely related to fungi, although their high level of specialization as intracellular parasites obscured this relationship for a long time (18, 25, 30). Gene phylogenies now firmly connect these two groups, although it remains uncertain whether microsporidia are sisters to the fungi or represent a lineage derived from within fungal diversity (21, 28).

A clear adaptive response to parasitism in microsporidia has been a reduction in cellular complexity. This was first recognized at an ultrastructural level with the apparent lack of peroxisomes, flagella, stacked Golgi bodies, and mitochondria (31). This reductive evolution is mirrored at a genomic level, with microsporidia containing the smallest eukaryotic genomes described to date (28, 29). The complete genomic sequence from the human microsporidian parasite Encephalitozoon cuniculi reveals a genome of only ~2.9 Mb containing approximately 2,000 genes, in contrast to the 6,000 genes found in the genome of the model fungus Saccharomyces cerevisiae. The minimal genome of E. cuniculi has been achieved through three mechanisms in concert: (i) gene loss, resulting in broad loss of biochemical pathways and capabilities, including much basic energy metabolism and numerous anabolic pathways; (ii) gene compaction with an average intergenic space of ~130 bp; and (iii) gene shortening, with E. cuniculi genes being on average 14% shorter than their homologues in fungi such as S. cerevisiae (28, 45). Thus, microsporidian evolution has apparently been shaped by a very strong trend to eliminate superfluous molecular and biochemical complexity.

Despite earlier suppositions that microsporidia lacked mitochondria, genome and expressed sequence tag data from microsporidia suggested the presence of several proteins typically targeted to this organelle (3, 19, 20, 24, 28, 38). Immunolocalization of a mitochondrial Hsp70 to small double membrane-bound organelles in Trachipleistophora hominis provided strong evidence for the existence of a mitochondrion in microsporidia, albeit a simplified organelle that lacks cisternae (48). Annotation of genomic data from E. cuniculi provided compelling matches for only 22 proteins implicated in mitochondrial function, suggesting that the metabolism of this relict mitochondrion (or mitosome) is also significantly reduced compared to that of canonical mitochondria (28). Further, no mitochondrial genome has been retained; thus, biogenesis of this organelle is wholly dependent on nucleus-encoded proteins. Based on these 22 proteins, a major role for the mitosome is iron-sulfur cluster assembly (22, 28). No genes have been found for ATP synthesis via oxidative phosphorylation, suggesting loss of this activity in mitosomes (28, 46). While it is likely that further mitosome-targeted proteins will be identified, it is clear that compared to mitochondria from fungal relatives, which are known to import ~1,000 proteins (40, 44), microsporidian mitosomes represent organelles with highly reduced proteomes, a feature consistent with other traits of cellular reduction.

The highly reduced state of the microsporidian mitosome,
requiring only a fraction of the protein diversity of other mitochondrial, presents an interesting case for studying organelle biogenesis—particularly the machinery for protein import of nucleus-encoded proteins. Mitochondrial protein import has been best characterized in fungi, and in these systems most proteins are imported via four major import complexes: a TOM (translocase of the outer mitochondrial membrane), a SAM (sorting and assembly machinery), and one of two TIMs (translocase of the inner mitochondrial membrane), TIM23 or TIM22 (see Fig. 5A) (5, 36). These complexes are broadly conserved throughout fungi as well as animals (15). Mitochondrial proteins can take one of several routes to the mitochondrion via this apparatus (5, 36). Broadly, soluble matrix proteins are recognized at the TOM complex by the receptor protein Tom20 through the binding of N-terminal presequences with characteristic features (1, 5, 7, 8, 36). These proteins are passed through the pore protein Tom40 of the TOM to the TIM23 complex and then driven into the mitochondrial matrix by way of the presequence translocase-associated motor (PAM) complex, where their presequences are subsequently removed. Some membrane proteins can also be released into the inner membrane from the TIM23 complex. Mitochondrial proteins that apparently lack such an extension, notably including many of the membrane proteins, are recognized by internal sequence elements. Tom70 has a greater role in recognizing these internal signals and thus the import of hydrophobic proteins (4, 11, 32, 39, 47). Such hydrophobic proteins are often bound by cytosolic molecular chaperones (Hsp70 and/or Hsp90) en route to the mitochondrion, and Tom70 is known to independently bind to both the chaperone and the substrate protein (7, 23, 33, 52). While a measure of substrate overlap between Tom20 and Tom70 occurs, the division of responsibility between these two receptors has likely evolved in response to the wide range of substrate proteins that must be imported into mitochondria and the need to handle this complexity.

For microsporidia little is known of the protein import apparatus for their relict mitochondrion, the mitosome. Has the very reduced organelle proteome, in concert with a genome-wide trend of the loss of redundant or superfluous genes, resulted in a smaller and/or derived import apparatus? In this study we have investigated the microsporidian mitosome protein import apparatus from *E. cuniculi* in order to evaluate how this apparatus has responded to the reduction in the number of proteins required to be imported and the overall radical reduction in the number and size of proteins encoded in the nuclear genome. A putative homologue of the outer membrane receptor protein Tom70 is of particular interest as the only receptor for the TOM complex and, given the known structure of Tom70 proteins, provides a highly informative example of how proteins can be shortened in the course of genome reduction.

**MATERIALS AND METHODS**

Analyses of HMMs. Hidden Markov models (HMMs) were generated for the following mitochondrial import protein-related families: AlphaMPP, BetaMPP, Hsp70, MDP, Mdm10, Metax1, Metax2, Osxl, Patm16, Patm18, Sam15, Sam31, Sam40, Secf14amma, SecC, SecE, SecG, SecY, Skp, ScarA, TaxC, Tim1, Tim13, Tim17, Tim18, Tim21, Tim22, Tim23, Tim44, Tim50, Tim54, Tim58, Tim8, Tim9, Tim9-10, Tom20, Tom22, Tom40, Tom5, Tom6, Tom70, and Tom7. Models were based on known homologues in as broad a taxon sample as possible (protein training sets are available upon request). The building of HMMs and the search of the *E. cuniculi* genomic data were performed with the software package HMMER (16, 17), version 3.2.3. *E. cuniculi* genomic data were downloaded from the Genoscope web site with the EMBL file (http://www.genoscope.cns.fr/extern/sequences/banque_Projet_AD). Coding sequences were extracted with the EMBOSS program “coderet” (41). The HMM search was performed with the program “hmmpfam.” The initial candidate hits were extracted as matches with E values of <0.01 with in-house tools. Each candidate sequence was manually reviewed and searched against the NCBI Protein database (BLASTp) to eliminate proteins with strong matches to nontarget proteins.

Expression vector construction. *E. cuniculi* genomic DNA, kindly provided by E. S. Didier (Tulane University, Louisiana), was harvested from purified spores grown in tissue culture using the QiAmp mini-DNA extraction kit (Qiagen, California). By using PCR the complete EcTom70 open reading frame was amplified with in-frame restriction sites with the forward primer 5’-GCT CTT CTG ATC TAA CAC CGC-3’ and the reverse primer 5’-GAC TGT CGA CGT AAA TAC GTC TCT GTA AAT GAA CAT CAG G-3’. PCR fragments were ligated into a modified version of the yeast expression vector p416 MET25 HDEL (37), and sequences were confirmed by DNA sequencing. Plasmids were transformed into haploid yeast strain MH272 and the Δtom70Δtom71 double deletion strain (NCY 0404) (11) and plated onto uracil-deficient selective medium.

In vivo localization. Yeast cells expressing EcTom70-GFP fusion proteins were generated with Mitotagger or Red CM-H2XRos (Molecular Probes). Yeast growth assays were performed with *S. cerevisiae* (S288c-2a) using the yeast medium BY4741 (supD ura3Δ 0 leu2Δ 0 hist1Δ 0 lys2Δ 0 trp1Δ 0) and plasmid p416 MET25 HDEL (37) harboring a 4.3-kb PstI fragment with the *E. cuniculi* Δtom70Δtom71 genomic DNA insert amplified with in-frame restriction sites with the forward primer 5’-ACT GGG ATC CAA ACG TCT TAA GAA CAT CAG G-3’. Yeast cells were grown to mid-logarithmic phase (optical density at 600 nm of 0.6) in selective minimal medium (SD-Ura) and diluted to an optical density at 600 nm of 0.2, and then 5-µl aliquots were serially diluted fivefold and spotted onto YPAD (yeast extract 1%, peptone 2%, adenine 0.1%, glucose 2%) and YEPG (yeast extract 1%, peptone 2%, ethanol 3%, glycerol 3%) plates. Plates were incubated at 25°C, 30°C, or 37°C for 2 to 4 days until colonies were visible and then photographed.

RESULTS

Structural conservation and divergence of *E. cuniculi* Tom70. Annotation of *E. cuniculi* genomic data identified an open reading frame with sequence similarity to Tom70 (28). This putative protein, EcTom70, is 477 residues long, 25% shorter than *S. cerevisiae* Tom70 (617 residues), and shares only 12% sequence identity with this fungal homologue. Given the size difference and low identity, EcTom70 is not reliably recovered by BLAST searches using fungal or animal Tom70s. To scrutinize the identity of this putative microsporidian Tom70, we

20 WALLER ET AL. EUKARYOT. CELL

Analyses of HMMs. Hidden Markov models (HMMs) were generated for the following mitochondrial import protein-related families: AlphaMPP, BetaMPP, Hsp70, MDP, Mdm10, Metax1, Metax2, Osxl, Patm16, Patm18, Sam15, Sam31, Sam40, Secf14amma, SecC, SecE, SecG, SecY, Skp, ScarA, TaxC, Tim1, Tim13, Tim17, Tim18, Tim21, Tim22, Tim23, Tim44, Tim50, Tim54, Tim58, Tim8, Tim9, Tim9-10, Tom20, Tom22, Tom40, Tom5, Tom6, Tom70, and Tom7. Models were based on known homologues in as broad a taxon sample as possible (protein training sets are available upon request). The building of HMMs and the search of the *E. cuniculi* genomic data were performed with the software package HMMER (16, 17), version 3.2.3. *E. cuniculi* genomic data were downloaded from the Genoscope web site with the EMBL file (http://www.genoscope.cns.fr/extern/sequences/banque_Projet_AD). Coding sequences were extracted with the EMBOSS program “coderet” (41). The HMM search was performed with the program “hmmpfam.” The initial candidate hits were extracted as matches with E values of <0.01 with in-house tools. Each candidate sequence was manually reviewed and searched against the NCBI Protein database (BLASTp) to eliminate proteins with strong matches to nontarget proteins.

Expression vector construction. *E. cuniculi* genomic DNA, kindly provided by E. S. Didier (Tulane University, Louisiana), was harvested from purified spores grown in tissue culture using the QiAmp mini-DNA extraction kit (Qiagen, California). By using PCR the complete EcTom70 open reading frame was amplified with in-frame restriction sites with the forward primer 5’-GCT CTT CTG ATC TAA CAC CGC-3’ and the reverse primer 5’-GAC TGT CGA CGT AAA TAC GTC TCT GTA AAT GAA CAT CAG G-3’. PCR fragments were ligated into a modified version of the yeast expression vector p416 MET25 HDEL (37), and sequences were confirmed by DNA sequencing. Plasmids were transformed into haploid yeast strain MH272 and the Δtom70Δtom71 double deletion strain (NCY 0404) (11) and plated onto uracil-deficient selective medium.

In vivo localization. Yeast cells expressing EcTom70-GFP fusion proteins were generated with Mitotagger or Red CM-H2XRos (Molecular Probes). Yeast growth assays were performed with *S. cerevisiae* (S288c-2a) using the yeast medium BY4741 (supD ura3Δ 0 leu2Δ 0 hist1Δ 0 lys2Δ 0 trp1Δ 0) and plasmid p416 MET25 HDEL (37) harboring a 4.3-kb PstI fragment with the *E. cuniculi* Δtom70Δtom71 genomic DNA insert amplified with in-frame restriction sites with the forward primer 5’-ACT GGG ATC CAA ACG TCT TAA GAA CAT CAG G-3’. Yeast cells were grown to mid-logarithmic phase (optical density at 600 nm of 0.6) in selective minimal medium (SD-Ura) and diluted to an optical density at 600 nm of 0.2, and then 5-µl aliquots were serially diluted fivefold and spotted onto YPAD (yeast extract 1%, peptone 2%, adenine 0.1%, glucose 2%) and YEPG (yeast extract 1%, peptone 2%, ethanol 3%, glycerol 3%) plates. Plates were incubated at 25°C, 30°C, or 37°C for 2 to 4 days until colonies were visible and then photographed.

RESULTS

Structural conservation and divergence of *E. cuniculi* Tom70. Annotation of *E. cuniculi* genomic data identified an open reading frame with sequence similarity to Tom70 (28). This putative protein, EcTom70, is 477 residues long, 25% shorter than *S. cerevisiae* Tom70 (617 residues), and shares only 12% sequence identity with this fungal homologue. Given the size difference and low identity, EcTom70 is not reliably recovered by BLAST searches using fungal or animal Tom70s. To scrutinize the identity of this putative microsporidian Tom70, we
used HMMs as search tools to see if this *E. cuniculi* sequence could be recovered. Using 35 diverse Tom70s from fungi and animals, an HMM was developed and used to screen the *E. cuniculi* genomic data. EcTom70 was recovered as the best match to the Tom70 model (E value, 4.00e^{-5}), supporting its identity as a Tom70 homologue.

The Tom70 from *S. cerevisiae* (ScTom70) contains an N-terminal transmembrane domain that anchors this protein in the outer mitochondrial membrane, with the remainder of the protein exposed to the cytosol (11). The crystal structure of this cytosolic portion of ScTom70 has been determined and consists of 26 α-helices (51). Twenty-two of these helices contribute to 11 tetratricopeptide repeat (TPR) motifs; this motif is a 34-residue motif consisting of two antiparallel α-helices. The first three of these TPRs and a seventh helix form an N-terminal “clamp” domain implicated in chaperone binding, and the remaining 19 helices create a C-terminal region (referred to as the “core” domain) that selectively binds mitochondrial preproteins (9, 11). A 27-residue region linking the clamp and core domains appears to provide a flexible interdomain loop and was unresolved in the crystal structure; however, the two domains contact via helices A7 and A25-26 arranged in antiparallel orientation (51).

Comparison of EcTom70 to ScTom70 in multiple sequence alignments, including homologues from diverse fungi and animals, allowed assessment of equivalent structures in the microsporidian protein. EcTom70 is predicted to contain a single transmembrane domain at the N terminus (using the TMPred algorithm [26]) (Fig. 1A). Prediction of repeated motifs (by REP algorithm [http://www.embl-heidelberg.de/]) (2) identifies seven likely TPR motifs that correspond in position to TPR1, -2, -5, -9, -10, and -11, with a similar helix-turn-helix structure predicted (with weak similarity to an ankyrin repeat) at the position equivalent to TPR7. Manual inspection of the alignment and prediction of α-helices by JPRED (12) indicates that paired helices corresponding to TPR3 and TPR5 might form TPR structures equivalent to these regions of EcTom70. Further, helices A7, A8, A25, and A26 in ScTom70, which contribute to clamp-core interactions, are also represented by helical segments in the EcTom70 protein sequence (Fig. 1A). Thus, several of the key structural features of ScTom70 are predicted for the EcTom70 sequence. Notable differences are the apparent loss of TRP6 by deletion, the presence of only one helical region, and the absence of two linker regions, (i) the region linking the transmembrane domain with the clamp domain and (ii) that linking the clamp and core domains. Together these changes contribute to the overall shortening of EcTom70 without substantial change to the predicted structure of the protein.

The role of the clamp domain in fungal and animal Tom70 proteins is to bind to molecular chaperones Hsp70 and/or Hsp90, which aid in delivering hydrophobic mitochondrial proteins to the import complex (52). This chaperone-receptor interaction is mediated by binding of the C terminus of the chaperones, characterized by the sequence EEVD, within a groove formed by the three TPRs of the clamp (6, 42, 52). Electrostatic interactions of the terminal aspartate contribute to a dicarboxylate anchor within the clamp. In PSI-BLAST searches of the Protein Data Bank (http://www.rcsb.org/pdb/home/home.do), TPR1 to TPR3 of EcTom70 recovered only significant matches (6.8e^{-2}) to other chaperone binding proteins, such as Tom70s, Hop, CHIP, and FKBP52 (52). A sequence-structural alignment of Tom70 clamps from *E. cuniculi* and select fungi and animals was generated (Fig. 1B). Contact residues in the known structures that interact with the EEVD motif, in particular the dicarboxylate anchor, are shown in red (Fig. 1B). These residues are conserved in the clamp sequences, indicating the importance of this dicarboxylate interaction (42, 52). Further, residues important for packing interactions of the TPRs are also conserved (Fig. 1B, bold) (42). While the residues for packing of the clamp domain are conserved in EcTom70, those residues that specifically interact with the terminal aspartate are not (Fig. 1B and C). These data suggest that while EcTom70 would likely form the clamp structure, it would be unable to form a dicarboxylate anchor with the EEVD peptide of chaperones.

To investigate the implications of this unusual clamp domain, the genome sequence for *E. cuniculi* was searched for molecular chaperones that could interact with TPR-clamp domains. No chaperones were identified with C-terminal EEVD, and a search of all *E. cuniculi* open reading frames equal to or greater than 40 codons revealed no such terminal sequence encoded in *E. cuniculi*. The *E. cuniculi* genome encodes three isoforms of Hsp70 and one of Hsp90, but none have a C-terminal sequence of EEVD. The only candidate with a similar motif is EcHsp90 (accession no. NP_584635), with the C-terminal sequence EEVQ. The presence of a cytosolic chaperone ending in EEVD (either Hsp70, Hsp90, or both) is virtually ubiquitous in eukaryotes (Fig. 1D) (6, 42), implying a critical function for this chaperone motif. Thus, the microsporidian *E. cuniculi* is exceptional: the binding site of the clamp domain of EcTom70 is modified, and the C terminus of chaperone EcHsp90 ends in the sequence EEVQ. Notably, loss of a negative charge in the chaperone C terminus coincides with loss of conserved positively charged residues in the clamp binding pocket (Fig. 1C and D).

EcTom70 targets mitochondria in yeast. Microsporidia represent an experimentally intractable group, with no systems currently available to genetically transform or manipulate these parasites. We therefore used *S. cerevisiae* as a model fungus in which to express EcTom70 and test for targeting and function consistent with a role in mitochondrial protein import. EcTom70 was fused with the reporter protein GFP, and fluorescence microscopy showed that EcTom70 colocalized with the mitochondrial stain MitoTracker in yeast (Fig. 2), with no detectable localization in other cell compartments. The protein is present in membrane extracts from yeast, specifically in mitochondrial membranes (Fig. 2). To test if EcTom70 is targeted to the yeast mitochondrial outer membrane, with cytosolic orientation consistent with TOM function, we analyzed the topology of EcTom70. Mitochondria were purified from yeast cells expressing EcTom70 and treated with trypsin, with immunoblotting being used to test for exposure of EcTom70-GFP to protease degradation. Under these conditions both EcTom70-GFP and endogenous ScTom70 are degraded by trypsin, leaving intact mitochondrial (Fig. 3A). With the same treatment, the intermembrane space protein cytochrome *b*2 (Cyb2) is not degraded, implying that the outer membrane is intact and, thus, EcTom70-GFP, like ScTom70, is attached to the cytosolic face of the outer membrane. Only with disrup-
tion of the outer mitochondrial membrane by mild osmotic shock is Cyb2 susceptible to trypsin (Fig. 3A). Sodium carbonate extraction of mitochondrial membranes shows that EcTom70 is anchored as an integral membrane protein, not as a peripheral membrane protein (Fig. 3B). This was also the case for endogenous ScTom70 and the integral membrane protein porin, whereas Hsp70, a peripheral component of the mitochondrial inner membrane and matrix, was enriched in the soluble fraction.

**EcTom70 does not complement ScTom70/Tom71 knockouts.** Though EcTom70 is assembled correctly in the mitochondrial outer membrane, it cannot replace the function of ScTom70.

*S. cerevisiae* contains two paralogues of Tom70 (ScTom70 and ScTom71), and so the *tom70/tom71* strain was used for

**FIG. 1.** (A) Schematic of the conserved domain structure of Tom70s from yeast (*S. cerevisiae*) and *E. cuniculi*. The three functional domains consist of (i) the transmembrane anchor (TM), (ii) the clamp domain, and (iii) the core domain. TPRs determined by crystal structure (ScTom70) or predicted by REP (EcTom70) are shown in yellow, α-helices are shown in green, and helix-turn-helix motifs are shown in orange. (B) Sequence-structural alignment of TPR clamp domains that bind chaperones Tom70 and/or Tom90. Conserved residues implicated in electrostatic interactions with the terminal EEVD motif of chaperones are shown in red. Conserved residues implicated in TRP packing interactions between α-helices are shown in bold black. **(C)** EcTom70 clamp domain (residues 34 to 143) modeled on the structure of human Hop TPR1 clamp domain in complex with an Hsp70 peptide (EEVD) (42) (PDB 1ELW). The bound EEVD peptide is shown in thick-stick mode. The five conserved clamp domain residues highlighted red in panel B are shown in narrow-stick mode with the conserved residue identity shown before the slash and the EcTom70 residue after the slash. **(D)** Hsp70 and Hsp90 alignment of C termini representing diverse eukaryotes. The conservation of the terminal EEVD is almost ubiquitous with the exception of microsporidian *E. cuniculi*.
complementation tests (11). These cells show a strong growth defect at 37°C on nonfermentable media (YPEG). When EcTom70 was expressed in Δtom70/tom71 cells to test for complementation, these cells showed growth equivalent to that of the Δtom70/tom71 cells under all growth conditions (Fig. 4).

**Further Tom/Tim homologues identified from E. cuniculi genomic data.** Given that microsporidia are most closely related to fungi and therefore diverged within the fungal-animal lineage, the ancestral microsporidia most likely possessed the common Tom and Tim proteins seen in both fungi and animals today. However, previous sequence similarity searches (BLAST) identified putative homologues of only three such proteins in *E. cuniculi*—Tom70, Tom40, and Tim22 (10, 28). Since many TOM and TIM proteins show weak conservation of primary sequence but greater conservation of structural features, HMM searches have proven a superior tool for identifying diverse TOM/TIM homologues in other lineages (15). Therefore, HMM searches were employed to comprehensively screen the *E. cuniculi* genomic data for candidate proteins of the mitochondrial protein import machinery. HMMs were generated for 26 known homologues of mitochondrial import proteins, focusing on those proteins characterized in fungi (see Materials and Methods).

Even using 26 HMMs, only seven candidate proteins were identified from *E. cuniculi* with significant matches. These include recovery of Tom70, Tom40, and Tim22 and the mitochondrial matrix chaperone Hsp70. In addition to these previously annotated gene products, candidates were identified for an essential component of the TIM23 complex in the mitochondrial inner membrane, Tim50; a J protein of the PAM complex, Pam16; and the major component of the SAM complex, Sam50. While some of the smaller proteins (e.g., the small Toms and tiny Tims) may be difficult to recognize because of the relative simplicity of these short sequences, the absence of matches for many of the import proteins is conspicuous given that they are readily recovered from other fungal and animal genomes. For example, Tom20 and Tom22 are conserved proteins that occur in all fungi and animals (34, 35). To validate the negative outcomes in *E. cuniculi*, we searched the UniProt database with the Tom20 and Tom22 models and recovered numerous strong matches to new sequences from diverse fungal and animal taxa (e.g., for Tom20 and Tom22...
matches were found in basidomycete fungus *Ustilago maydis* [9.20e⁻⁵⁴ and 4.20e⁻²⁵, respectively], the starfish *Nematostella vectensis* [1.50e⁻⁵⁶ and 1.70e⁻³⁴, respectively], and even the ancestral lineage of choanoflagellate *Monosiga brevicollis* [1.20e⁻¹⁷ and 8.40e⁻¹³, respectively]). Thus, failure to identify any candidates from *E. cuniculi* likely indicates their absence and indicates that the import apparatus for mitochondrial proteins in *E. cuniculi* is heavily reduced in complexity (Fig. 5).

**DISCUSSION**

The discovery in anaerobic eukaryotes of mitosomes that represent organelles derived from mitochondria provides a fascinating opportunity to examine reductive evolution of an organelle (49). The mitosomes of microsporidia provide an especially useful model due to microsporidia having diverged from within the fungal and animal lineages, both of which serve as well-studied models for mitochondrial biogenesis and biology. Organelle biogenesis is heavily dependent on the import of proteins encoded in the nucleus and translated in the cytoplasm. In the case of mitosomes, which now lack a genome, they are wholly dependent on this process. Common features of this organelle protein import machinery shared by fungi and animals allow us to predict the features of this machinery that likely occurred in the ancestral mitochondria of microsporidia. Thus, from examination of the existing machinery in microsporidia we can infer the nature of any change that has occurred during organelle reduction. Given that the microsporidian *E. cuniculi* has demonstrated a very strong trend for genomic and biochemical reduction and that the function and proteome of the mitosome are also apparently heavily reduced (28), we speculate that the protein import machinery of the mitosome might also show great change.

**Modification of *E. cuniculi* Tom70.** The protein receptors of the TOM complex are the first points of discrimination for proteins to be imported to mitochondria (5, 36). Tom70 serves as one of the two major receptors in fungi and animals, with a dual binding capacity for mitochondrial preproteins and, independently, molecular chaperones bound to these preproteins (9, 11, 52). A crystal structure of the Tom70 from yeast shows that TPRs are central to this structure and assemble as two semi-independent domains (51). Protein sequence similarity was used to identify a putative Tom70 homologue from *E. cuniculi* genomic data (28), and we have substantiated this identification by screening the *E. cuniculi* data with a Tom70 HMM and recovering “EcTom70” with a high confidence value. Targeting of heterologously expressed EcTom70 to the yeast mitochondrion, where it correctly inserts as an integral protein of the outer membrane, provides further strong evidence that this is a genuine Tom70 homologue and that it likely functions in the microsporidian mitosome. Moreover, it indicates that the requirements for targeting proteins to the outer membrane and membrane insertion are conserved between fungi and microsporidia.

The reduced size of EcTom70 compared to that of yeast is consistent with overall shortening of proteins observed in *E. cuniculi* (28). The elimination of the two linker regions between the structurally conserved domains in EcTom70 likely reflects preservation of protein functional regions under a

---

**FIG. 5.** Schematics of the protein import machinery and pathways in yeast mitochondria (A) and *E. cuniculi* mitosome (B) based on identified homologues of the general fungal/animal pathways. Protein components of the yeast system were all represented by HMMs used to search the microsporidian genomic data and represent the major presequence-dependent and presequence-independent pathways. Homologues identified in *E. cuniculi* indicate a severely reduced import apparatus utilizing elements of the presequence-independent pathway.
strong pressure to reduce size. The further loss of one of eight TPRs of the core domain might also reflect a tolerable loss without affecting the function of this region in binding to mitosomal preproteins. Despite these losses, all other major structural features identified in the yeast Tom70 structure appear to have been conserved. Lack of function of EcTom70 in the yeast TOM complex might reflect divergence of this more compact Tom70 and loss of EEVD-mediated chaperone binding. We note, however, that human Tom70, which can bind the EEVD sequences, is also unable to complement yeast Tom70 mutants (52).

**Loss of TOM/TIM homologues.** Our use of HMM to search *E. cuniculi* data shows that candidates for Tom70, Tom40, and Tim22 are identified with high confidence, and additional new components of the TIM complex (Tim50 and Pam16) and the SAM complex (Sam50) were discovered. These results offer further evidence of a conserved import system, albeit in relatively minimal form (Fig. 5).

Conspicuously absent from the predicted import machinery in *E. cuniculi* are the other outer membrane receptors Tom20 and Tom22. Tom20 is specific to the fungus-animal lineage, is a well-conserved protein throughout this group, and is reliably recovered with HMMs throughout fungal and animal lineages (34). Tom70 is chiefly responsible for recognizing presequences at the N terminus of many mitochondrial proteins. Tom22 assists the function of Tom20, passing precursor proteins on to the translocation pore, Tom40 (reviewed in reference 27). Tom22 is widely conserved throughout eukaryotes and is reliably recovered with similarity searches (35). Tom70, on the other hand, has a greater role in importing membrane proteins, since it binds hydrophobic stretches, e.g., those characteristic of mitochondrial transmembrane solute carrier proteins. Microsporidian mitosomal proteins have shown a tendency for loss of N-terminal targeting extensions and a greater reliance of internal signals for targeting, based on heterologous targeting to yeast mitochondria (10). It is conceivable that EcTom70 is sufficient for recognition of the remnant mitosomal proteins and facilitates translocation through the EcTom40 channel.

A further reduction of the mitosomal import machinery is indicated by identification of only one member of the Tim23/Tim17/Tim22 family of proteins in *E. cuniculi*. Only the HMM representing Tim22 identified a homologue, and therefore this protein is designated EcTim22. Ancient duplications of the inner membrane pore protein are represented widely throughout eukaryotes by paralogues Tim22, Tim23, and Tim17 (15, 43). In yeast and other fungi and in animals Tim23 and Tim17 together contribute to presequence-mediated import across the inner membrane, while Tim22 contributes to a separate complex dedicated to insertion of membrane proteins such as the carrier proteins of the inner membrane (Fig. 5) (5, 36). The specialization of two TIM complexes early in eukaryotic evolution enabled diverse mitochondrial protein traffic. The presence of only a single inner membrane pore in *E. cuniculi* suggests that microsporidia have reversed this specialization in the face of reduced protein traffic and overall cellular reduction. We suggest that the protein called EcTim22 is probably assisted by EcTim50 to form a TIM23 complex: Tim50 both regulates the inner membrane barrier and binds to emergent proteins from the TOM complex and hence is essential for TIM function. A single J-protein, EcPam16, is likely responsible for regulation of the ATP-dependent role of mHsp70 during final passage across the inner membrane (Fig. 5). As is the case in trypanosomatids (43), the single TIM complex in *E. cuniculi* mitosomes might drive translocation of proteins into the matrix and assembly of inner membrane proteins, too.

Identification of a Sam50 homologue, for insertion of β-barrel proteins (e.g., Tom40), reflects the essential function of the SAM complex in mitosomes, as in mitochondria (5, 36). It is unclear whether failure to identify further SAM complex proteins (Sam35, Sam37, and Mdm10) is due to poor conservation of these proteins or to a capacity of Sam50 to act alone. Another lone complex component is Erv1, of the MIA (mitochondrial intermembrane space import and assembly) machinery (5, 36). Erv1, identified as a likely mitosomal protein in the genome annotation of *E. cuniculi*, is implicated in promoting the sequential formation of intramembrane disulfides in imported intermembrane space proteins. These molecules provide further evidence of relics of a mitochondrial-type protein import system.

The skeletal form of the mitosomal protein import machinery identified here might reflect the difficulty in identifying homologues of many of the import proteins of fungi and animals, particularly several of the small proteins. Presently few genomic data exist for microsporida aside from *E. cuniculi*, limiting the opportunity to look more broadly for mitosomal homologues in this group. It is possible, however, that our observations reflect an import machinery that has been reduced in response to dramatic reductionism seen throughout microsporidian biology. It is notable that, of the subset of import machinery that the HMMs have identified, all major essential functions are represented, although only once rather than in duplicate specialist form as seen in fungi and animals (Fig. 5). Reduction and change in EcTom70 and apparent loss of Tom20 and one of the inner membrane pore complexes are all consistent with conversion to a minimal apparatus. Moreover, use of Tom70 as the principal outer membrane receptor is supported by the loss of presequences from many mitosomal proteins and a greater role of internal signals for organelle import. If such changes hold across microsporidia, these insights offer new approaches to tackling microsporidia as human and animal pathogens. A peptide (ending in EEVQ) that specifically targets and blocks the EcTom70 clamp domain, for instance, could offer scope for perturbing mitosomal protein import as an antimicrosporidal chemotherapeutic strategy.

**ACKNOWLEDGMENTS**

We are grateful to E. S. Didier (Tulane University, Louisiana) for providing *E. cuniculi* genomic DNA.

This work was supported by grants from the Australian Research Council (R.F.W., T.D.M., and T.L.). R.F.W. was supported by an Australian National Health and Medical Research Council Peter Doherty Fellowship (no. 145896). T.L. is an ARC Federation Fellow.

**REFERENCES**

1. Abe, Y., T. Shodai, T. Muto, K. Mihara, H. Torii, S. Nishikawa, T. Endo, and D. Kohda. 2000. Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. Cell 100:551–560.

2. Andrade, M. A., C. P. Ponting, T. J. Gibson, and P. Bork. 2000. Homology-based method for identification of protein repeats using statistical significance estimates. J. Mol. Biol. 298:521–537.

3. Arisz, N., L. B. Sanchez, L. M. Weiss, M. Muller, and T. Hashimoto. 2002.
Mitochondrion-type hsp70 genes of the amitochondriate protists, *Giardia intestinalis*, *Entamoeba histolytica* and two microsporidia. Parasitol. Int. 51:9–16.

4. Bauer, M. F., S. Hofmann, W. Neupert, and M. Brunner. 2000. Protein translocation into mitochondria: the role of TIM complexes. Trends Cell Biol. 10:25–31.

5. Bolender, N., A. Sickmann, R. Wagner, C. Meisinger, and N. Pfanner. 2008. Multiple pathways for sorting mitochondrial precursor proteins. EMBO Rep. 9:842–49.

6. Brinker, A., C. Scheufler, F. Von Der Mulbe, B. Fleckenstein, C. Herrmann, G. Jung, I. Moarefi, and F. U. Hartl. 2002. Ligand discrimination by TPR domains. Relevance and selectivity of EYEV-recognition in Hsp70 x Hsp90 complexes. J. Biol. Chem. 277:19265–19275.

7. Brix, J., K. Dietmeier, and N. Pfanner. 1997. Differentiation of precursors by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. J. Biol. Chem. 272:20730–20735.

8. Brix, J., S. Rudiger, B. Bukan, J. Schneider-Mergener, G. E. Schulz, and N. Pfanner. 1999. Distribution of binding sequences for the mitochondrial import receptors Tom20, Tom22, and Tom70 in a precursor-carrying preprotein and a non-cleavable preprotein. J. Biol. Chem. 274:16522–16530.

9. Chan, N. C., V. A. Likic, R. F. Waller, T. D. Mulhern, and T. Lithgow. 2000. The mitochondrial import receptor Tom70: identification of a 25 kDa core domain with a specific binding site for preproteins. J. Mol. Biol. 303:479–488.

10. Burri, L., B. Walli, C. D. Bursac, T. Lithgow, and P. J. Keeling. 2006. Microsporidian mitosomes retain elements of the general mitochondrial targeting system. Proc. Natl. Acad. Sci. USA 103:15916–15920.

11. Chan, C. N., V. A. Likic, R. F. Waller, T. D. Mulhern, and T. Lithgow. 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.

12. Cuff, J., M. Clamp, A. Siddiqui, M. Finlay, and G. Barton. 1998. IPred: a consensus secondary structure prediction server. Bioinformatics 14:892–893.

13. Daum, G. P., C. Bohni, and G. Schatz. 1982. Import of proteins into mitochondria. Cytochrome b2 and cytochrome c peroxidase are located in the intermembrane space of yeast mitochondria. J. Biol. Chem. 257:13026–13033.

14. Didier, E. S., and L. M. Weiss. 2006. Microsporidiosis: current status. Curr. Opin. Infect. Dis. 19:485–492.

15. Dolezal, P., V. Likic, J. Tacheye, and T. Lithgow. 2006. Evolution of the molecular machines for protein import into mitochondria. Science 313:514–518.

16. Eddy, S. R. 1996. Hidden Markov models. Curr. Opin. Struct. Biol. 6:361–365.

17. Eddy, S. R. 1998. Profile hidden Markov models. Bioinformatics 14:755–763.

18. Edlund, T. D., J. Li, S. Vivesvara, M. H. Vodkin, G. L. McLaughlin, and S. K. Katiyar. 1996. Phylogenetic analysis of β-tubulin sequences from amitochondrial protozoa. Mol. Phylogen. Evol. 4:359–367.

19. Fast, N. M., and P. J. Keeling. 2001. Alpha and beta subunits of pyruvate dehydrogenase from the microsporidian *Nosema locustae*: mitochondrial-derived carbon metabolism in microsporidia. Mol. Biochem. Parasitol. 117:201–209.

20. Gernot, A., H. Philippe, and H. Le Guyader. 1997. Evidence for loss of mitochondria in *Microsporidia* from a mitochondrial-type HSP70 in *Nosema locustae*. Mol. Biochem. Parasitol. 87:159–168.

21. Gill, E. E., and N. M. Fast. 2006. Assessing the microsporidium-fungi relationship: combined phylogenetic analysis of eight genes. Gene 375:103–109.

22. Goldberg, A. V., S. Mokli, A. D. Tsoumis, K. Neumann, G. Kuhnke, F. Delbac, C. P. Vivares, R. P. Hirt, R. Lill, and T. M. Embley. 2008. Localization and functionality of microsporidian iron-sulphur cluster assembly proteins. Nature 452:624–628.

23. Hauke, V., K. Horst, G. Schatz, and T. Lithgow. 1996. The Mas20p and Mas70p subunits of the protein import receptor of yeast mitochondria interact via the tetratricopeptide repeat motif in Mas20p: evidence for a single hetero-oligomeric receptor. EMBO J. 15:1231–1237.

24. Hirt, R. P., B. Healy, C. R. Vossbrinck, E. U. Canning, and T. M. Embley. 1997. A mitochondrial Hsp70 orthologue in *Vesicomya necatrix*: molecular evidence that microsporidia once contained mitochondria. Curr. Biol. 7:595–599.

25. Hirt, R. P., J. M. Logdon, Jr., B. Healy, M. W. Dories, W. F. Doolittle, and T. M. Embley. 1999. Microsporidia are related to fungi: evidence from the largest subunit of RNA polymerase II and other proteins. Proc. Natl. Acad. Sci. USA 96:580–585.

26. Hofmann, K., and W. Stoffel. 1993. TMbase—a database of membrane spanning proteins segments. Biol. Chem. 374:166.