Charles University, Faculty of Science

Univerzita Karlova, Přírodovědecká fakulta

Ph.D. study program: Parasitology

Doktorský studijní program: Parazitologie

Abhijith R. Makki

Protein translocation into hydrogenosomes of *Trichomonas vaginalis*

Translokace proteinů do hydrogenosomů *Trichomonas vaginalis*

Ph.D. Thesis

Thesis supervisor: Prof. RNDr. Jan Tachezy, Ph.D.

Prague, 2019

Declaration of the author

I declare that I have prepared this thesis independently. I also proclaim that the literary sources have been properly cited and neither this work nor the substantial part of it has been used to reach the same or any other academic degree.

Abhijith R. Makki

Declaration of the thesis supervisor

The data presented in this thesis resulted from a team collaboration at the Laboratory of Molecular and Biochemical Protistology and from a cooperation with our collaborators. I declare that the involvement of Mr. Abhijith R. Makki in this work was substantial and that he contributed significantly to obtain the results.

Prof. RNDr. Jan Tachezy, Ph.D.

Acknowledgements

The period spent during my Ph.D. study was fruitful and very satisfying. First and foremost, I would like to thank my supervisor and mentor, Jan Tachezy for his advice, unwavering support and encouragement. Many thanks to my colleagues in the laboratory, core facilities and our collaborators who helped me during the work. I am also grateful to the Grant Agency of the Charles University (GAUK) and the Mobility Funds for providing me funds for chemicals, conferences, research stay and travel during my doctoral study. My wife and my parents have always supported me, and their role is often underappreciated. So, I take this opportunity to thank them.

TABLE OF CONTENTS

ABSTRACT

Mitochondria carry out several important functions in eukaryotic cells such as energy metabolism, iron-sulfur cluster assembly, apoptosis, signaling pathways, protein quality control etc. Most mitochondrial proteins are synthesized on the cytosolic ribosomes and transported to the organelles by the cytosolic chaperones and mitochondrial protein import machinery based on specific targeting signals. Although, the basic principles of protein import have been explained, many questions remain unanswered, particularly for highly modified mitochondria such as hydrogenosomes. The aim of the study was to investigate protein translocation into hydrogenosomes of a human parasite, *Trichomonas vaginalis* (Tv) with a focus on the composition, function and structure of protein translocases and the role of targeting signals.

The translocase of the outer membrane (TOM) is responsible for the import of most proteins into the organelle. Even though, the presence of a TOM complex in trichomonad hydrogenosomes was predicted, its components were not known. Moreover, the generic structure of the mitochondrial TOM complex was not resolved. This study showed that the TvTOM complex is highly divergent consisting of two modified core subunits – channel-forming TvTom40 isoforms and a Tom22-like protein, and two lineage-specific subunits – Tom36 and Tom46 that most likely, function as receptors. Additionally, TvTOM forms a stable supercomplex with Sam50 that is involved in the biogenesis of β-barrel proteins. Electron microscopy revealed that the translocase has a triplet-pore structure with a unique skull shape.

Mitochondrial matrix preproteins carry an N-terminal targeting sequence (NTS). Interestingly, a glycolytic enzyme, ATP-dependent phosphofructokinase (ATP-PFK) that does not contain a predictable NTS localizes to hydrogenosomes. Localization experiments suggested that TvATP-PFK and its homologous ATP-PFKs from yeast and *E. coli* possess unknown internal targeting signal (ITS) that is possibly recognized by the protein import machinery. From an evolutionary perspective, the ability of mitochondria and hydrogenosomes to recognize proteins such as ATP-PFK may represent an ancient mechanism from the early phases of organelle evolution whereas, NTS-dependent import might have evolved later. *T. vaginalis* has several unique tail-anchored (TA) proteins, a class of integral membrane proteins that localize to the hydrogenosomal outer membrane, including the newly characterized TvTOM subunits. Analyses of physico-chemical properties

and localization experiments identified new traits for hydrogenosomal TA protein targeting such as higher net positive charges in the C-terminal segment which, otherwise are primarily for peroxisomal TA proteins in aerobic eukaryotes, and a slightly longer transmembrane domain when compared to mitochondrial TA proteins.

Taken together, these studies show that the protein import into hydrogenosomes is rather divergent compared to that of mitochondria. The triplet-pore TOM complex, composed of conserved core subunits was present in the last common eukaryotic ancestor while, the peripheral receptors evolved independently in different eukaryotic lineages. The changes observed in the protein translocases and the targeting signals most likely reflect the adaptation of hydrogenosomes to anaerobic conditions, particularly, the loss of respiratory chain complexes that resulted in low or absence of membrane potential.

ABSTRAKT (CZECH)

Mitochondrie plní řadu významných funkcí v eukaryotických buňkách, jako je energetický metabolismus, syntéza železo-sirných center, apoptóza, bunečná signalizace, kontrola kvality proteinů atd. Většina mitochondriálních proteinů je syntetizována na cytosolických ribozomech a transportována do organel za pomocí cytosolických chaperonů a mitochondriálních membránových translokáz, které rozpoznávají specifické adresové sekvence. Přestože základní principy importu proteinů jsou známé, mnoho otázek zůstává nezodpovězeno, zejména u vysoce modifikovaných mitochondrií, jako jsou hydrogenosomy. Cílem této studie bylo prozkoumat translokaci proteinů do hydrogenosomů lidského parazita *Trichomonas vaginalis* (Tv), se zaměřením na složení, funkci a strukturu proteinových translokáz a roli adresových sekvencí.

Translokáza vnější mitochondriální membrány (TOM) je zodpovědná za import většiny proteinů do mitochondrií. Ačkoliv přítomnost komplexu TOM v hydrogenosomech trichomonád byla predikována na základě analýzy genomu, jednotlivé složky komplexu nebyly známy. Navíc ani celková struktura mitochondriálního komplexu TOM nebyla zcela vyřešena. Tato studie ukázala, že komplex TvTOM je velmi divergentní, sestávající se ze dvou modifikovaných základních podjednotek – TvTom40, který tvoří translokační kanál a protein podobný Tom22, a dále ze dvou podjednotek specifických pro linii trichomonád – Tom36 a Tom46, které pravděpodobně fungují jako receptory. TvTOM navíc tvoří stabilní superkomplex se Sam50, který se podílí na biogenezi β-barelových proteinů. Elektronová mikroskopie odhalila, že studovaná translokáza obsahuje triplet pórů a má neobvyklý tvar připomínající lebku.

Preproteiny, které jsou určeny pro transport do mitochondrialní matrix obvykle nesou N-terminální adresovou sekvenci (NTS). Je proto zajímavé, že glykolytický enzym, ATP-dependentní fosfofruktokináza (ATP-PFK) je importován do hydrogenosomů bez predikovatelné NTS. Lokalizační experimenty prokázaly, že TvATP-PFK a homologní ATP-PFK z kvasinek a *E. coli* mají neznámé interní adresové signály (ITS), které jsou rovněž rozpoznávány importní mašinerií. Z evolučního hlediska může být schopnost mitochondrií a hydrogenosomů rozpoznávat proteiny jako je ATP-PFK na základě ITS původním mechanismem, zatímco import, který závisí na NTS se objevil později. *T. vaginalis* má řadu proteinů kotvených C-terminální doménou (tail-anchored, TA) ve vnější hydrogenosomální membráně, včetně nově charakterizovaných podjednotek TvTOM. Analýzy fyzikálně-

chemických vlastností a lokalizační experimenty ukázaly specifické vlastnosti hydrogenosomálních TA proteinů, jako je vyšší kladný náboj v C-terminálním segmentu, který je v ostatních eukaryotických buňkách charakteristický pouze pro peroxisomální TA proteiny, a ve srovnání s mitochondriálními TA proteiny, hydrogenosomální TA proteiny mají delší transmembránovou doménu.

Celkově tyto studie ukazují, že import proteinů do hydrogenosomů je ve srovnání s mitochondriemi v mnoha aspektech odlišný. Ačkoliv komplex TOM s tripletem pórů složených ze tří centrálních podjednotek, byl patrně přítomen již u posledního společného předka eukaryotických organismů, periferní receptory se vyvíjely nezávisle v různých eukaryotických liniích. Změny pozorované ve struktuře translokáz a adresových sekvencích u hydrogenosomů s největší pravděpodobností odrážejí adaptaci hydrogenosomů k anaerobním podmínkám, zejména ztrátu respiračních komplexů, která vedla ke snížení nebo ztrátě membránového potenciálu.

1. INTRODUCTION

1.1 Mitochondrial research history

Mitochondria were first observed as intracellular structures in the 1840s. However, it was not until 1890 that Richard Altmann recognised them because of their structure, referred to as "bioblasts" and concluded them to be "elementary organisms" living inside cells, carrying out some important functions [1]. In 1898, Carl Benda named the compartment as Mitochondrion ("*mitos*" - thread, "*chondrion*" - granule-like/grain-like in Greek language). For the next few decades, various biochemical studies deduced that mitochondria could be the sites of energy production. This led many scientists to turn their focus towards mitochondrial research and procedures were developed to isolate mitochondria through subcellular fractionation [1]. In 1950s, the electron micrograph showed that these organelles have a double membrane.

1.2 Mitochondrial origin and evolution

Mitochondria are centres of numerous important functions in the eukaryotic cell such as energy metabolism, combat oxidative stress, protein quality control, cellular signalling, iron-sulfur cluster assembly etc. Mitochondrion originated from a proteobacterium, that formed an endosymbiotic relationship with an early-eukaryotic cell or archaebacterium [2,3]. Gradually, the endosymbiont lost its capacity to function as an independent entity when most of the endosymbiotic genes were either transferred to the host nucleus or simply lost [4]. Consequently, the proteins required for the endosymbiont (pre-mitochondrion) were synthesized on the cytoplasmic ribosomes and transported to the evolving organelle. A few pioneering reviews have pointed out that the loss of genes in the endosymbiont acted as a selected pressure to engineer and install protein import machinery, and because most subunits do not have homologues in bacteria, it has been hypothesized that the eukaryotic system developed these modules *de novo* [5,6].

1.3 Mitochondria and mitochondria-related organelles (MROs)

Most eukaryotic model organisms possess mitochondria, which synthesize ATP through aerobic respiration with the help of proton-pumping electron transport chain and ATP synthase, conduct tricarboxylic acid (TCA) cycle, have ADP/ATP carriers, a genome and

translational machinery and form cristae. However, a considerable number of organisms that dwell in oxygen-limited conditions have highly modified forms of mitochondria such as hydrogenosomes and mitosomes, grouped under mitochondria-related organelles (MROs) [7]. Hydrogenosomes were first discovered as a hydrogen-producing and pyruvate-metabolizing organelle in a bovine parasite, *Tritrichomonas foetus* [8,9]. Subsequent studies found that these organelles are bound by a double-membrane, have heat shock proteins, Hsp70, Hsp60 and Hsp10, protein import pathways that are characteristic to mitochondria, mitochondrial iron-sulfur cluster assembly machinery and remnant mitochondrial complex I [10–17]. Hydrogenosomes are present in various anaerobic parasitic and free-living eukaryotes such as trichomonads (*Trichomonas vaginalis* and *T. foetus*), archamoebae (*Mastigamoeba balamuthi*), diplomonads (*Spironucleus salmonicida*), ciliates (*Nyctotherus ovalis, Trimyema sp., Dasytricha ruminantium, Plagiopyla frontata*, *Metopus concortus*, *Sonderia* sp., *Isotricha* sp. and *Entodinium* sp.), fungi (*Neocallimastix frontalis* and *Piromyces sp.* E2) and heteroloboseans (*Sawyeria marylandensis* and *Psalteriomonas lanterna*) [7,8,18–29]. All hydrogenosomes are not biochemically identical nevertheless, they are grouped under one class as they produce hydrogen and do not have an electron transport chain. Mitosomes are found in some parasitic species namely, *Giardia intestinalis*, *Entamoeba histolytica* and *Cryptosporidium parvum*, and are essential for iron-sulfur cluster assembly (Giardia) or sulfate-activation pathway (Entamoeba) [30,31]. Both hydrogenosomes and mitosomes are related to mitochondria and they originated from the same ancestral organelle or were derived from aerobic mitochondria [12,14,30,31]. These organelles have undergone significant functional reduction to suit the organism's lifestyle and environment that is reflected in their proteome.

Müller et al classified mitochondria and MROs into five classes – aerobic mitochondria, anaerobic mitochondria, hydrogen-producing mitochondria, hydrogenosomes and mitosomes based on the ability to synthesize ATP, produce hydrogen, use oxygen as terminal electron acceptor and presence of respiratory chain complexes [7]. However, the characterization of MRO in mainly free-living organisms revealed that they are rather mosaic featuring continuum traits from aerobic mitochondria to mitosomes. For example, the MRO in *Blastocystis hominis* has features from both mitochondria and hydrogenosomes and similarly, *Dysnectes brevis* has hydrogen-producing organelles (Hydrogenosomes) but, they cannot synthesize ATP and hence, resembling mitosomes [32,33]. Thus, recent investigations on MROs have created a blur between different classes of mitochondria, and it has been suggested that all MROs should be simply referred to as mitochondria rather than being

classified into subtypes [34]. More recently, it was discovered that *Monocercomonoides exilis* and other members of oxymonads completely lack a mitochondrial organelle although, it appears that the mitochondrion was lost secondarily [35].

The number of proteins in mitochondria present vary greatly in different organisms, for example 1,158 – 1,900 in humans, 1,098 in mouse, 1,008 in *Trypanosoma brucei* (procyclic form) and 901 in yeast [36–41]. Interestingly, the size of the mitochondrial proteome varies from one tissue to another within the same organism [36,42]. On the other hand, the proteome remains the same even if there is a switch in the metabolic activity [43]. Compared to mitochondria, hydrogenosomes of *T. vaginalis* has a smaller proteome with around 600 proteins [44,45]. The proteomes of mitosomes in *E. histolytica* and *G. intestinalis* are minimalistic, made up of only 95 and 139 putative proteins respectively [31,46].

1.4 Protein import into mitochondria

The protein import mechanism is indispensable for mitochondrial functioning and biogenesis. Mitochondria have four different sub-compartments where proteins are localized: outer membrane, intermembrane space (IMS), inner membrane and matrix. Most mitochondrial proteins are nuclear-encoded, synthesized on the cytosolic ribosomes and are brought to the mitochondrial surface by different cytosolic molecular chaperones of Hsp70 and Hsp90 families [47–49]. Protein translocation into mitochondria is mediated by the multi-subunit complexes of translocase of the outer membrane (TOM), sorting and assembly machinery (SAM) and mitochondrial import (MIM) complex, translocases of the inner membrane (TIM22/TIM23), presequence-associated motor (PAM) and mitochondrial IMS assembly machinery (MIA) (Fig 1) [50,51]. So far, five protein import pathways have been reported: presequence pathway, carrier pathway, MIA pathway, β-barrel pathway and MIM pathway (Fig 1) [51]. Matrix proteins are directed to TIM23 complex via TOM and then, are pulled into the matrix by PAM. Inner membrane proteins like metabolite carriers first cross the outer membrane via TOM, channelled towards TIM22 complex by the IMS-localized chaperones, Tim9-Tim10 and Tim8-Tim13 complexes and further, laterally released into the inner membrane by the TIM22 complex. Proteins that localize to the IMS cross the OM via TOM and are folded by Mia40, and Essential for respiration and vegetative growth protein 1 (Erv1). The β-barrel proteins residing in the mitochondrial outer membrane (MOM) are first recognised and internalized by the TOM complex, received in the IMS by Tim9-Tim10 chaperones and delivered to the SAM, where the β-barrels are assembled and released in the

membrane. Polytopic outer membrane proteins are received by Tom70, a receptor of the TOM complex, and in conjunction with MIM complex are assembled in the outer membrane. Tail-anchored (TA) and signal-anchored proteins which carry a single transmembrane domain at the C- or N-termini respectively are inserted to the MOM without the aid of the TOM complex. However, if TA and signal-anchored proteins are a part of TOM, their biogenesis requires the pre-existing components of TOM complex [52].

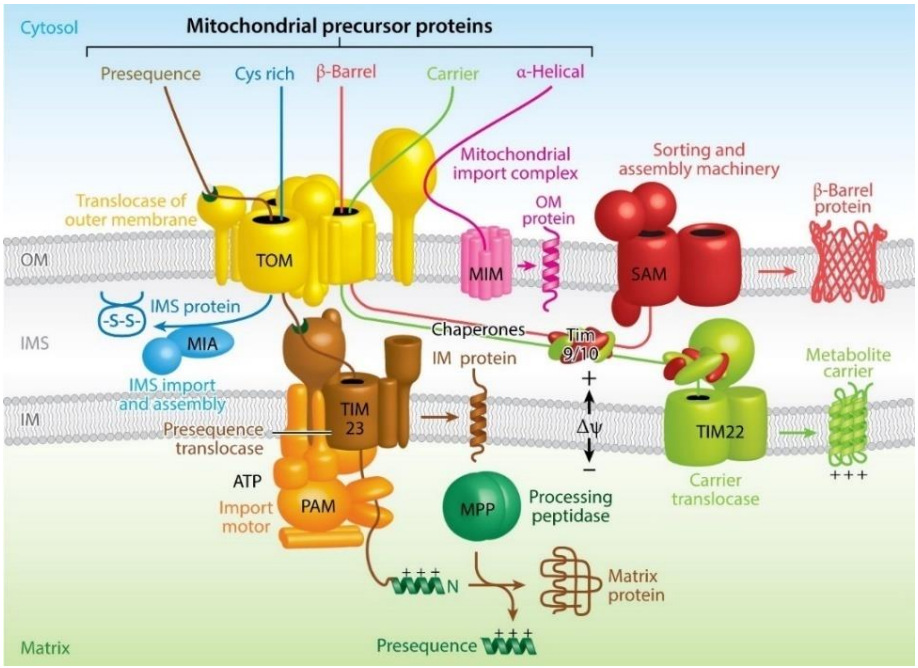


Figure 1. Overview of the five major protein import pathways of mitochondria. Presequence-carrying preproteins are imported by the TOM and TIM23 complexes. Proteins with a hydrophobic sorting signal can be released into the inner membrane, whereas hydrophilic proteins are imported into the matrix with the help of PAM. The mitochondrial processing peptidase (MPP) removes the presequences. Cysteine-rich proteins of the IMS are imported by TOM and MIA, which inserts disulfide bonds in the imported proteins. The precursors of β-barrel proteins are translocated through TOM to the small TIM chaperones of the IMS and are inserted into the outer membrane by SAM. The precursors of metabolite carriers of the inner membrane are imported via TOM, small TIM chaperones, and the TIM22 complex. Several α-helical outer membrane proteins are imported by the MIM complex. The membrane potential (ψ) across the inner membrane drives protein translocation by the TIM23 and TIM22 complexes. Referenced from [51].

1.4.1 Mitochondrial targeting signals

Mitochondrial matrix precursor proteins and in some cases, preproteins of the IMS and inner membrane carry a cleavable N-terminal targeting sequence (NTS) while, proteins of the IMS,

inner and outer membranes carry non-cleavable internal targeting sequence (ITS), which acts as an entry ticket for their delivery to the organelle (Fig 2). The NTS contains positively charged, hydrophobic and hydroxylated amino acid residues that form an amphipathic α-helix to present a positively charged surface on one side and a hydrophobic surface on the other (Fig 2A) [53–55]. Once inside the matrix, the NTS of the preproteins is cleaved off by the mitochondrial processing peptidase (MPP) (Fig 2Ai). The origin and distribution of presequence on genes are quite intriguing. A few pioneering works have shown that the synthetic mitochondrial sequences could translocate passenger proteins across either artificial or bacterial lipid bilayers [53,56]. These reports led to a speculation that the presequence either existed or were developed prior to the existence of the mitochondrial protein translocases. The NTS in the hydrogenosomal proteins are considerably shorter and have significantly lower positive charge in comparison with the mitochondrial NTS (Fig 2Aii) [57]. Moreover, in the last few years, many proteins without any readily identifiable NTS were found to be targeted to *T. vaginalis* hydrogenosomes and yeast mitochondria [58,59]. Even the mutants with a deleted N-terminal segment ranging from 10-30 residues were found to be targeted to the organelles [58–60]. The NTS was also reported to be absent for proteins targeted to hydrogenosomes of *S. salmonicida* entirely and for several mitosomal matrix proteins in *G. intestinalis* [20,61]. These results revealed a new perspective that some matrix proteins could have unknown ITS to reach the mitochondrial matrix and this could be an ancient trait connected with the endosymbiotic origin (Fig 2Aiii) [59].

ITS is a loose term and includes a varied repertoire of signals that are non-cleavable either present as a part of primary sequence or formed as a secondary structure depending on the class of protein. Some inner membrane proteins have both an NTS and a hydrophobic sorting or stop-transfer signal that allow their insertion to the inner membrane (Fig 2Bi) [62,63]. Metabolite carriers of the inner membrane are very hydrophobic and carry typical multiple ITS along the primary sequence (Fig 2Bii) [50,64]. Some IMS proteins carry an NTS, which is first cleaved by MPP exposing a second region to be cleaved by the inner membrane peptidase (IMP) (Fig 2Ci) [62,65]. Many IMS proteins carry a cysteine rich Cx_3C, Cx_9C or similar motif with or without NTS (Fig 2Cii) [63].

The MOM hosts several α-helical single- (TA and signal-anchored proteins) and multi-spanning membrane proteins, which have their targeting information in the TMD and its flanking regions (Fig 2Di, ii and iii) [66–70]. β-barrel proteins that also reside in the MOM are recognised via a specialized β-hairpin motif, and their targeting to mitochondria depends on the hydrophobicity of the β-hairpin motif (Fig 2Div, β-hairpin motif – green box)

[71]. The last β-strand of all mitochondrial β-barrel proteins contains a conserved β-signal motif, PxGxxHxH (P – polar amino acid, x – any, G – glycine, H – hydrophobic amino acid) for their assembly to the MOM (Fig 2Div, TMD coloured in orange) [72].

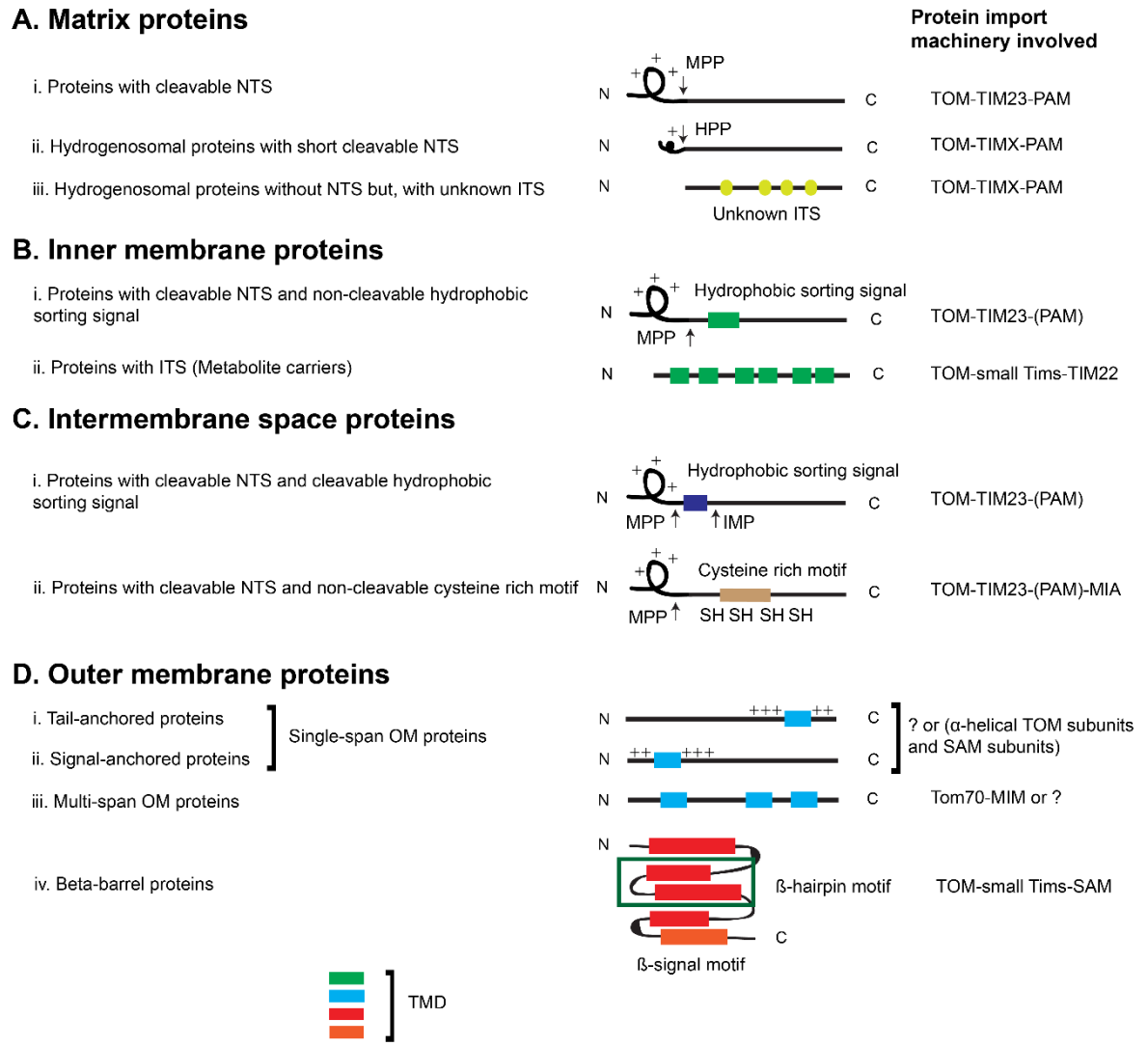


Figure 2. Targeting and sorting signals of mitochondrial precursor proteins. The mitochondrial precursor proteins carry cleavable or non-cleavable targeting signals. (A) The matrix proteins contain an N-terminal sequence (NTS) that is cleaved off by the MPP or hydrogenosomal processing peptidase (HPP) in the matrix of mitochondria and hydrogenosomes respectively. Some hydrogenosomal matrix do not have NTS but, are recognised by unknown ITS. (B) The inner proteins which have non-cleavable hydrophobic sorting signal in addition to an NTS are sorted to the inner membrane. Metabolite carriers of the inner membrane carry ITS for their targeting and insertion. (C) The inner proteins which have cleavable hydrophobic sorting signal in addition to an NTS is removed by the inner membrane peptidase (IMP). The biogenesis of some IMS proteins requires cysteine rich motif with or without NTS. (D) The MOM α-helical proteins have their targeting information in the TMD and its flanking regions. The β-barrel proteins are targeted to the MOM based on β-hairpin motif and β-signal motif. X = 17/22/23, ? = unknown.

1.4.2 Cytosolic chaperones and co-chaperones

Cytosolic chaperones and their co-chaperones are vital for the targeting of proteins to the correct organelle and other than sorting, they also prevent proteins from misfolding and aggregation and thus, maintaining the cellular homeostasis. Because the cytosolic factors bind to a variety of substrates, it was assumed that their interaction is mostly non-specific. However, an increasing number of evidences indicates that it is a much more specific and regulated process [47–49]. The biogenesis of mitochondrial proteins is promoted by the chaperones of Hsp70 and Hsp90 families and their co-chaperones of Hsp40 family. In addition to the protein precursors, the chaperones also interact with the protein import receptors [47–49]. The chaperones present the precursors in such a manner that the import receptors on mitochondria can recognise the targeting signals and can initiate the import process.

1.4.3 Mitochondrial protein import machinery

For an efficient import of proteins, mitochondria employ TOM complex, SAM and MIM in the outer membrane and TIM23 and TIM22 complexes, PAM in the inner membrane and MIA and the hetero-oligomeric TIM chaperones in the IMS. It is increasingly becoming clear that the mitochondrial protein import machineries do not work in isolation as once assumed but, are highly dynamic in real time and keep associating and dissociating with different proteins or other protein translocases in the membranes for specialized functions and crosslink the protein biogenesis with various mitochondrial processes [41].

A. Translocase of the Outer Membrane (TOM)

Most proteins targeted to mitochondria enter through a high molecular weight complex known as the TOM complex. The TOM complex acts as a sorting centre and directs the translocating proteins to their respective destination – outer membrane, IMS, inner membrane or matrix. At the core of the complex is a β-barrel protein, Tom40 that forms a protein-conducting channel across the outer membrane [73,74]. Both Tom40 and voltage-dependent anionic channel (VDAC) belong to the family of mitochondrial porins; are composed of 19 anti-parallel β-strands and share a common ancestry [75]. The stretches forming the Tom40 channel contain both positively charged acidic and hydrophobic amino acid residues. A detailed cross-linking approach by Shiota et al showed that a soluble matrix protein carrying

positively charged presequence takes an acidic path while, the mitochondrial carrier protein takes a hydrophobic path and thus, the two different classes of proteins follow distinct paths inside the channel [76]. Every second residue of the Tom40 barrel structure faces the interior of the channel whereas, the other residue faces the exterior [76]. The long N-terminal part of Tom40 was observed to pass through the Tom40 channel itself to recruit the chaperones in the IMS [76]. Besides Tom40, the yeast TOM complex consists of 6 α-helical proteins - Tom5, Tom6, Tom7, Tom20, Tom22 and Tom70 (Fig 3).

The mitochondrial TOM complex in animals and fungi have two primary receptors – Tom20 and Tom70 that identify proteins with NTS and ITS [50,77,78]. Cytosolic chaperones, Hsp70 and Hsp90 deliver substrate proteins in an unfolded form to Tom70 [47]. The C-terminal region of both Tom70 and Tom20 have 11 and 1 tetratricopeptide repeat (TPR) domains respectively, which interact with cytosolic chaperones and substrates, while their N-terminus has a single transmembrane helix anchored in the MOM. Tom70 is present in animals, fungi and members of Stramenopiles, Alveolata and Rhizaria (SAR, represented by *Blastocystis hominis*) supergroup (Table 1) [79]. The canonical Tom20 is present in animals and fungi (Opisthokonta) (Table 1). *Arabidopsis thaliana* (Plantae) has a Tom20 that evolved independently with a C-terminal anchor (Table 1) [80]. *Entamoeba* has a lineage-specific soluble TPR-carrying protein named, Tom60 (Table 1) [81]. The TOM complex in *T. brucei* named archaic translocase of the outer membrane (ATOM) has two lineage-specific receptors, ATOM69 and ATOM46 (Table 1). ATOM69 is a TA protein with Hsp20 chaperone-like and TPR domains while, ATOM46 is signal-anchored with armadillo repeats [82].

Tom22 is the second most conserved subunit of the TOM complex (Table 1) [6,83]. It has three domains: (i) *cis* or cytosolic domain that plays a chaperoning role during the transfer of protein from the receptor to the import pore [84–86], (ii) conserved TMD that has a tryptophan residue at the second position, a few hydroxylic residues and an invariant proline residue that forms a kink to tether two Tom40 molecules [83,86,87] and (iii) *trans* or IMS domain that binds to positively charged presequence of the translocating protein exiting the Tom40 pore and hands off to the Tim50 receptor of the TIM23 complex [88].

During the import of a presequence-carrying protein, the hydrophobic groove of Tom20 interacts with the hydrophobic side of the amphiphilic α-helix while, the acidic residues-rich cytosolic domain of Tom22 binds to the positively charged side of the amphiphilic α-helix [86,89]. Although, Tom22 and Tom20 co-operate for the simultaneous binding of the preproteins, Tom22 forms a stable association with Tom40 but, not with

Tom20 [86,90]. While, the deletion of both *TOM20* and *TOM70* genes is lethal in yeast, overexpression of *TOM22* gene can confer viability, although with a defective growth phenotype [90]. Presumably, Tom70, Tom20 and Tom22 have multiple protein-protein interaction and docking sites that assist in the effective recognition and transfer of the substrate to the import channel [91]. Tom20 can bind to both proteins with an NTS and ITS while, Tom70 prefers proteins with ITS. Tom22 selectively binds to presequence-carrying preproteins [91]. Tom5 has been shown to link the TOM receptors to the TOM channel [92] but, supportive follow up studies are so far lacking.

Table 1: Distribution of TOM complex modules in representatives of main eukaryotic supergroups. *S. cerevisiae* (Opisthokonta), *A. thaliana* (Plantae), *B. hominis* (SAR), *E. histolytica* (Archamoebae), *T. brucei* (Excavata), *T. vaginalis* (Excavata); ? = unknown, - = absent

TOM modules	*S. cerevisiae*	*A. thaliana*	*B. hominis*	*Entamoeba* sp.	*T. brucei*	*T. vaginalis*
Translocation channel	Tom40	Tom40	Tom40	Tom40	Tom40	Tom40
Central receptor	Tom22	Tom22'	?	?	ATOM14	?
Primary receptors	Tom20 and Tom70	Plant Tom20	Tom70	Tom60	ATOM69 and ATOM46	?
Small Toms	Tom5, Tom6 and Tom7	Tom5, Tom6 and Tom7	?	?	ATOM11 and ATOM12	?
Other	-	-	-	-	pATOM36	-

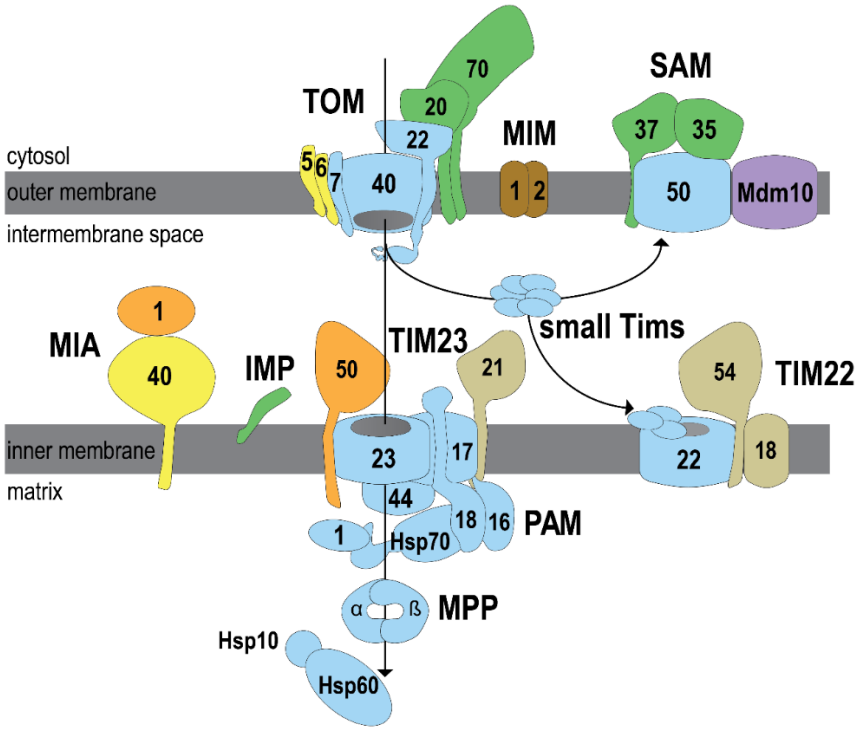


Figure 3. The mitochondrial protein import machinery in *S. cerevisiae*. Various subunits of yeast TOM, SAM, MIA, TIM and PAM subunits are colour-coded to indicate their distribution in different eukaryotes. Tom40, Tom22 and Tom7 of TOM, Sam50 of SAM, Tim17/22/23 and small Tims of TIM complexes, Tim44, Pam16, Pam18, Hsp70 and Mge1 of PAM, MPP and chaperones Hsp60-Hsp10 are present in most eukaryotes (Blue). Tom20 and Tom70 of TOM, Sam35 and Sam37 of SAM and IMP are present in metazoans (Green). Tom5 and Tom6 of TOM and Mia40 of MIA are present in metazoans and plants (Yellow). Tim50 of TIM23 complex and Erv1 of MIA are present in metazoans, plants, SAR supergroup and Kinetoplastida (Orange). Tim21 of TIM23 complex and Tim54 and Tim18 of TIM22 complex are present in metazoans (Khaki). Mdm10 is present in fungi and in certain members of amoebozoans, excavates and plants (Violet). The MIM complex are specific to fungi (Brown). Figure modified from [93].

How is the TOM complex assembled in the MOM? Tom40 is the central subunit around which the complex is built, and together with Tom22, Tom5, Tom6 and Tom7, it forms a core complex whereas, receptors Tom20 and Tom70 are loosely associated giving rise to the mature complex [76,90,94]. A new molecule of Tom40 (monomer) is assembled onto a pre-existing TOM complex corresponding to the size of the dimeric core complex implying that the core complex acts as a platform for the formation of the mature complex [76,95]. The cytosolic chaperones deliver Tom40 precursor in a partially folded conformation, in contrast to the soluble matrix proteins that are maintained in an unfolded state, and the release of the Tom40 precursor from the chaperones requires ATP [95]. The initial steps of Tom40 biogenesis involves its interaction with Tom20 and Tom22 receptors and translocation across the outer membrane through the TOM channel [52,95]. Next, the

Tom40 precursor forms a ~250 kDa intermediate with Tom5 with the precursor exposed to the IMS. Then, the Tom40 precursor forms a ~100 kDa intermediate with Tom5 and Tom6 before being assembled to a pre-existing complex to form a ~400 kDa mature complex [52,95]. Why does the assembly of Tom40 first involve a ~250 kDa intermediate and then, a ~100 kDa intermediate? Other than the Tom40 precursor and Tom5, the ~250 kDa intermediate does not seem to contain any other TOM subunit [52]. Nevertheless, it is possible that this intermediate contains SAM that is known to fold and assemble Tom40 precursor to the MOM. After its release from SAM, the Tom40 precursor may form a ~100 kDa intermediate. A single molecule of Tom22 can interact with two molecules of Tom40 via its TMD contributing to the formation of the TOM complex (Shiota et al 2011, Shiota et al 2015). Tom5, Tom6 and Tom7 are involved in the maintenance of the TOM complex [52,92,96]. In yeast, Tom6 initiates the interaction between Tom40 and Tom22 as an assembly factor but, it is not required once the interaction has been established. In the absence of Tom6, Tom40 predominantly stays in a ~100 kDa subcomplex with Tom7 and Tom5 [90].

By the late 1990s, the method to obtain intact TOM complexes from mitochondria was worked out. The first observations via electron microscopy (EM) revealed that the isolated TOM complex was composed of two or three pores, which were referred to as the core- and holo-complexes respectively [94,97]. Later, cryo-EM showed that a ~550 kDa translocase from *S. cerevisiae* is triangular shaped, with a three-fold symmetry measuring 138 Å in diameter and forms three pores [98]. Recently, a high-resolution study accounted that a ~148 kDa TOM core complex from *Neurospora crassa* has two pores, measuring 130 X 100 Å in size, forming a shallow funnel on the cytoplasmic side to allow an efficient translocation of proteins. Each Tom40 channel is surrounded by the TMDs of Tom22, Tom5, Tom6 and Tom7. The TMD of Tom22 has been demonstrated to connect two molecules of Tom40s and thus, it forms a dimer interface (Fig 4) [86,99]. It has been hypothesized that the TOM complex is a dynamic structure switching between dimeric and trimeric forms with the TMD of Tom22 tethering two molecules of Tom40s (Fig 5) [76]. Two crosslinking studies in the recent times have suggested that the pool of dimeric TOMs do not have Tom22 [76,100]. However, cryo-EM structure for dimeric TOMs demonstrated the interaction of Tom22 with two Tom40 molecules [99].

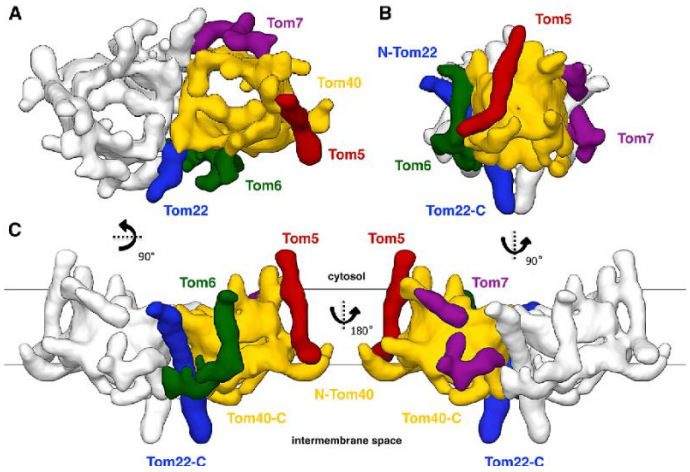


Figure 4. Structure of *N. crassa* TOM core complex. (A) The β-barrel of Tom40 is yellow, Tom22 is blue, Tom5 is red, Tom6 is green and Tom7 is purple. (B) Side view of the complex with Tom5 in front. (C) Two side views indicate the orientations of the small α-helical subunits relative to the lipid bilayer. Referenced from [99].

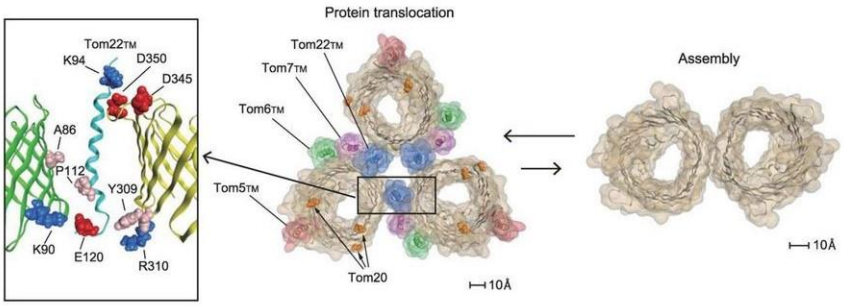


Figure 5. Subunit organization of the yeast TOM complex. Subunit arrangement of the Tom40 β-barrel and TM α-helices of Tom5, Tom6, Tom7, and Tom22. The Tom22 transmembrane α-helix, possibly bent at Pro112, tethers two Tom40 molecules. Referenced from [76].

In addition to Tom40, the MOM has mitochondrial porin Por1 that is involved in ion transport [101]. Earlier, it was thought that the TOM complex and the mitochondrial porin have entirely independent functions without interacting with each other. However, a recent study has shown that Tom22 under certain conditions dissociates from the TOM pool partnering with Por1 and then may re-associate back with TOM. Thus, Por1 modulates Tom22 integration to TOM [100]. The TOM complex is also involved in maintaining protein quality control. In yeast, the TOM complex in cooperation with the cytosolic Hsp104 imports misfolded protein aggregates into mitochondria for either refolding or proteolysis [102]. When there is a defect in the mitochondrial protein import, various quality control mechanisms are activated to restore proteostatsis. The mitochondrial compromised protein

import response (mitoCPR) recruits both ATPase Msp1 and proteasome via Cis1 to remove any unimported precursors [103]. Similarly, the mitochondrial protein translocation-associated degradation (mitoTAD) triggers the binding of Ubx2 to TOM to recruit the AAA-ATPase Cdc48 that clears the trapped precursor protein from the TOM complex [104].

B. Sorting and Assembly Machinery (SAM)

The β-barrel proteins of the MOM such as Tom40 and VDAC are guided by the small Tim complexes (Tim9-Tim10 and Tim8-Tim13) in the IMS to SAM following their translocation through the TOM pore [50,105,106]. The SAM recognizes a β-signal motif, PxGxxHxH (P - polar amino acid, x – any, G - glycine and H - hydrophobic amino acid), present in the last strand of all β-barrels [72] and facilitates the insertion of β-barrels to the lipid bilayer. The core subunit of SAM is Sam50, a β-barrel protein of the outer membrane protein 85 (Omp85) family that is conserved from bacteria to humans [6,105,106]. However, so far, both bioinformatic searches and proteomic studies have failed to identify a Sam50 homologue in the parasitic protist, *G. intestinalis* [107]. The mitochondrial SAM is homologous to the bacterial β-barrel assembly machinery (BAM complex) involved in the assembly of various OMPs [105,106]. Sam50 has an N-terminal polypeptide transport-associated (POTRA) domain present in the IMS that recognises β-barrel precursors and a C-terminal "bacterial surface antigen" domain that forms a 16 β-stranded barrel structure [6,75,106,108]. During the precursor assembly, first, the β-signal of the precursor initiates the lateral opening of Sam50 barrel. Then, Sam50 mediates the folding and lateral release of the precursor to the lipid phase. This represents a general mechanism present in mitochondria, chloroplast and Gram-negative bacteria [109]. In fungi, SAM consists of two more subunits, Sam35 and Sam37 (Fig 3) [110–112]. Sam35 assists the binding of β-barrel precursor to Sam50 while, Sam37 mediates the release of the substrate from SAM to the membrane [113]. Cryo-EM showed that SAM exists in 2 forms, one with two pores formed by 2 molecules of Sam50 and the other composed of Sam50:Sam35:Sam37 in the ratio 1:1:1 [112].

Although, SAM is primarily involved in the β-barrel biogenesis, it is also is required for the assembly of α-helical TOM proteins. The insertion of Tom22 into the TOM complex required all three subunits, Sam50, Sam37 and Sam35. However, for the insertion of small Toms – Tom5, Tom6 and Tom7, the role of Sam37 was sufficient [114]. The cytosolic domain of Tom22 interacts with Sam37 resulting in the formation of a transient TOM-SAM supercomplex in yeast [115,116]. Mitochondrial distribution and morphology protein 10

(Mdm10), a β-barrel protein found in the MOM is a dynamic interaction partner of SAM (Fig 3). Mdm10 promotes the biogenesis of some α-helical and β-barrel proteins, including Tom40 [117]. Under certain conditions, Tom7 plays an antagonistic role to dissociate Mdm10 from SAM [118]. The second role of Mdm10 is to anchor a mitochondrial component Mmm2 of the endoplasmic reticulum–mitochondria encounter structure (ERMES) to modulate the inter-organellar contact sites [119]. The human mitochondria have Metaxin 1 and Metaxin 2 proteins for β-barrel assembly, and they share limited homology with Sam37 and Sam35 respectively [120,121]. Sam50, along with the metaxins was reported to be functionally connected with the mitochondrial contact site and cristae organizing system (MICOS) proteins Mitofilin (Mic60) and CHCHD3 (Mic19) forming a Mitochondrial IMS Bridging (MIB) complex to link the outer and inner membranes and for the maintenance of cristae structure. Interestingly, Sam50 seems to be involved during the assembly of respiratory complexes I, III and IV and it plays a role in mitophagy [122,123]. These findings point out that the SAM functions are not just limited to assembling β-barrel proteins in mitochondria.

C. Mitochondrial Import (MIM) complex

Some α-helical single- and multi-span membrane proteins are not imported to mitochondria through the TOM channel. Instead, these proteins are inserted to the MOM by the MIM complex in cooperation with Tom70 receptor [124,125]. The MIM complex is specific only to fungi, composed of an oligomer of Mim1 and 1-2 units of Mim2 (Fig 3) [126]. Both Mim1 and Mim2 have an N-terminal cytosolic domain and an IMS-localized C-terminus with a single TMD passing through the MOM [111,126,127]. Mim1 promotes the insertion of two signal-anchored TOM receptors Tom20 and Tom70 to the TOM core complex [124]. A latest investigation has shown cation-sensitive channel activity for Mim1 from electrophysiological experiments making it the first protein to be reported with a channel-forming α-helical structure in the MOM [128]. The oligomerization of Mim1 is crucial for the formation of the channel to facilitate the lateral release of the proteins to the lipid bilayer. The deletion of *MIM1* gene in yeast cells leads to reduced growth, abnormal mitochondrial morphology, impaired assembly of the TOM complex and decreased biogenesis of the mitochondrial proteins [111,127]. Subsequent characterization of *MIM2* gene resulted in the same set of defects [126] indicating that both Mim1 and Mim2 function in the same pathway. A recent study showed that Mim1 and/or Mim2 could be functionally replaced by pATOM36 of *T.*

brucei and vice versa. Mim1/Mim2 and pATOM36 do not share any sequence similarity and have different topologies. This seems to be a case of convergent evolution [129].

D. Translocases of the Inner Membrane (TIM)

The protein translocases in the inner membrane are comprised of TIM22 complex and TIM23-PAM complex. The former is supported by the molecular chaperone complexes of Tim9-Tim10 and Tim8-Tim13, which sort the translocating proteins from TOM to TIM22.

i. TIM22 complex (Carrier translocase of the inner membrane)

The mitochondrial inner membrane contains a large amount of metabolite carrier proteins to transport molecules across the membrane. They have non-cleavable ITS present along the primary structure of the proteins [50]. These newly synthesized hydrophobic α-helical proteins are immediately bound by the cytosolic Hsp70 and Hsp90 to prevent aggregation and are delivered to Tom70 receptors [47,64]. The carrier proteins enter TOM in such a manner that both their N- and C-termini are toward the cytosolic side and the mid-portion passes through the channel [64,130]. Further, the precursor is chaperoned by the small TIMs in the IMS following which, the TIM22 complex mediates their insertion to the inner membrane (Fig 3) [130–132]. The TIM22 complex is composed of the channel-forming α-helical Tim22 and the receptor protein Tim54. Tim9-Tim10 along with Tim12 form a heterocomplex of Tim9-Tim10-Tim12 and is recognized by the receptor of the TIM22 complex, Tim54 at the exterior side of the complex (Fig 3 and Table 2) [133]. The carrier protein substrate is then inserted onto the TIM22 complex which, has two channels and hence, called twin-pore translocase [131]. Later, the substrate is laterally released into the lipid phase of the inner membrane. Besides Tim22 and Tim54, the TIM22 complex contains a Tim18-Sdh3 module. Tim18 is evolutionarily related to Sdh4 of the succinate dehydrogenase (SDH) complex or respiratory complex II. A metazoan-specific component, Tim29 was found to be vital for the stability of the translocase, to form a contact site between TOM and TIM22 for a more efficient transfer of hydrophobic proteins in the aqueous IMS [134]. The biogenesis of mitochondrial carrier protein is very complicated and still a topic of research. Recently, it was shown that the mitochondrial porin plays a role to facilitate the recruitment of the TIM22 complex to the substrate-TOM-small TIMs translocation intermediate during

the hand off step [135]. In human mitochondria, MICOS associates and positions the TIM22 complex at the cristae junction to promote the insertion of mitochondrial carriers [136].

ii. TIM23-PAM complex (Presequence translocase of the inner membrane and its associated motor)

The presequence translocase of the inner membrane and its associated motor imports preproteins into the matrix. The TIM23 complex is formed by four proteins: Tim50, Tim21, Tim23 and Tim17. Tim23 is the channel-forming subunit (Fig 3 and Table 2). The channel is hydrophilic, sensitive to membrane potential and presequences [137]. The N-terminal of Tim23 seems to be localized in the IMS and the C-terminal membrane-embedded region forms the channel. However, there is also some evidence that the N-terminal part can insert itself to the outer membrane to bridge with TOM when the substrates are being channelled across the membranes [138,139]. Tim17, a core subunit of TIM23 complex regulates the opening and closing of the Tim23 channel [140]. Patch clamping experiments showed that TIM23 is a twin-pore translocase and depletion of Tim17 will collapse the twin-pores to single-pore entities suggesting that Tim17 might be required to hold the two channels together [140]. Both Tim23 and Tim17 have four α-helical TMDs and contain GxxxG motif that is important for their structural integrity.

The presequence pathway first involves receiving the incoming protein from the *trans* side of Tom22 and Tom40 of the TOM complex by the Tim50 receptor [141]. Tim21 tethers TIM23 complex with TOM for an efficient passing of the translocating protein [142]. Tim50 has a TMD, a short N-terminus segment towards the matrix and a C-terminal functional receptor domain that faces the IMS [141]. The TIM23 complex is also responsible for the import and assembly of proteins of the respiratory complexes. Tim21 links the TIM23 complex to the respiratory complexes III and IV and this coupling promotes the import of newly synthesized protein, its lateral release and assembly [143]. A fungi-specific subunit, Mgr2 binds to a hydrophobic stop-transfer sorting signal and regulates the release of proteins to the lipid phase [144]. A newly identified subunit of TIM23 complex, reactive oxygen species modulator 1 (ROMO1) controls the distribution of Tim21 between TIM23 complex and respiratory complexes [145]. Recently, it was shown that the TIM23 complex is associated with MICOS in yeast [136].

The TIM23 complex recruits PAM proteins – Tim44, Pam18, Pam16, mtHsp70 and Mge1 to pull the translocating protein into the matrix (Fig 3 and Table 2) [142,146]. The

TIM23 complex is dynamic and exists in two forms: TOM tethering and PAM binding states [142]. The central subunit of PAM is a molecular chaperone, mtHsp70 that functions at the expense of ATP. It exists in two forms: a membrane-associated form that behaves as a protein import motor and a soluble form that has chaperone activity to prevent misfolding of newly imported proteins [147,148]. The N-terminal region of Tim44 interacts with the import motor and the C-terminal region interacts with the TIM23 complex [149]. Tim44 acts a docking station for mtHsp70 and their dynamic interaction is modulated by the presence of ADP/ATP [150]. Mge1, a nucleotide-exchange factor and a homologue of bacterial GrpE removes ADP molecule from mtHsp70 and thus, recycles mtHsp70 for a new round of ATP hydrolysis [50,151]. Pam18 promotes ATPase activity of mtHsp70 while, Pam16 works in an antagonistic manner to control the functioning of Pam18. Both Pam18 and Pam16 are co-chaperones and have J-domains [152,153]. Besides ATP being a driving force, membrane potential plays a significant role in the translocation of proteins across the inner membrane. The negative potential on the matrix side creates an electrophoretic effect on the positively charged NTS. The membrane potential also initiates the voltage-dependant activation of the Tim23 channel [137]. For an efficient import of mitochondrial proteins, a combination of two mechanisms – passive trapping (Brownian ratchet model) for loosely folded proteins and active pulling (pulling model) for tightly folded proteins are necessary [51,154].

iii. A general TIM complex

The canonical mitochondria have two TIM complexes, TIM22 and TIM23 for the import of inner membrane and matrix proteins. The channel-forming components of both complexes belong to Tim17 protein family: Tim17 and Tim23 (TIM23), and Tim22 (TIM22). However, certain protists have only one type of protein from Tim17 family and a general TIM complex that seems to import different classes of proteins. *T. brucei* (Tb) has a ~ 1,100 kDa TIM that can import both proteins with an NTS as well as membrane proteins [155,156]. TbTIM contains a protein encoded by a single gene of Tim17 family that was probably derived from Tim22 (Table 2), Tim50 and five novel proteins namely, TbTim47, TbTim54, TbTim62, TimRhom I and TimRhom II [155–157]. Both *G. intestinalis* and *C. parvum* mitosomes have a TIM based on a single type protein similar to Tim17 (Table 2) [157,158]. The mitochondrion-like organelle in *Paratrimastix pyriformis* has a member of Tim17 protein family [159]. *T. vaginalis* has 5 paralogues of Tim17 protein family (Table 2) [44]. The MRO of *Blastocystis* species contains a Tim17-type protein, Tim50 and Tim21 [32]. The

microsporidian pathogen *Encephalitozoon cuniculi* has a Tim17 protein related to Tim22 (Table 2) [160]. So far, not a single TIM subunit member has been identified in *E. histolytica* (Table 2) [161].

E. Mitochondrial Intermembrane space Assembly machinery (MIA)

Proteins that localize to the mitochondrial IMS proteins, namely small TIM chaperones, cytochrome c oxidase etc. contain conserved cysteine motifs that are oxidized to disulfide bridges by the redox-regulated import receptor Mia40 [63,162]. This process also induces folding of the newly synthesized protein in the IMS and the mechanism is thus known as oxidative protein folding [163]. During the formation of disulfide bridges, the electrons are transferred from the reduced substrate to Mia40 that is re-oxidized by sulfhydryl oxidase, Erv1 (Fig 3). Further, Erv1 gets re-oxidized by the transfer of electrons to cytochrome c oxidase of the respiratory chain, or directly to molecular oxygen. Mia40 is essential in animals and fungi while, in plants, Mia40 is dispensable as Erv1 can perform the same task. Because plant Erv1 was able to compensate the deletion of yeast Mia40 and some protists have only Erv1 but not Mia40, it seems that Mia40 was added to the sulfide-relay system for better substrate specificity [164,165]. When the oxidative folding is defective, the substrates are retrotranslocated through TOM back to the cytosol and degraded by the proteasome machinery. This pathway is part of a quality control surveillance system [166].

Table 2: Distribution of TIM and PAM subunits in organisms of various eukaryotic supergroups. *Homo sapiens* (Metazoa), *S. cerevisiae* (Fungi), *E. cuniculi* (Fungi), *B. hominis* (SAR), *T. vaginalis* (Parabasalia, Excavata), *G. intestinalis* (Diplomonadida, Excavata), *T. brucei* (Kinetoplastida, Excavata), *E. histolytica* (Archamoebae); Tim22* = Tim17 derived from Tim22, ? = unknown, + = present, - = absent.

Modules	*H. sapiens*	*S. cerevisiae*	*E. cuniculi*	*B. hominis*	*T. vaginalis*	*G. intestinalis*	*T. brucei*	*E. histolytica*
TIM23 complex								
Translocation channel	Tim23, Tim17	Tim23, Tim17	Tim22*	Tim17	Tim17/23?	Tim17	Tim22*	?
Receptor	Tim50	Tim50	Tim50	Tim50	?	?	Tim50	?
ROMO1	+	+	?	?	?	?	?	?
Tim21	+	+	+	+	-	-	-	-
PAM								
Tim44, Pam16, Pam18 and mtHsp70	+	+	+	+	+	+	+	?
Mge1	+	+	+	+	+	+	+	?
MPP	α/β	α/β	?	α/β	α/β HPP	β	α/β	?
TIM22 complex								
Translocation channel	Tim22	Tim22	Tim22*	-	Tim22?	-	Tim22*	?
Receptor	Tim54	Tim54	Tim54	?	?	?	-	?
Small Tims: Tim9-10-12, Tim8-13	+	+	+	+	+	?	+	?
Tim29	+	-	-	-	-	-	-	-
Tim18	+	+	+	-	-	-	-	-

F. Mitochondrial peptidases and molecular chaperones

When the presequence of the translocating preprotein reaches the matrix, it is cleaved off by MPP (Fig 3) [167]. MPP is a heterodimer composed of two subunits, the larger α-subunit with a glycine-rich loop for substrate recognition and the smaller β-subunit for catalytic activity. MPP evolved from a pre-existing bacterial protease and during the mitochondrial evolution, it was initially integrated into cytochrome bc1 complex of the respiratory chain. This status quo has been maintained in plants while, in fungi and mammals, the two MPP subunits separated from cytochrome bc1 complex allowing independent regulation of protein processing and respiration [167,168]. Interestingly, MPP in *G. intestinalis* mitosomes consists of only βMPP subunit with a catalytic activity [57]. Both α/βMPP of canonical mitochondria and Giardia βMPP evolved from the same ancestor and the latter seems to have followed a divergent reductive path [57]. The characteristics of Giardia enzyme are reflected in the presequences found in mitosomal preproteins which are short and deprived of positive charges [57]. Trichomonad hydrogenosomes have a canonical α/βMPP that is similar to those found in most eukaryotes [57]. The N-terminal presequences in most eukaryotes that are recognised by MPP have three features: overall positive charge, ability to form amphiphilic α-helix and an arginine residue at - 2 position (R-2 rule) from the cleavage site [167]. The N-terminal part of βMPP forms a conserved negatively charged amphiphilic α-helix structure to interact with the positively charged amphiphilic α-helix formed by presequence of the translocating preproteins [169]. Once inside in the matrix, the proteins are folded by Hsp60-Hsp10 molecular chaperones (Fig 3). Some IMS proteins like cytochromes b1 and c2 carry two cleavable presequences. First, they are translocated via TOM and TIM23, and the NTS is cleaved off by the MPP which exposes a "stop-transfer signal" resulting in their retrotranslocation to the IMS where, IMP cleaves off the second segment [62].

1.4.4 Evolution of mitochondrial protein translocases

What were the constituents of the protein translocases in proto-mitochondria of the last common eukaryotic ancestor (LECA)? It has been speculated that protomitochondria had a primitive set of protein translocases [6]. Most of the protein import pathways are common to all forms of mitochondria. Heterologous expression of hydrogenosomal and mitosomal proteins have resulted in their localization to mitochondria and vice versa implying that the overall mechanism is conserved and the targeting signals can be recognised by the protein import machinery of different eukaryotic lineages [44,59,61,161].

The proto-TOM complex seemed to be composed of at least three conserved subunits Tom40, Tom22 and Tom7 (Fig 6) [6,83,170]. Tom40 is conserved across all eukaryotic lineages that have mitochondria [171]. Because of the endosymbiotic origin of mitochondria, it seems likely that Tom40 and other related proteins such as VDAC/porin and Mdm10 were derived from a common bacterial protein. A primitive Tom40 structural model reconstructed based on the phylogenetic profiles of Tom40 from different organisms displayed the presence of both acidic and hydrophobic amino acid residues inside the channel to facilitate the translocation of both preproteins with presequence and hydrophobic proteins respectively. Tom22 functions as a receptor and also tethers two Tom40 subunits to form a trimeric TOM complex [76,84]. Hence, it was speculated that the proto-TOM complex could have been trimeric in nature (Fukawasa et al 2017). The N-terminus of Tom22 in plants, *Toxoplasma gondii* and *T. brucei* (ATOM14) is shorter compared to that of opisthokont Tom22 and does not seem to bind to presequences [83,170,172,173]. Thus, the elongation of N-terminus of Tom22 by the addition of an acidic cluster might have coincided with the gain of Tom20 in opisthokonts for the cooperative recognition of presequences [170]. Various peripheral lineage-specific receptors were most likely added to the evolving TOM complex after the divergence of proto-mitochondria. Both Sam50 and its bacterial homologue, BamA form a 16 β-stranded structure and function via a similar lateral gate opening mechanism to assemble β-barrel proteins [109,174]. However, Sam50 has a single N-terminal POTRA domain while, BamA, has five POTRA domains. Homologues of four other subunits of the BAM complex, BamB, BamC, BamD and BamE seem to be absent in eukaryotes [6]. Sam50 was probably derived from BamA, and four POTRA domains as well as four other BAM subunits were lost during the changeover.

The core subunits of TIM complexes, Tim17, Tim22 and Tim23 belong to the same family of proteins and most likely evolved by gene duplication from a common ancestral protein. All three proteins show sequence similarity to the bacterial amino acid transporter LivH and outer envelope protein (OEP) 16 of chloroplasts [175]. Because of their sequence similarity, these proteins are grouped under preprotein and amino acid transporters (PRAT) family. Zarsky and Dolezal suggested that the ancestral mitochondria had all 3 proteins – Tim17, Tim22 and Tim23, and some lineages retained one protein and lost the other two through reductive evolution [157]. Tim50 is present in many metazoans, plants, SAR supergroup and Kinetoplastida (Fig 6). Tim44 and Pam18 seem to have evolved from two bacterial proteins named TimA and TimB that share similar structural features respectively [176].

Taking into consideration both experimental and bioinformatic data, it can be speculated that the proto-mitochondria of the LECA most likely had Tom40, Tom22 and Tom7 constituting a primitive TOM to import both presequence-carrying and hydrophobic proteins, Sam50 to assemble outer membrane β-barrels, small Tims as chaperones in the IMS and a TIM-PAM complex composed of a channel-forming Tim17/22/23, Tim50, Tim44, Pam18, mtHsp70, Mge1 mediating the import of matrix proteins and insertion of inner membrane proteins (Fig 6).

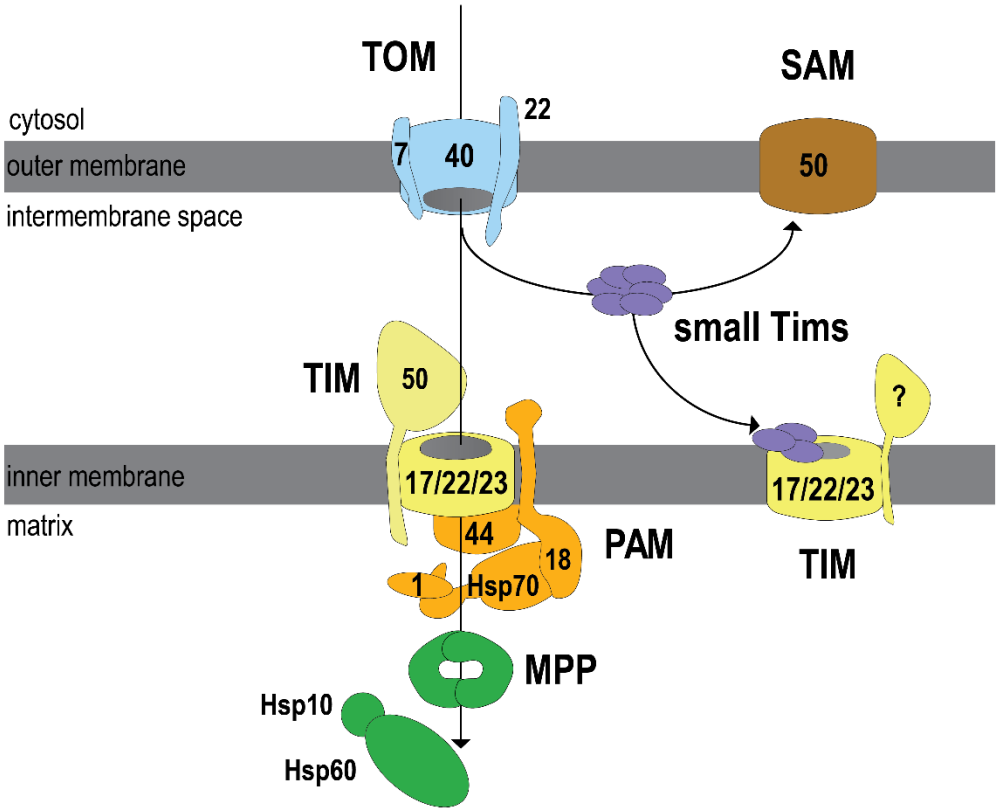


Figure 6. Scheme for the proto-mitochondrial protein import machinery in the LECA. Subunits of TOM, SAM, TIM and PAM subunits are colour-coded. Tom40, Tom22 and Tom7 (Blue) of TOM, Sam50 (Brown) of SAM, small Tims (Violet), Tim17/22/23 and Tim50 (Yellow) of TIM complex, Tim44, Pam18, Hsp70 and Mge1 (Orange) of PAM, MPP and chaperones Hsp60-Hsp10 (Green) were predicted to have been present in proto-mitochondria of the LECA. ? = unknown.

1.4.5 Targeting of tail-anchored (TA) proteins

TA proteins are a distinct group of integral membrane proteins anchored by a single TMD near their C-terminus. TA proteins perform diverse essential functions such as import of proteins into organelles, division of organelles, cellular signaling, apoptosis, enzymatic activity, vesicular trafficking etc. Some of the notable members of this family include

cytochrome b5, Bcl-2, Bax, Tom22, small Toms - Tom5, Tom6 and Tom7, Sec61β, Sec61γ, Pex15, Pex26, SNARE proteins etc. TA proteins are found in mitochondria, endoplasmic reticulum (ER), peroxisomes, chloroplasts of plants and in bacteria. How the cellular machinery distinguishes targeting signals and sorts the precursor proteins to the right organelle in eukaryotes is not fully understood. While many aspects of TA protein targeting to mitochondria remain unknown, the Guided Entry of TA proteins (GET) and TMD Recognition Complex (TRC) pathways that deliver proteins to the ER in yeast and humans respectively have been studied in great detail [177–179].

Most mitochondrial proteins cross the MOM via the TOM pore. However, TA proteins are targeted to and assembled in the MOM independent of TOM channel [180]. The TA proteins are transported to their destination post-translationally as their targeting signals are usually found in the C-terminal region. The ER seems to be the default destination for TA proteins in eukaryotes as mitochondrial TA proteins lacking specific targeting signals are localized to the ER [181]. The TA proteins of mitochondria have an N-terminal functional domain that faces the cytosol, a transmembrane α-helix and a C-terminal segment (CTS) containing basic amino acid residues that localizes in the IMS. Mitochondrial TA proteins have a moderately hydrophobic TMD that is flanked by positively charged residues [68,182]. Changes by deleting the TMD, increasing the TMD length, lowering the net positive charge of the CTS, inserting a linker between the TMD and the CTS have impaired the localization of mitochondrial TA proteins [68]. Following their release from the exit tunnel of the cytosolic ribosomes, the hydrophobic segments are masked by the cytosolic molecular chaperones to avoid aggregation [180]. In yeast, Ssa chaperones of the Hsp70 family, Sti1 co-chaperone and peroxisomal protein import factor, Pex19 were found to be involved in the biogenesis of two mitochondrial TA proteins, Fis1 and Gem1 [183]. When the cell synthesizes a high amount of mitochondrial TA proteins, the competition between the molecular chaperones delivering proteins to mitochondria and the factors from the ER-associated GET pathway decides the outcome [129]. Usually, this results in the mistargeting of a portion of mitochondrial TA proteins to the ER. So, in principle, an active mitochondrial targeting system must be present to deter other pathways from capturing the substrates [129].

When heterologously expressed in yeast cells, two *E. coli* TA proteins, ElaB and YqjD localized to mitochondria, and a small fraction were found in the ER. This shows that the targeting machinery for TA proteins is conserved from bacteria to eukaryotes and it was suggested that the TA proteins acquired through bacteria could have driven the protein targeting in eukaryotes [184]. An often-overlooked factor is the difference in the lipid

composition of the mitochondrial outer and the ER membranes as lipids have been shown to regulate the TA protein insertion as well [185].

1.5 *Trichomonas vaginalis* and hydrogenosomes

T. vaginalis is an anaerobic (microaerophilic), unicellular, parasitic protist that belongs to the supergroup Excavata. This parasite is responsible for a sexually transmitted infection called Trichomoniasis in humans with over 270 million cases every year worldwide [186]. *T. vaginalis* has a large ~160 Mb genome with close to 60,000 genes owing to a high number of gene duplication [187]. *T. vaginalis* has hydrogenosomes that metabolize mostly pyruvate and/or malate to carbon dioxide, acetate and hydrogen with concomitant synthesis of ATP by substrate-level phosphorylation. They are devoid of genome, protein synthesis machinery, membrane-bound respiratory chain complexes, cristae in the inner membrane and TCA cycle. However, they play vital roles in iron-sulfur cluster assembly, amino acid metabolism and detoxification etc. As per the proteomic data, the organelle has around 600 proteins with 70 of them predicted to be in the outer and inner membranes [44,45].

1.6 Protein import into *T. vaginalis* hydrogenosomes

The mechanism of protein import into hydrogenosomes is poorly understood and limited experimental data available are exclusively on the hydrogenosomes of *T. vaginalis*. Despite the presence of a common mode of protein import [14], there are remarkable differences between mitochondria and hydrogenosomes. The NTS in the hydrogenosomal proteins are considerably shorter (10-20 amino acid residues) than the mitochondrial NTS (10-80 amino acid residues); and the shortest hydrogenosomal NTS (5 amino acid residues) as in the case of a matrix protein pyruvate:ferredoxin oxidoreductase (PFO), cannot form an amphiphilic α-helix [57,188,189]. Moreover, the hydrogenosomal NTSs have significantly lower positive charge in comparison with the mitochondrial NTS [57]. It has been suggested that this difference reflects low (if any) hydrogenosomal inner membrane potential due to lack of the respiratory chain complexes [7,14]. Moreover, recent studies have shown that the NTS is not essential for the translocation of several hydrogenosomal matrix proteins at all [58,59].

Further, the hydrogenosomal proteome had revealed only a few components of protein import machinery [44]. Concerning the TOM complex, six isoforms of Tom40-like proteins were identified. However, it was difficult to distinguish between putative Tom40 and

VDAC that have similar structure and belong to the mitochondrial porin superfamily. Also, none of the prototypical mitochondrial receptors (Tom20 and Tom70), and small Toms (Tom5, Tom6 and Tom7) were found in the proteome [44]. Regarding SAM, only the core subunit Sam50 was identified. The hydrogenosomal proteome also showed the presence of two proteins with homology to Tim9-Tim10 that are possibly located in the IMS and on the other hand, there are no genes encoding for Mia40 and Erv1 proteins [44]. In the inner membrane, the presence of putative TIM was proposed, which is supported by the finding of five divergent homologues of Tim17/22/23 family proteins (Rada et al 2011). However, it is still unknown whether these proteins form a single general translocase as observed in other protists [155,158] or distinct TIM23 and TIM22 complexes as in animals and fungi. Although, it is difficult to distinguish between Tim17 family proteins, it has been suggested that trichomonads have a Tim23-type translocase based on phylogenetic analysis [44,157]. The only complete part of import machinery seems to be PAM, that is composed of Pam16, Pam18, Tim44 and mtHsp70 [44]. Trichomonad hydrogenosomes have α/β heteromeric HPP [57]. The absence of certain components in the protein import machinery is likely to be a result of reductive evolution during adaptation of trichomonads to anaerobic environment and/or parasitic lifestyle. However, it is also possible that some components might be highly divergent and were not been identified by conventional homology searches.

2. AIMS AND OBJECTIVES:

1. To identify and characterize the components of the TOM complex of hydrogenosomes in *T. vaginalis*

2. To investigate the interaction between the hydrogenosomal proteins (substrates) and the subunits of TvTOM complex

3. To elucidate the structure of TvTOM complex

4. To understand the evolution of TOM complex

5. To examine the NTS-independent transport of proteins to *T. vaginalis* hydrogenosomes

6. To study the targeting of TA proteins in *T. vaginalis*

3. LIST OF PUBLICATIONS AND CONTRIBUTIONS

3.1 Makki A, Rada P, Žárský V, Kereïche S, Kováčik L, Novotný M, Jores T, Rapaport D, Tachezy J (2019) **Triplet-pore structure of a highly divergent TOM complex of hydrogenosomes in *Trichomonas vaginalis***. PLOS Biol 17(1):e3000098.
Contributions: Design of the project and experiments, cloning and preparation of Trichomonas and yeast strains, immunofluorescence microscopy, subcellular fractionation, protease protection assay, alkaline carbonate extraction, blue native-PAGE, complementation study in yeast, immunoprecipitations, mass spectrometry (MS) data analyses and curation of MS data in PRIDE, bioinformatics, isolation of TvTOM complex and sample preparation for electron microscopy, cloning and synthesis of radiolabeled precursors, *in vitro* protein import experiments, expression and purification of His-tagged Trichomonas proteins from *E. coli*, binding assay for TvTOM receptor candidates, data analyses and interpretation, manuscript preparation and submission

3.2 Rada P, **Makki A**, Zimorski V, Garg S, Hampl V, Hrdý I, Gould SB, Tachezy J (2015) **N-terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis***. *J Eukaryot Cell* 14(12):1264–1275.
Contributions: Cloning and preparation of Trichomonas strains, immunofluorescence microscopy, subcellular fractionation, protease protection assay, figures for the manuscript

3.3 Rada P, **Makki A**, Žárský V, Tachezy J (2019) **Targeting of tail-anchored proteins to *Trichomonas vaginalis* hydrogenosomes**. *Mol Microbiol* 111(3):588–603.
Contributions: Cloning and preparation of Trichomonas strains, immunofluorescence microscopy, subcellular fractionation, protease protection assay, figures for the manuscript

3.4 Dolezal P, **Makki A**, Dyall SD (2019) **Protein import into hydrogenosomes and mitosomes. Hydrogenosomes and Mitosomes: Mitochondria of Anaerobic Eukaryotes**, ed Tachezy J (Springer International Publishing, Cham), pp 31–84.
Contribution: Co-author of the book chapter

4. RESULTS AND CONCLUSIONS:

4.1 The TOM complex of *T. vaginalis* hydrogenosomes (Makki et al. 2019)

The mitochondrial protein translocases are central to the organellar functioning and biogenesis, and they played a key role during mitochondrial evolution. Most of our experimental knowledge on TOM complex is limited to a few model organisms and additionally, the generic structure of TOM was still unresolved since a TOM with three pores was reported in yeast and a TOM with two pores in *N. crassa* [98,99]. Although, *T. vaginalis* hydrogenosomes were shown to have a mitochondrial-type protein translocase, the proteome showed that other than the putative Tom40-like proteins, none of the known TOM components was present (Bradley et al 1997, Rada et al 2011). In this project, different subunits of TvTOM were identified and functionally characterized, and the structure of the translocase was elucidated.

4.1.1 Tom40-like proteins

Among the seven putative Tom40-like proteins in *T. vaginalis*, TvTom40-2 was found to be the most conserved as per bioinformatic searches and modelling [190]. TvTom40-2 forms a typical 19 β-stranded structure but, with only one N-terminal α-helix while, *N. crassa* Tom40 has two α-helices. TvTom40-2 has both negatively and positively charged amino acid residues inside the barrel to provide distinct paths for the translocation of preproteins with presequences and hydrophobic proteins similar to yeast Tom40 [76,190]. TvTom40-2 was found to be present in the hydrogenosomal outer membrane in high molecular weight complexes of ~570 and ~330 kDa [190].

4.1.2 TvTom40-2 could partially complement yeast Tom40

TvTom40-2 and ScTom40 share low sequence similarity. Nevertheless, heterologous expression of TvTom40-2 in yeast resulted in its localization in the MOM and surprisingly, it had the same topology as in hydrogenosomes [190]. This supports two studies that demonstrated that the targeting and assembly signals for mitochondrial porins are conserved in their β-hairpin motifs and β-motif in the last β-strand respectively [71,72]. Further, TvTom40-2 could partially complement the function of yeast Tom40 suggesting that the basic characteristics of Tom40 channel in different eukaryotes are similar for facilitating the transport of proteins and all Tom40s most likely evolved from a common ancestral protein [190].

4.1.3 Components of the TvTOM complex

The composition of TvTOM complex was investigated using co-immunoprecipitations (coIP) coupled with proteomic and bioinformatics analyses. Using TvTom40-2 as bait, its interacting partners - other isoforms of TvTom40s, three TA proteins - Tom36, Tom46 and Homp19, Sam50 and its paralogue, Sam50p were identified via mass spectrometry (MS) [190]. Even though Tom22 is conserved in diverse organisms [6,83,173], this method could not find a homologue in Trichomonas. However, a more sensitive hidden Markov model-based bioinformatic approach identified a Tom22-like protein. It has a short cytosolic *cis* domain and a conserved Tom22 transmembrane segment but, lacks an IMS-localized *trans* domain (Makki et al 2019).

In total, all four TA proteins including Tom22-like protein interact with TvTom40-2. Two of them, Tom36 and Tom22-like protein are present in high molecular weight complexes of ~570 and ~330 kDa whereas, the other two TA proteins, Tom46 and Homp19 are present only in ~330 kDa complex in the hydrogenosomal outer membrane similar to TvTom40-2 [190]. Tom36 and Tom46 are paralogues and have an N-terminal Hsp20 chaperone-like domain followed by three TPRs. TOM receptors namely, Tom70, ATOM69, Tom60, Tom20 of other eukaryotes carry TPR domains that are involved in protein-protein interactions and moreover, the domain architecture of Tom36 and Tom46 resembles that of ATOM69 of *T. brucei* [81,82,190]. Hence, we reasoned that Tom36 and Tom46 could function as TOM receptors in trichomonad hydrogenosomes. Reciprocal coIP showed that Tom36 interacts with multiple isoforms of TvTom40 including TvTom40-2, Tom46, Homp19 and Sam50. The fourth TA protein, Homp19 on the other hand, does not have any homologue. Interestingly, TvTom40-2 interacts with other isoforms of TvTom40s indicating that either two or three different isoforms of Tom40 can be present in a single TvTOM complex [190].

4.1.4 Interaction between TvTOM subunits and hydrogenosomal protein substrates

Out of eight putative proteins of mitochondrial porin family, six were classified as TvTom40-1 to -6 and the other two as porins based on cluster analysis [44]. However, it was difficult to distinguish Tom40-like proteins from porins relying solely on bioinformatics. Hence, the function of TvTom40-2 was verified via *in vitro* import assays and coIP. The translocation of a hydrogenosomal ferredoxin-dihydrofolate reductase (DHFR) fusion protein used as a substrate was arrested at the hydrogenosomal protein import site via methotrexate-induced folding of the DHFR region. The translocation-arrested ferredoxin-DHFR substrate was co-

purified with TvTom40-2 demonstrating that TvTom40-2 mediates protein import into hydrogenosomes. This also confirmed that similar to other forms of mitochondria, hydrogenosomes import soluble preproteins in an unfolded or loosely folded manner [107,156,190,191].

To test whether Tom36 and Tom46 can bind to hydrogenosomal preproteins, binding assays were employed. Recombinant cytosolic domains of both Tom36 and Tom46, which have Hsp20 chaperone-like and TPR domains, were immobilized on resins and incubated with various proteins. Both the receptor candidates could bind to two hydrogenosomal preproteins – frataxin and α-subunit of succinyl coA synthetase but, not cytosolic cytochrome b_5. Based on overall data, Tom36 and Tom46 most likely function as TvTOM receptors [190].

4.1.5 TvTOM and Sam50 form a stable supercomplex

Reciprocal coIPs show that Sam50 is tightly associated with different isoforms of TvTom40 including TvTom40-2 and Tom36, and Sam50 is present in a ~570 kDa complex similar to TvTom40-2 implying that TvTOM and Sam50 form a stable supercomplex [190]. Sam50 is the only known subunit of SAM in *T. vaginalis* hydrogenosomes. In yeast, Sam37 interacts with Tom22 to form a transient TOM-SAM supercomplex [115,116]. Since, Sam37 is absent in trichomonads, a different mode of interaction can be expected between TvTOM and Sam50 [190].

4.1.6. Skull-shaped TvTOM complex has a triplet-pore structure

Electron microscopic visualization of the isolated TvTOM complex revealed three types of particles – TOM with one, two and three pores [190]. The triplet-pore structures represent the mature holo complex and the particles with two pores represent the TOM core complex. While some features of TvTOM are similar to TOMs in fungi [97,99], the dimensions of TvTOMs with one and three pores (70 X 125 Å and 150 X 175 Å respectively) are quite different because of the presence of an extra component outside the TvTom40 barrel that seems to form a stable association and thus, giving a skull shape to the triplet-pore TvTOM structure [190]. Based on this study and the previous reports on TOMs in fungi, the triplet-pore structure seems to be generic for all mitochondrial TOMs.

4.1.7 Evolution of mitochondrial TOMs

Due to their omnipresence in eukaryotes, Tom40, Tom22 and Tom7 were proposed to be the constituents of the earliest TOM [6,83,171]. For some time, Tom22 and Tom7 were not

identified in many excavates providing no support to the proposition. However, the identification of Tom22 homologues in *T. brucei*, *T. vaginalis*, *Carpediemonas membranifera*, *N. gruberi*, *Euglena gracilis* and *Stygiella incarcarata* and Tom7 homologues in *C. membranifera*, *S. incarcerata, N. gruberi*, *N. fowleri* and *E. gracilis* changed the status quo [170,173,190]. Absence of Tom7 in *T. vaginalis* suggests that Tom7 might have been secondarily lost, or an improved method is required to identify proteins encoded by small open reading frames. Collectively, the proto-TOM complex present in the LECA was most likely trimeric in nature [190] and consisted of at least three components Tom40, Tom22 and Tom7. The TOM receptors are mostly lineage-specific and were gained later in evolution. Although, Tom36 and Tom46 resemble ATOM69 in domain features, they seemed to have evolved independently [190].

4.2 NTS-independent import of phosphofructokinase into *T. vaginalis* hydrogenosomes (Rada et al. 2015)

In this project, the import of TvATP-dependent phosphofructokinase (PFK) into *T. vaginalis* hydrogenosomes was examined. The parasite expresses both PPi- and ATP-dependent enzymes, wherein the former is present in the cytosol while, the latter is compartmentalized in the hydrogenosomes. The PPi-dependent enzyme activity is about 50-fold higher than the ATP-PFK, rendering the metabolic significance of the latter unclear [192]. Our previous study on the hydrogenosomal proteome had shown the presence of a few isoforms of ATP-PFK which is rather unusual considering that PFK is a glycolytic enzyme and none of the enzymes involved in the upstream or downstream reactions are present in hydrogenosomes [44]. What is even more interesting is that ATP-PFK1 does not carry a predictable NTS and is still capable of being imported into hydrogenosomes. An N-terminal 16 residue truncated version could localize to hydrogenosomes as well. These data suggest that it might carry an unknown ITS. Further, when heterologously expressed in yeast, TvATP-PFK localized to mitochondria [192]. To understand more, *S. cerevisiae* and *E. coli* ATP-PFKs were expressed in trichomonads and their localization was analysed. The yeast ATP-PFK (ScPFK) is cytosolic and has three regions: an N-terminal segment, a catalytic domain that is homologous to *T. vaginalis* ATP-PFK1 and a C-terminal regulatory domain. Both full-length ScPFK and a mutant without the regulatory domain localized in the cytosol but, was observed to be associated with the hydrogenosomal surface [192]. When the catalytic domain of ScPFK alone was expressed, it localized to hydrogenosomes. The *E. coli* ATP-PFK localized to hydrogenosomes as well. These results imply that both the catalytic domain of ScPFK,

which is similar to TvATP-PFK1, and *E. coli* ATP-PFK have some targeting signal or they carry an ancient feature that facilitates their import to hydrogenosomes and perhaps, the features in the N-terminal segment of ScPFK prevents its transport to mitochondria or in this case, hydrogenosomes [192]. Despite the deletion of NTS from some mitochondrial and hydrogenosomal matrix proteins, their localization to the organelles was not affected [58]. However, TvATP-PFK is the first case of a soluble protein that localizes to hydrogenosomes (mitochondria) without a predictable NTS [192].

In mitochondria, the positive charge of the NTS contributes to the membrane potential ($\Delta\psi$)-driven translocation of preproteins across the inner membrane [193] and this $\Delta\psi$ is generated by the respiratory chain complexes. However, the loss of respiratory chain in hydrogenosomes has most likely made the positive charge of NTS expendable. Most hydrogenosomal NTSs are shorter with only one positively charged residue [57], and in many cases, they are either not essential for preprotein import [58,59] or simply absent [192]. Thus, the import of these proteins is based on recognition of poorly understood ITSs. Perhaps, such changes in the targeting signals are likely reflected by the presence of lineage-specific TvTOM receptors, the absence of an elongated *cis* and acidic *trans* domains in Tom22-like protein and the divergence of TIM machinery [44,190].

4.3 Targeting of TA proteins (Rada et al. 2018)

In this project, the targeting of TA proteins in *T. vaginalis* was investigated. Trichomonas has two primary locations where the TA proteins can be localized: hydrogenosomal outer membrane and ER membrane. They carry targeting signals in the C-terminal TMD and its flanking regions. TA proteins are imported to mitochondria independent of TOM complex [68,180]. The proteome of *T. vaginalis* hydrogenosomes had shown the presence of 12 unique putative TA proteins [44]. First, the localization and topology of some candidates was confirmed using biochemical and microscopy experiments [194]. Taking these proteins as templates, the features that are essential for targeting were defined, which showed the presence of a transmembrane α-helix of 18–22 residues, a C-terminal mean hydrophobicity of 1.95 (range 1.5–2.64), a short CTS of 3–16 residues, mean net positive charges of 0.82 (range -2 to 4) and 3.27 (range 1–5) for N-terminal and C-terminal TMD flanking regions respectively and the presence of a lysine-arginine motif in the CTS such as R/K-K/R or KRRK or RKKK etc [44,194]. This approach identified 120 putative hydrogenosomal TA proteins. When the TMD was extended with 3-7 valine (hydrophobic) residues, the targeting of Tom5 mutants to mitochondria was reduced gradually but, never abolished and led to their

mis-localization to the ER [68]. When 9 valine residues were introduced in the TMD of *T. vaginalis* TA4, the mutant TA4 did not localize to hydrogenosomes but, instead remained in the cytosol [194]. Removal of the TMD and the IMS-localized CTS of TA4 also resulted in its localization in the cytosol. However, both removal of the CTS and replacement of five lysine residues in the CTS to serine residues caused a dual localization of the mutants to both hydrogenosomes and ER [194]. Trichomonads possess flattened ER sacs around the nucleus and interestingly, in the TA4 mutant without the CTS, the ER was dissipated into multiple vesicles [194]. Further, to understand the dynamics of TA protein targeting to hydrogenosomes and ER, domain swapping experiments were performed. A chimeric protein carrying the soluble part of hydrogenosomal TA4 fused to the C-terminal region of ER TA-protein disulfide isomerase (TA-PDI) (flanking region 1-TMD-flanking region 2) localized to the ER. When the TMD of TA4 in this fused protein was replaced with the TMD of TA-PDI, it had no effect and was still targeted to the ER suggesting that the charges on the flanking regions of the TMD played a dominant role in the targeting [194]. The targeting of mitochondrial and hydrogenosomal TA proteins share some properties: (i) the C-terminal TMD is of a defined length, (ii) the TMD is flanked by basic residues at the N-terminus, C-terminus or both and (iii) the TMD domain is of moderate hydrophobicity. Despite these, there are notable differences in hydrogenosomal TA protein targeting: (i) the TMD is longer, (ii) the net positive charge on the CTS is higher and (iii) the difference in the net positive charge of CTS between hydrogenosomal and ER proteins is higher. In fungi and mammals, the net positive charge of CTS of TA proteins is of the decreasing order: peroxisomes followed by mitochondria and then, ER. Since, peroxisomes seem to be absent in Trichomonas, the hydrogenosomal proteins have a higher range of net positive charge in the CTS [194].

Adaptation to operate under anaerobic conditions has resulted in an enormous reduction of both mitochondrial functions as well as proteome of *T. vaginalis* hydrogenosomes [44,45,187]. The following publications point out that these adaptations, particularly the loss of respiratory chain complexes that led to a low membrane potential, or its absence were seminal for the shaping of protein import into *T. vaginalis* hydrogenosomes.

5. REFERENCES

1. Ernster L, Schatz G. Mitochondria: a historical review. J Cell Biol. 1981;91: 227s LP-255s. doi:10.1083/jcb.91.3.227s
2. Sagan L. On the origin of mitosing cells. J Theor Biol. 1967;14: 225-IN6. doi:https://doi.org/10.1016/0022-5193(67)90079-3
3. Martin W, Müller M. The hydrogen hypothesis for the first eukaryote. Nature. 1998;392: 37–41. doi:10.1038/32096
4. Timmis JN, Ayliff MA, Huang CY, Martin W. Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 2004;5: 123–135. doi:10.1038/nrg1271
5. Dyall SD, Brown MT, Johnson PJ. Ancient invasions: From endosymbionts to organelles. Science (80-). 2004;304: 253–257. doi:10.1126/science.1094884
6. Dolezal P, Likic V, Tachezy J, Lithgow T. Evolution of the molecular machines for protein import into mitochondria. Science (80-). 2006;313: 314–318. doi:10.1126/science.1127895
7. Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, et al. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol Mol Biol Rev. American Society for Microbiology; 2012;76: 444–495. doi:10.1128/MMBR.05024-11
8. Lindmark DG, Müller M. Hydrogenosome, a cytoplasmic organelle of the anaerobic flagellate *Tritrichomonas foetus*, and its role in pyruvate metabolism. J Biol Chem. 1973;248: 7724–7728. Available: http://www.jbc.org/content/248/22/7724.abstract
9. Cerkasovova A, Lukasova G, Cerkasov J, Kulda J. Biochemical characterization of large granule fraction of *Tritrichomonas foetus* (KV1 strain). J Protozool. 1973;20.
10. Benchimol M, Souza W De. Fine structure and cytochemistry of the hydrogenosome of *Tritrichomonas foetus*. J Protozool. John Wiley & Sons, Ltd (10.1111); 1983;30: 422–425. doi:10.1111/j.1550-7408.1983.tb02942.x
11. Horner DS, Hirt RP, Kilvington S, Lloyd DG, Embley TM. Molecular data suggest an early acquisition of the mitochondrion endosymbiont. Proc R Soc London Ser B Biol Sci. 1996;263: 1053 LP – 1059. Available: http://rspb.royalsocietypublishing.org/content/263/1373/1053.abstract
12. Bui ET, Bradley PJ, Johnson PJ. A common evolutionary origin for mitochondria and

hydrogenosomes. Proc Natl Acad Sci U S A. United States; 1996;93: 9651–9656.

13. Germot A, Philippe H, Le Guyader H. Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. Proc Natl Acad Sci. 1996;93: 14614 LP – 14617. Available: http://www.pnas.org/content/93/25/14614.abstract
14. Bradley PJ, Lahti CJ, Plumper E, Johnson PJ. Targeting and translocation of proteins into the hydrogenosome of the protist *Trichomonas*: Similarities with mitochondrial protein import. EMBO J. 1997;16: 3484–3493. doi:10.1093/emboj/16.12.3484
15. Embley M, der Giezen Mark V, Horner D, Dyal P, Foster P. Mitochondria and hydrogenosomes are two forms of the same fundamental organelle. Philos Trans R Soc London Ser B Biol Sci. Royal Society; 2003;358: 191–203. doi:10.1098/rstb.2002.1190
16. Sutak R, Dolezal P, Fiumera HL, Hrdy I, Dancis A, Delgadillo-Correa M, et al. Mitochondrial-type assembly of FeS centers in the hydrogenosomes of the amitochondriate eukaryote *Trichomonas vaginalis*. Proc Natl Acad Sci U S A. 2004;101: 10368–73. doi:10.1073/pnas.0401319101
17. Hrdy I, Hirt RP, Dolezal P, Bardonová L, Foster PG, Tachezy J, et al. *Trichomonas* hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. Nature. 2004;432: 618–622. doi:10.1038/nature03149
18. Fritz-Laylin LK, Prochnik SE, Ginger ML, Dacks JB, Carpenter ML, Field MC, et al. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. Cell. 2010;140: 631–642. doi:https://doi.org/10.1016/j.cell.2010.01.032
19. Barberà MJ, Ruiz-Trillo I, Tufts JYA, Bery A, Silberman JD, Roger AJ. *Sawyeria marylandensis* (Heterolobosea) has a hydrogenosome with novel metabolic properties. Eukaryot Cell. 2010;9: 1913 LP – 1924. doi:10.1128/EC.00122-10
20. Jerlström-Hultqvist J, Einarsson E, Xu F, Hjort K, Ek B, Steinhauf D, et al. Hydrogenosomes in the diplomonad *Spironucleus salmonicida*. Nat Commun. Nature Pub. Group; 2013;4: 2493. doi:10.1038/ncomms3493
21. Nývltová E, Stairs CW, Hrdý I, Rídl J, Mach J, Pačes J, et al. Lateral gene transfer and gene duplication played a key role in the evolution of *Mastigamoeba balamuthi* hydrogenosomes. Mol Biol Evol. 2015/01/07. Oxford University Press; 2015;32: 1039–1055. doi:10.1093/molbev/msu408
22. Yarlett N, Hann AC, Lloyd D, Williams A. Hydrogenosomes in the rumen protozoon *Dasytricha ruminantium* Schuberg. Biochem J. 1981;200: 365–372.

doi:10.1042/bj2000365

23. Yarlett N, Hann AC, Lloyd D, Williams AG. Hydrogenosomes in a mixed isolate of *Isotricha prostoma* and *Isotricha intestinalis* from ovine rumen contents. Comp Biochem Physiol Part B Comp Biochem. 1983;74: 357–364. doi:https://doi.org/10.1016/0305-0491(83)90025-1
24. van Bruggen JJA, Zwart KB, Hermans JGF, van Hove EM, Stumm CK, Vogels GD. Isolation and characterization of *Methanoplanus endosymbiosus* sp. nov., an endosymbiont of the marine sapropelic ciliate *Metopus contortus* quennerstedt. Arch Microbiol. 1986;144: 367–374.
25. Lloyd D, Hillman K, Yarlett N, Williams AG. Hydrogen production by rumen holotrich protozoa: Effects of oxygen and implications for metabolic control by in situ conditions. J Protozool. John Wiley & Sons, Ltd (10.1111); 1989;36: 205–213. doi:10.1111/j.1550-7408.1989.tb01075.x
26. Dyer BD. Metopus, Cyclidium and Sonderia: ciliates enriched and cultured from sulfureta of a microbial mat community. Biosystems. 1989;23: 41–51. doi:https://doi.org/10.1016/0303-2647(89)90007-5
27. Paul RG, Williams AG, Butler RD. Hydrogenosomes in the rumen entodiniomorphid ciliate *Polyplastron multivesiculatum*. J Gen Microbiol. Microbiology Society; 1990;136: 1981–1989. doi:10.1099/00221287-136-10-1981
28. Fenchel T, Finlay BJ. The biology of free-living anaerobic ciliates. Eur J Protistol. 1991;26: 201–215. doi:https://doi.org/10.1016/S0932-4739(11)80143-4
29. de Graaf RM, Duarte I, van Alen TA, Kuiper JWP, Schotanus K, Rosenberg J, et al. The hydrogenosomes of *Psalteriomonas lanterna*. BMC Evol Biol. 2009;9: 287. doi:10.1186/1471-2148-9-287
30. Tovar J, León-Avila G, Sánchez LB, Sutak R, Tachezy J, van der Giezen M, et al. Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. Nature. 2003;426: 172–6. doi:10.1038/nature01945
31. Mi-ichi F, Abu Yousuf M, Nakada-Tsukui K, Nozaki T. Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. Proc Natl Acad Sci U S A. 2009;106: 21731–6. doi:10.1073/pnas.0907106106
32. Stechmann A, Hamblin K, Pérez-Brocal V, Gaston D, Richmond GS, van der Giezen M, et al. Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. Curr Biol. Elsevier; 2008;18: 580–585. doi:10.1016/j.cub.2008.03.037

33. Leger MM, Kolisko M, Kamikawa R, Stairs CW, Kume K, Čepička I, et al. Organelles that illuminate the origins of *Trichomonas* hydrogenosomes and *Giardia* mitosomes. Nat Ecol Evol. 2017;1: 0092. doi:10.1038/s41559-017-0092
34. Tachezy J, editor. Hydrogenosomes and Mitosomes: Mitochondria of Anaerobic Eukaryotes [Internet]. Springer, Cham; 2019. doi:https://doi.org/10.1007/978-3-030-17941-0
35. Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petrželková R, Eme L, et al. A eukaryote without a mitochondrial organelle. Curr Biol. Elsevier; 2016;26: 1274–1284. doi:10.1016/j.cub.2016.03.053
36. Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, Ong S-E, et al. A mitochondrial protein compendium elucidates complex I disease biology. Cell. Elsevier; 2008;134: 112–123. doi:10.1016/j.cell.2008.06.016
37. Panigrahi AK, Ogata Y, Zíková A, Anupama A, Dalley RA, Acestor N, et al. A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome. Proteomics. WILEY-VCH Verlag; 2009;9: 434–450. doi:10.1002/pmic.200800477
38. Smith AC, Robinson AJ. MitoMiner v3.1, an update on the mitochondrial proteomics database. Nucleic Acids Res. 2015;44: D1258–D1261. doi:10.1093/nar/gkv1001
39. Calvo SE, Clauser KR, Mootha VK. MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res. 2015;44: D1251–D1257. doi:10.1093/nar/gkv1003
40. Morgenstern M, Stiller SB, Lübbert P, Peikert CD, Dannenmaier S, Drepper F, et al. Definition of a high-confidence mitochondrial proteome at quantitative scale. Cell Rep. Elsevier; 2017;19: 2836–2852. doi:10.1016/j.celrep.2017.06.014
41. Pfanner N, Warscheid B, Wiedemann N. Mitochondrial proteins: from biogenesis to functional networks. Nat Rev Mol Cell Biol. 2019;20: 267–284. doi:10.1038/s41580-018-0092-0
42. Forner F, Foster LJ, Campanaro S, Valle G, Mann M. Quantitative proteomic comparison of rat mitochondria from muscle, heart, and liver. Mol & Cell Proteomics. 2006;5: 608 LP – 619. doi:10.1074/mcp.M500298-MCP200
43. Ohlmeier S, Kastaniotis AJ, Hiltunen JK, Bergmann U. The yeast mitochondrial proteome, a study of fermentative and respiratory growth. J Biol Chem. 2004;279: 3956–3979. Available: http://www.jbc.org/content/279/6/3956.abstract
44. Rada P, Doležal P, Jedelský PL, Bursac D, Perry AJ, Šedinová M, et al. The core components of organelle biogenesis and membrane transport in the hydrogenosomes of

Trichomonas vaginalis. PLoS One. 2011;6. doi:10.1371/journal.pone.0024428

45. Schneider RE, Brown MT, Shiflett AM, Dyall SD, Hayes RD, Xie Y, et al. The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. Int J Parasitol. 2011;41: 1421–1434. doi:10.1016/j.ijpara.2011.10.001
46. Jedelský PL, Doležal P, Rada P, Pyrih J, Smíd O, Hrdý I, et al. The minimal proteome in the reduced mitochondrion of the parasitic protist *Giardia intestinalis*. PLoS One. Public Library of Science; 2011;6: e17285–e17285. doi:10.1371/journal.pone.0017285
47. Young JC, Hoogenraad NJ, Hartl FU. Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell. Elsevier; 2003;112: 41–50. doi:10.1016/S0092-8674(02)01250-3
48. Jores T, Lawatscheck J, Beke V, Franz-Wachtel M, Yunoki K, Fitzgerald JC, et al. Cytosolic Hsp70 and Hsp40 chaperones enable the biogenesis of mitochondrial β-barrel proteins. J Cell Biol. 2018;217: 3091 LP – 3108. doi:10.1083/jcb.201712029
49. Opaliński Ł, Song J, Priesnitz C, Wenz L-S, Oeljeklaus S, Warscheid B, et al. Recruitment of cytosolic J-proteins by TOM receptors promotes mitochondrial protein biogenesis. Cell Rep. Elsevier; 2018;25: 2036-2043.e5. doi:10.1016/j.celrep.2018.10.083
50. Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N. Importing mitochondrial proteins: Machineries and mechanisms. Cell. 2009;138: 628–644. doi:10.1016/j.cell.2009.08.005
51. Wiedemann N, Pfanner N. Mitochondrial machineries for protein import and assembly. Annu Rev Biochem. Annual Reviews; 2017;86: 685–714. doi:10.1146/annurev-biochem-060815-014352
52. Model K, Meisinger C, Prinz T, Wiedemann N, Truscott KN, Pfanner N, et al. Multistep assembly of the protein import channel of the mitochondrial outer membrane. Nat Struct Biol. 2001;8: 361–370. doi:10.1038/86253
53. Roise D, Horvath SJ, Tomich JM, Richards JH, Schatz G. A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphiphilic helix and perturb natural and artificial phospholipid bilayers. EMBO J. 1986;5: 1327–34. Available: http://www.ncbi.nlm.nih.gov/pubmed/3015598
54. von Heijne G, Steppuhn J, Herrmann RG. Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem. 1989;180: 535–45. Available: http://www.ncbi.nlm.nih.gov/pubmed/2653818

55. Abe Y, Shodai T, Muto T, Mihara K, Torii H, Nishikawa S, et al. Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20. Cell. Elsevier; 2000;100: 551–560. doi:10.1016/S0092-8674(00)80691-1

56. Maduke M, Roise D. Import of a mitochondrial presequence into protein-free phospholipid vesicles. Science. 1993;260: 364–7. Available: http://www.ncbi.nlm.nih.gov/pubmed/8385804

57. Šmíd O, Matušková A, Harris SR, Kučera T, Novotný M, Horváthová L, et al. Reductive evolution of the mitochondrial processing peptidases of the unicellular parasites *Trichomonas vaginalis* and *Giardia intestinalis*. PLOS Pathog. Public Library of Science; 2008;4: e1000243. Available: https://doi.org/10.1371/journal.ppat.1000243

58. Zimorski V, Major P, Hoffmann K, Brás XP, Martin WF, Gould SB. The N-terminal sequences of four major hydrogenosomal proteins are not essential for import into hydrogenosomes of *Trichomonas vaginalis*. J Eukaryot Microbiol. 2013;60: 89–97. doi:10.1111/jeu.12012

59. Garg S, Stölting J, Zimorski V, Rada P, Tachezy J, Martin WF, et al. Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. Genome Biol Evol. 2015;7: 2716–2726. doi:10.1093/gbe/evv175

60. Rada P, Kellerová P, Verner Z, Tachezy J. Investigation of the secretory pathway in *Trichomonas vaginalis* argues against a moonlighting function of hydrogenosomal enzymes. J Eukaryot Microbiol. John Wiley & Sons, Ltd (10.1111); 2019;0. doi:10.1111/jeu.12741

61. Dolezal P, Smíd O, Rada P, Zubácová Z, Bursać D, Suták R, et al. Giardia mitosomes and trichomonad hydrogenosomes share a common mode of protein targeting. Proc Natl Acad Sci U S A. 2005;102: 10924–9. doi:10.1073/pnas.0500349102

62. Glick BS, Brandt A, Cunningham K, Müller S, Hallberg RL, Schatz G. Cytochromes c1 and b2 are sorted to the intermembrane space of yeast mitochondria by a stop-transfer mechanism. Cell. 1992;69: 809–822. doi:https://doi.org/10.1016/0092-8674(92)90292-K

63. Chacinska A, Pfannschmidt S, Wiedemann N, Kozjak V, Sanjuán Szklarz LK, Schulze-Specking A, et al. Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. EMBO J. John Wiley & Sons, Ltd; 2004;23: 3735–3746. doi:10.1038/sj.emboj.7600389

64. Wiedemann N, Pfanner N, Ryan MT. The three modules of ADP/ATP carrier

cooperate in receptor recruitment and translocation into mitochondria. EMBO J. 2001;20: 951–60. doi:10.1093/emboj/20.5.951

65. Mossmann D, Meisinger C, Vögtle F-N. Processing of mitochondrial presequences. Biochim Biophys Acta - Gene Regul Mech. 2012;1819: 1098–1106. doi:https://doi.org/10.1016/j.bbagrm.2011.11.007

66. Kanaji S, Iwahashi J, Kida Y, Sakaguchi M, Mihara K. Characterization of the signal that directs Tom20 to the mitochondrial outer membrane. J Cell Biol. 2000;151: 277 LP – 288. doi:10.1083/jcb.151.2.277

67. Suzuki H, Maeda M, Mihara K. Characterization of rat TOM70 as a receptor of the preprotein translocase of the mitochondrial outer membrane. J Cell Sci. 2002;115: 1895 LP – 1905. Available: http://jcs.biologists.org/content/115/9/1895.abstract

68. Horie C, Suzuki H, Sakaguchi M, Mihara K. Characterization of signal that directs C-tail–anchored proteins to mammalian mitochondrial outer membrane. Fox TD, editor. Mol Biol Cell. The American Society for Cell Biology; 2002;13: 1615–1625. doi:10.1091/mbc.01-12-0570

69. Waizenegger T, Stan T, Neupert W, Rapaport D. Signal-anchor domains of proteins of the outer membrane of mitochondria: Structural and functional characteristics. J Biol Chem. 2003;278: 42064–42071. doi:10.1074/jbc.M305736200

70. Ahting U, Waizenegger T, Neupert W, Rapaport D. Signal-anchored proteins follow a unique insertion pathway into the outer membrane of mitochondria. J Biol Chem. 2005;280: 48–53. doi:10.1074/jbc.M410905200

71. Jores T, Klinger A, Groß LE, Kawano S, Flinner N, Duchardt-Ferner E, et al. Characterization of the targeting signal in mitochondrial β-barrel proteins. Nat Commun. The Author(s); 2016;7: 12036. Available: http://dx.doi.org/10.1038/ncomms12036

72. Kutik S, Stojanovski D, Becker L, Becker T, Meinecke M, Krüger V, et al. Dissecting membrane insertion of mitochondrial β-barrel proteins. Cell. 2008;132: 1011–1024. doi:https://doi.org/10.1016/j.cell.2008.01.028

73. Hill K, Model K, Ryan MT, Dietmeier K, Martin F, Wagner R, et al. Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. Nature. 1998;395: 516–521. doi:10.1038/26780

74. Ahting U, Thieffry M, Engelhardt H, Hegerl R, Neupert W, Nussberger S. Tom40, the pore-forming component of the protein-conducting TOM channel in the outer membrane of mitochondria. J Cell Biol. 2001;153: 1151–1160.

doi:10.1083/jcb.153.6.1151

75. Zeth K. Structure and evolution of mitochondrial outer membrane proteins of β-barrel topology. Biochim Biophys Acta - Bioenerg. 2010;1797: 1292–1299. doi:https://doi.org/10.1016/j.bbabio.2010.04.019
76. Shiota T, Imai K, Qiu J, Hewitt VL, Tan K, Shen H-H, et al. Molecular architecture of the active mitochondrial protein gate. Science (80-). 2015;349: 1544–1548. doi:10.1126/science.aac6428
77. Söllner T, Griffiths G, Pfaller R, Pfanner N, Neupert W. MOM19, an import receptor for mitochondrial precursor proteins. Cell. 1989;59: 1061–1070. doi:10.1016/0092-8674(89)90762-9
78. Söllner T, Pfaller R, Griffiths G, Pfanner N, Neupert W. A mitochondrial import receptor for the ADP/ATP carrier. Cell. 1990;62: 107–115. doi:10.1016/0092-8674(90)90244-9
79. Tsaousis AD, Gaston D, Stechmann A, Walker PB, Lithgow T, Roger AJ. A functional Tom70 in the human parasite Blastocystis sp.: Implications for the evolution of the mitochondrial import apparatus. Mol Biol Evol. 2011;28: 781–791. doi:10.1093/molbev/msq252
80. Perry AJ, Hulett JM, Likić VA, Lithgow T, Gooley PR. Convergent evolution of receptors for protein import into mitochondria. Curr Biol. 2006;16: 221–229. doi:10.1016/j.cub.2005.12.034
81. Makiuchi T, Mi-Ichi F, Nakada-Tsukui K, Nozaki T. Novel TPR-containing subunit of TOM complex functions as cytosolic receptor for Entamoeba mitosomal transport. Sci Rep. 2013;3: 1–7. doi:10.1038/srep01129
82. Mani J, Desy S, Niemann M, Chanfon A, Oeljeklaus S, Pusnik M, et al. Mitochondrial protein import receptors in Kinetoplastids reveal convergent evolution over large phylogenetic distances. Nat Commun. 2015;6: 6646. doi:10.1038/ncomms7646
83. Maćašev D, Whelan J, Newbigin E, Silva-Filho MC, Mulhern TD, Lithgow T. 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. 2004;21: 1557–1564. doi:10.1093/molbev/msh166
84. Kiebler M, Keil P, Schneider H, van der Klei IJ, Pfanner N, Neupert W. The mitochondrial receptor complex: A central role of MOM22 in mediating preprotein transfer from receptors to the general insertion pore. Cell. 1993;74: 483–492. doi:10.1016/0092-8674(93)80050-O

85. van Wilpe S, Ryan MT, Hill K, Maarse AC, Meisinger C, Brix J, et al. Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase. Nature. 1999;401: 485–489. doi:10.1038/46802

86. Shiota T, Mabuchi H, Tanaka-Yamano S, Yamano K, Endo T. In vivo protein-interaction mapping of a mitochondrial translocator protein Tom22 at work. Proc Natl Acad Sci. 2011;108: 15179–15183. doi:10.1073/pnas.1105921108

87. Allen R, Egan B, Gabriel K, Beilharz T, Lithgow T. A conserved proline residue is present in the transmembrane-spanning domain of Tom7 and other tail-anchored protein subunits of the TOM translocase. FEBS Lett. 2002;514: 347–350. doi:10.1016/S0014-5793(02)02433-X

88. Kanamori T, Nishikawa S, Nakai M, Shin I, Schultz PG, Endo T. Uncoupling of transfer of the presequence and unfolding of the mature domain in precursor translocation across the mitochondrial outer membrane. Proc Natl Acad Sci. 1999;96: 3634 LP – 3639. doi:10.1073/pnas.96.7.3634

89. Yamano K, Kuroyanagi-Hasegawa M, Esaki M, Yokota M, Endo T. Step-size analyses of the mitochondrial Hsp70 import motor reveal the Brownian ratchet in operation. J Biol Chem. 2008;283: 27325–27332. doi:10.1074/jbc.M805249200

90. Dekker PJ, Ryan MT, Brix J, Muller H, Honlinger A, Pfanner N. Preprotein translocase of the outer mitochondrial membrane: molecular dissection and assembly of the general import pore complex. Mol Cell Biol. United States; 1998;18: 6515–6524.

91. Brix J, Dietmeier K, Pfanner N. Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. J Biol Chem. 1997;272: 20730–20735. doi:10.1074/jbc.272.33.20730

92. Dietmeier K, Honlinger A, Bomer U, Dekker PJ, Eckerskorn C, Lottspeich F, et al. Tom5 functionally links mitochondrial preprotein receptors to the general import pore. Nature. England; 1997;388: 195–200. doi:10.1038/40663

93. Dolezal P, Makki A, Dyall SD. Protein import into hydrogenosomes and mitosomes. In: Tachezy J, editor. Hydrogenosomes and Mitosomes: Mitochondria of Anaerobic Eukaryotes. Cham: Springer International Publishing; 2019. pp. 31–84. doi:10.1007/978-3-030-17941-0_3

94. Ahting U, Thun C, Hegerl R, Typke D, Nargang FE, Neupert W, et al. The TOM core complex: The general protein import pore of the outer membrane of mitochondria. J Cell Biol. 1999;147: 959–968. doi:10.1083/jcb.147.5.959

95. Rapaport D, Neupert W. Biogenesis of Tom40, core component of the TOM complex of mitochondria. J Cell Biol. 1999;146: 321 LP – 332. doi:10.1083/jcb.146.2.321
96. Hönlinger A, Bömer U, Alconada A, Eckerskorn C, Lottspeich F, Dietmeier K, et al. Tom7 modulates the dynamics of the mitochondrial outer membrane translocase and plays a pathway-related role in protein import. EMBO J. 1996;15: 2125–37. doi:10.1002/j.1460-2075.1996.tb00566.x
97. Künkele K, Heins S, Dembowski M, Nargang FE, Benz R, Thieffry M, et al. The preprotein translocation channel of the outer membrane of mitochondria. Cell. 1998;93: 1009–1019. doi:10.1016/S0092-8674(00)81206-4
98. Model K, Meisinger C, Kühlbrandt W. Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase. J Mol Biol. Elsevier Ltd; 2008;383: 1049–1057. doi:10.1016/j.jmb.2008.07.087
99. Bausewein T, Mills DJ, Langer JD, Nitschke B, Nussberger S, Kühlbrandt W. Cryo-EM structure of the TOM core complex from *Neurospora crassa*. Cell. Elsevier; 2017;170: 693–700. doi:10.1016/j.cell.2017.07.012
100. Sakaue H, Shiota T, Ishizaka N, Kawano S, Tamura Y, Tan KS, et al. Porin associates with Tom22 to regulate the mitochondrial protein gate assembly. Mol Cell. 2019;73: 1044-1055.e8. doi:https://doi.org/10.1016/j.molcel.2019.01.003
101. Lee AC, Xu X, Blachly-Dyson E, Forte M, Colombini M. The role of yeast VDAC genes on the permeability of the mitochondrial outer membrane. J Membr Biol. 1998;161: 173–181. doi:10.1007/s002329900324
102. Ruan L, Zhou C, Jin E, Kucharavy A, Zhang Y, Wen Z, et al. Cytosolic proteostasis through importing of misfolded proteins into mitochondria. Nature. Macmillan Publishers Limited, part of Springer Nature. All rights reserved.; 2017;543: 443. Available: https://doi.org/10.1038/nature21695
103. Weidberg H, Amon A. MitoCPR—A surveillance pathway that protects mitochondria in response to protein import stress. Science (80-). 2018;360: eaan4146. doi:10.1126/science.aan4146
104. Mårtensson CU, Priesnitz C, Song J, Ellenrieder L, Doan KN, Boos F, et al. Mitochondrial protein translocation-associated degradation. Nature. 2019;569: 679–683. doi:10.1038/s41586-019-1227-y
105. Kozjak V, Wiedemann N, Milenkovic D, Lohaus C, Meyer HE, Guiard B, et al. An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane. J Biol Chem. 2003;278: 48520–48523.

doi:10.1074/jbc.C300442200

106. Paschen SA, Waizenegger T, Stan T, Preuss M, Cyrklaff M, Hell K, et al. Evolutionary conservation of biogenesis of β-barrel membrane proteins. Nature. Macmillian Magazines Ltd.; 2003;426: 862. Available: http://dx.doi.org/10.1038/nature02208
107. Martincová E, Voleman L, Pyrih J, Žárský V, Vondráčková P, Kolísko M, et al. Probing the biology of *Giardia intestinalis* mitosomes using in vivo enzymatic tagging. Mol Cell Biol. 2015;35: 2864 LP – 2874. doi:10.1128/MCB.00448-15
108. Habib SJ, Waizenegger T, Niewienda A, Paschen SA, Neupert W, Rapaport D. The N-terminal domain of Tob55 has a receptor-like function in the biogenesis of mitochondrial β-barrel proteins. J Cell Biol. 2007;176: 77 LP – 88. doi:10.1083/jcb.200602050
109. Höhr AIC, Lindau C, Wirth C, Qiu J, Stroud DA, Kutik S, et al. Membrane protein insertion through a mitochondrial β-barrel gate. Science. 2018;359: eaah6834. doi:10.1126/science.aah6834
110. Milenkovic D, Kozjak V, Wiedemann N, Lohaus C, Meyer HE, Guiard B, et al. Sam35 of the mitochondrial protein sorting and assembly machinery is a peripheral outer membrane protein essential for cell viability. J Biol Chem. 2004;279: 22781–22785. doi:10.1074/jbc.C400120200
111. Ishikawa D, Yamamoto H, Tamura Y, Moritoh K, Endo T. Two novel proteins in the mitochondrial outer membrane mediate β-barrel protein assembly. J Cell Biol. 2004;166: 621 LP – 627. Available: http://jcb.rupress.org/content/166/5/621.abstract
112. Klein A, Israel L, Lackey SWK, Nargang FE, Imhof A, Baumeister W, et al. Characterization of the insertase for β-barrel proteins of the outer mitochondrial membrane. J Cell Biol. 2012;199: 599–611. Available: http://jcb.rupress.org/content/199/4/599.abstract
113. Chan NC, Lithgow T. The peripheral membrane subunits of the SAM complex function codependently in mitochondrial outer membrane biogenesis. Mol Biol Cell. 2007/10/31. The American Society for Cell Biology; 2008;19: 126–136. doi:10.1091/mbc.e07-08-0796
114. Stojanovski D, Guiard B, Kozjak-Pavlovic V, Pfanner N, Meisinger C. Alternative function for the mitochondrial SAM complex in biogenesis of α-helical TOM proteins. J Cell Biol. 2007;179: 881–893. doi:10.1083/jcb.200706043
115. Qiu J, Wenz L-S, Zerbes RM, Oeljeklaus S, Bohnert M, Stroud DA, et al. Coupling of

mitochondrial import and export translocases by receptor-mediated supercomplex formation. Cell. Elsevier; 2013;154: 596–608. doi:10.1016/j.cell.2013.06.033

116. Wenz L-S, Ellenrieder L, Qiu J, Bohnert M, Zufall N, van der Laan M, et al. Sam37 is crucial for formation of the mitochondrial TOM–SAM supercomplex, thereby promoting β-barrel biogenesis. J Cell Biol. 2015;210: 1047–1054. Available: http://jcb.rupress.org/content/210/7/1047.abstract
117. Meisinger C, Rissler M, Chacinska A, Szklarz LKS, Milenkovic D, Kozjak V, et al. The mitochondrial morphology protein Mdm10 functions in assembly of the preprotein translocase of the outer membrane. Dev Cell. 2004;7: 61–71. doi:10.1016/j.devcel.2004.06.003
118. Meisinger C, Wiedemann N, Rissler M, Strub A, Milenkovic D, Schönfisch B, et al. Mitochondrial protein sorting: Differentiation of β-barrel assembly Tom7-mediated segregation of Mdm10. J Biol Chem. 2006;281: 22819–22826. doi:10.1074/jbc.M602679200
119. Ellenrieder L, Opaliński Ł, Becker L, Krüger V, Mirus O, Straub SP, et al. Separating mitochondrial protein assembly and endoplasmic reticulum tethering by selective coupling of Mdm10. Nat Commun. The Author(s); 2016;7: 13021. Available: https://doi.org/10.1038/ncomms13021
120. Humphries AD, Streimann IC, Stojanovski D, Johnston AJ, Yano M, Hoogenraad NJ, et al. Dissection of the mitochondrial import and assembly pathway for human Tom40. J Biol Chem. 2005;280: 11535–11543. doi:10.1074/jbc.M413816200
121. Kozjak-Pavlovic V, Ross K, Benlasfer N, Kimmig S, Karlas A, Rudel T. Conserved roles of Sam50 and metaxins in VDAC biogenesis. EMBO Rep. 2007/05/18. 2007;8: 576–582. doi:10.1038/sj.embor.7400982
122. Ott C, Ross K, Straub S, Thiede B, Götz M, Goosmann C, et al. Sam50 functions in mitochondrial intermembrane space bridging and biogenesis of respiratory complexes. Mol Cell Biol. 2012;32: 1173 LP – 1188. doi:10.1128/MCB.06388-11
123. Jian F, Chen D, Chen L, Yan C, Lu B, Zhu Y, et al. Sam50 regulates PINK1-Parkin-mediated mitophagy by controlling PINK1 stability and mitochondrial morphology. Cell Rep. 2018;23: 2989–3005. doi:https://doi.org/10.1016/j.celrep.2018.05.015
124. Becker T, Pfannschmidt S, Guiard B, Stojanovski D, Milenkovic D, Kutik S, et al. Biogenesis of the mitochondrial TOM complex: Mim1 promotes insertion and assembly of signal-anchored receptors. J Biol Chem. 2008;283: 120–7. doi:10.1074/jbc.M706997200

125. Becker T, Wenz L-S, Krüger V, Lehmann W, Müller JM, Goroncy L, et al. The mitochondrial import protein Mim1 promotes biogenesis of multispanning outer membrane proteins. J Cell Biol. 2011;194: 387 LP – 395. Available: http://jcb.rupress.org/content/194/3/387.abstract
126. Dimmer KS, Papić D, Schumann B, Sperl D, Krumpe K, Walther DM, et al. A crucial role for Mim2 in the biogenesis of mitochondrial outer membrane proteins. J Cell Sci. 2012;125: 3464 LP – 3473. doi:10.1242/jcs.103804
127. Waizenegger T, Schmitt S, Zivkovic J, Neupert W, Rapaport D. Mim1, a protein required for the assembly of the TOM complex of mitochondria. EMBO Rep. 2005;6: 57–62. doi:10.1038/sj.embor.7400318
128. Krüger V, Becker T, Becker L, Montilla-Martinez M, Ellenrieder L, Vögtle F-N, et al. Identification of new channels by systematic analysis of the mitochondrial outer membrane. J Cell Biol. 2017;216: 3485 LP – 3495. doi:10.1083/jcb.201706043
129. Vitali DG, Käser S, Kolb A, Dimmer KS, Schneider A, Rapaport D. Independent evolution of functionally exchangeable mitochondrial outer membrane import complexes. Pfanner N, editor. Elife. eLife Sciences Publications, Ltd; 2018;7: e34488. doi:10.7554/eLife.34488
130. Curran SP, Leuenberger D, Schmidt E, Koehler CM. The role of the Tim8p–Tim13p complex in a conserved import pathway for mitochondrial polytopic inner membrane proteins. J Cell Biol. 2002;158: 1017 LP – 1027. doi:10.1083/jcb.200205124
131. Rehling P, Model K, Brandner K, Kovermann P, Sickmann A, Meyer HE, et al. Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. Science (80-). 2003;299: 1747–1751. doi:10.1126/science.1080945
132. Neupert W, Herrmann JM. Translocation of proteins into mitochondria. Annu Rev Biochem. Annual Reviews; 2007;76: 723–749. doi:10.1146/annurev.biochem.76.052705.163409
133. Kerscher O, Holder J, Srinivasan M, Leung RS, Jensen RE. The Tim54p–Tim22p complex mediates insertion of proteins into the mitochondrial inner membrane. J Cell Biol. 1997;139: 1663 LP – 1675. Available: http://jcb.rupress.org/content/139/7/1663.abstract
134. Kang Y, Baker MJ, Liem M, Louber J, McKenzie M, Atukorala I, et al. Tim29 is a novel subunit of the human TIM22 translocase and is involved in complex assembly and stability. Elife. 2016;5. doi:10.7554/eLife.17463
135. Ellenrieder L, Dieterle MP, Doan KN, Mårtensson CU, Floerchinger A, Campo ML, et

al. Dual role of mitochondrial porin in metabolite transport across the outer membrane and protein transfer to the inner membrane. Mol Cell. Elsevier; 2019;73: 1056-1065.e7. doi:10.1016/j.molcel.2018.12.014

136. Callegari S, Müller T, Schulz C, Lenz C, Jans DC, Wissel M, et al. A MICOS–TIM22 association promotes carrier import into human mitochondria. J Mol Biol. 2019; doi:https://doi.org/10.1016/j.jmb.2019.05.015
137. Truscott KN, Kovermann P, Geissler A, Merlin A, Meijer M, Driessen AJ, et al. A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. Nat Struct Biol. 2001;8: 1074–82. doi:10.1038/nsb726
138. Donzeau M, Káldi K, Adam A, Paschen S, Wanner G, Guiard B, et al. Tim23 links the inner and outer mitochondrial membranes. Cell. 2000;101: 401–412. doi:https://doi.org/10.1016/S0092-8674(00)80850-8
139. Popov-Čeleketić D, Mapa K, Neupert W, Mokranjac D. Active remodelling of the TIM23 complex during translocation of preproteins into mitochondria. EMBO J. John Wiley & Sons, Ltd; 2008;27: 1469–1480. doi:10.1038/emboj.2008.79
140. Martinez-Caballero S, Grigoriev SM, Herrmann JM, Campo ML, Kinnally KW. Tim17p regulates the twin pore structure and voltage gating of the mitochondrial protein import complex TIM23. J Biol Chem. 2007;282: 3584–93. doi:10.1074/jbc.M607551200
141. Mokranjac D, Sichting M, Popov-Čeleketić D, Mapa K, Gevorkyan-Airapetov L, Zohary K, et al. Role of Tim50 in the transfer of precursor proteins from the outer to the inner membrane of mitochondria. Mol Biol Cell. American Society for Cell Biology (mboc); 2009;20: 1400–1407. doi:10.1091/mbc.e08-09-0934
142. Chacinska A, Lind M, Frazier AE, Dudek J, Meisinger C, Geissler A, et al. Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. Cell. Elsevier; 2005;120: 817–829. doi:10.1016/j.cell.2005.01.011
143. van der Laan M, Wiedemann N, Mick DU, Guiard B, Rehling P, Pfanner N. A role for Tim21 in membrane-potential-dependent preprotein sorting in mitochondria. Curr Biol. 2006;16: 2271–2276. doi:10.1016/j.cub.2006.10.025
144. Ieva R, Schrempp SG, Opaliński Ł, Wollweber F, Höß P, Heißwolf AK, et al. Mgr2 functions as lateral gatekeeper for preprotein sorting in the mitochondrial inner membrane. Mol Cell. 2014;56: 641–652. doi:https://doi.org/10.1016/j.molcel.2014.10.010

145. Richter F, Dennerlein S, Nikolov M, Jans DC, Naumenko N, Aich A, et al. ROMO1 is a constituent of the human presequence translocase required for YME1L protease import. J Cell Biol. 2019;218: 598 LP – 614. doi:10.1083/jcb.201806093
146. Lytovchenko O, Melin J, Schulz C, Kilisch M, Hutu DP, Rehling P. Signal recognition initiates reorganization of the presequence translocase during protein import. EMBO J. John Wiley & Sons, Ltd; 2013;32: 886–898. doi:10.1038/emboj.2013.23
147. Kang P-J, Ostermann J, Shilling J, Neupert W, Craig EA, Pfanner N. Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. Nature. 1990;348: 137–143. doi:10.1038/348137a0
148. Horst M, Oppliger W, Rospert S, Schönfeld H, Schatz G, Azem A. Sequential action of two hsp70 complexes during protein import into mitochondria. EMBO J. 1997;16: 1842 LP – 1849. Available: http://emboj.embopress.org/content/16/8/1842.abstract
149. Banerjee R, Gladkova C, Mapa K, Witte G, Mokranjac D. Protein translocation channel of mitochondrial inner membrane and matrix-exposed import motor communicate via two-domain coupling protein. Pfanner N, editor. Elife. eLife Sciences Publications, Ltd; 2015;4: e11897. doi:10.7554/eLife.11897
150. Slutsky-Leiderman O, Marom M, Iosefson O, Levy R, Maoz S, Azem A. The interplay between components of the mitochondrial protein translocation motor studied using purified components. J Biol Chem. 2007;282: 33935–33942. doi:10.1074/jbc.M704435200
151. Bolliger L, Deloche O, Glick BS, Georgopoulos C, Jeno P, Kronidou N, et al. A mitochondrial homolog of bacterial GrpE interacts with mitochondrial hsp70 and is essential for viability. EMBO J. England; 1994;13: 1998–2006.
152. Kozany C, Mokranjac D, Sichting M, Neupert W, Hell K. The J domain–related cochaperone Tim16 is a constituent of the mitochondrial TIM23 preprotein translocase. Nat Struct Mol Biol. Nature Publishing Group; 2004;11: 234. Available: http://dx.doi.org/10.1038/nsmb734
153. Li Y, Dudek J, Guiard B, Pfanner N, Rehling P, Voos W. The presequence translocase-associated protein import motor of mitochondria: Pam16 functions in an antagonistic manner to Pam18. J Biol Chem. 2004;279: 38047–38054. doi:10.1074/jbc.M404319200
154. Pfanner N, Truscott KN. Powering mitochondrial protein import. Nat Struct Biol. 2002;9: 234–236. doi:10.1038/nsb0402-234
155. Singha UK, Hamilton V, Duncan MR, Weems E, Tripathi MK, Chaudhuri M. Protein

translocase of mitochondrial inner membrane in *Trypanosoma brucei*. J Biol Chem. 2012;287: 14480–14493. doi:10.1074/jbc.M111.322925

156. Harsman A, Oeljeklaus S, Wenger C, Huot JL, Warscheid B, Schneider A. The non-canonical mitochondrial inner membrane presequence translocase of trypanosomatids contains two essential rhomboid-like proteins. Nat Commun. The Author(s); 2016;7: 13707. Available: https://doi.org/10.1038/ncomms13707
157. Žárský V, Doležal P. Evolution of the Tim17 protein family. Biol Direct. 2016;11: 54. doi:10.1186/s13062-016-0157-y
158. Pyrihová E, Motyčková A, Voleman L, Wandyszewska N, Fišer R, Seydlová G, et al. A single TIM translocase in the mitosomes of *Giardia intestinalis* illustrates convergence of protein import machines in anaerobic eukaryotes. Martin B, editor. Genome Biol Evol. 2018;10: 2813–2822. doi:10.1093/gbe/evy215
159. Zubáčová Z, Novák L, Bublíková J, Vacek V, Fousek J, Rídl J, et al. The mitochondrion-like organelle of *Trimastix pyriformis* contains the complete glycine cleavage system. PLoS One. Public Library of Science; 2013;8: e55417. Available: https://doi.org/10.1371/journal.pone.0055417
160. Katinka MD, Duprat S, Cornillot E, Méténier G, Thomarat F, Prensier G, et al. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature. The Author(s); 2001;414: 450. Available: http://dx.doi.org/10.1038/35106579
161. Dolezal P, Dagley MJ, Kono M, Wolynec P, Likić VA, Foo JH, et al. The essentials of protein import in the degenerate mitochondrion of Entamoeba histolytica. PLoS Pathog. 2010;6. doi:10.1371/journal.ppat.1000812
162. Terziyska N, Lutz T, Kozany C, Mokranjac D, Mesecke N, Neupert W, et al. Mia40, a novel factor for protein import into the intermembrane space of mitochondria is able to bind metal ions. FEBS Lett. John Wiley & Sons, Ltd; 2005;579: 179–184. doi:10.1016/j.febslet.2004.11.072
163. Banci L, Bertini I, Cefaro C, Ciofi-Baffoni S, Gallo A, Martinelli M, et al. MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. Nat Struct Mol Biol. Nature Publishing Group; 2009;16: 198. Available: https://doi.org/10.1038/nsmb.1553
164. Liu Z, Li X, Zhao P, Gui J, Zheng W, Zhang Y. Tracing the evolution of the mitochondrial protein import machinery. Comput Biol Chem. 2011;35: 336–340. doi:https://doi.org/10.1016/j.compbiolchem.2011.10.005

165. Peleh V, Zannini F, Backes S, Rouhier N, Herrmann JM. Erv1 of *Arabidopsis thaliana* can directly oxidize mitochondrial intermembrane space proteins in the absence of redox-active Mia40. BMC Biol. 2017;15: 106. doi:10.1186/s12915-017-0445-8
166. Bragoszewski P, Wasilewski M, Sakowska P, Gornicka A, Böttinger L, Qiu J, et al. Retro-translocation of mitochondrial intermembrane space proteins. Proc Natl Acad Sci. 2015;112: 7713 LP – 7718. doi:10.1073/pnas.1504615112
167. Gakh O, Cavadini P, Isaya G. Mitochondrial processing peptidases. Biochim Biophys Acta - Mol Cell Res. 2002;1592: 63–77. doi:https://doi.org/10.1016/S0167-4889(02)00265-3
168. Mach J, Poliak P, Matusková A, Zárský V, Janata J, Lukes J, et al. An advanced system of the mitochondrial processing peptidase and core protein family in *Trypanosoma brucei* and multiple origins of the core I subunit in eukaryotes. Genome Biol Evol. Oxford University Press; 2013;5: 860–875. doi:10.1093/gbe/evt056
169. Kleiber J, Kalousek F, Swaroop M, Rosenberg LE. The general mitochondrial matrix processing protease from rat liver: structural characterization of the catalytic subunit. Proc Natl Acad Sci. 1990;87: 7978 LP – 7982. doi:10.1073/pnas.87.20.7978
170. Fukasawa Y, Oda T, Tomii K, Imai K. Origin and evolutionary alteration of the mitochondrial import system in eukaryotic lineages. Mol Biol Evol. 2017;34: 1574–1586. doi:10.1093/molbev/msx096
171. Zarsky V, Tachezy J, Dolezal P. Tom40 is likely common to all mitochondria. Curr Biol. 2012;22: R479–R481. doi:https://doi.org/10.1016/j.cub.2012.03.057
172. van Dooren GG, Yeoh LM, Striepen B, McFadden GI. The import of proteins into the mitochondrion of *Toxoplasma gondii*. J Biol Chem. 2016;291: 19335–19350. doi:10.1074/jbc.M116.725069
173. Mani J, Rout S, Desy S, Schneider A. Mitochondrial protein import - Functional analysis of the highly diverged Tom22 orthologue of *Trypanosoma brucei*. Sci Rep. Nature Publishing Group; 2017;7: 40738. doi:10.1038/srep40738
174. Noinaj N, Kuszak AJ, Gumbart JC, Lukacik P, Chang H, Easley NC, et al. Structural insight into the biogenesis of β-barrel membrane proteins. Nature. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2013;501: 385. Available: https://doi.org/10.1038/nature12521
175. Rassow J, Dekker PJT, van Wilpe S, Meijer M, Soll J. The preprotein translocase of the mitochondrial inner membrane: function and evolution. J Mol Biol. 1999;286: 105–120. doi:https://doi.org/10.1006/jmbi.1998.2455

176. Clements A, Bursac D, Gatsos X, Perry AJ, Civciristov S, Celik N, et al. The reducible complexity of a mitochondrial molecular machine. Proc Natl Acad Sci U S A. 2009;106: 15791–5. doi:10.1073/pnas.0908264106
177. Stefanovic S, Hegde RS. Identification of a targeting factor for posttranslational membrane protein insertion into the ER. Cell. 2007;128: 1147–1159. doi:https://doi.org/10.1016/j.cell.2007.01.036
178. Schuldiner M, Metz J, Schmid V, Denic V, Rakwalska M, Schmitt HD, et al. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. Cell. Cell Press; 2008;134: 634–645. doi:10.1016/j.cell.2008.06.025
179. Wang F, Whynot A, Tung M, Denic V. The mechanism of tail-anchored protein insertion into the ER membrane. Mol Cell. 2011;43: 738–750. doi:https://doi.org/10.1016/j.molcel.2011.07.020
180. Setoguchi K, Otera H, Mihara K. Cytosolic factor- and TOM-independent import of C-tail-anchored mitochondrial outer membrane proteins. EMBO J. John Wiley & Sons, Ltd; 2006;25: 5635–5647. doi:10.1038/sj.emboj.7601438
181. Borgese N, Colombo S, Pedrazzini E. The tale of tail-anchored proteins. J Cell Biol. 2003;161: 1013 LP – 1019. doi:10.1083/jcb.200303069
182. Kuroda R, Ikenoue T, Honsho M, Tsujimoto S, Mitoma J, Ito A. Charged amino acids at the carboxyl-terminal portions determine the intracellular locations of two isoforms of cytochrome b5. J Biol Chem. 1998;273: 31097–31102. doi:10.1074/jbc.273.47.31097
183. Cichocki BA, Krumpe K, Vitali DG, Rapaport D. Pex19 is involved in importing dually targeted tail-anchored proteins to both mitochondria and peroxisomes. Traffic. John Wiley & Sons, Ltd (10.1111); 2018;19: 770–785. doi:10.1111/tra.12604
184. Lutfullahoğlu-Bal G, Keskin A, Seferoğlu AB, Dunn CD. Bacterial tail anchors can target to the mitochondrial outer membrane. Biol Direct. 2017;12: 16. doi:10.1186/s13062-017-0187-0
185. Brambillasca S, Yabal M, Makarow M, Borgese N. Unassisted translocation of large polypeptide domains across phospholipid bilayers. J Cell Biol. 2006/11/27. The Rockefeller University Press; 2006;175: 767–777. doi:10.1083/jcb.200608101
186. World Health Organization. Global incidence and prevalence of selected curable sexually transmitted infections. 2008;
187. Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, et al. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science (80-).

2007;315: 207–212. doi:10.1126/science.1132894

188. Hrdý I, Müller M. Primary structure and eubacterial relationships of the pyruvate:ferredoxin oxidoreductase of the amitochondriate eukaryote *Trichomonas vaginalis*. J Mol Evol. 1995;41: 388–396. doi:10.1007/BF00186551
189. Pfanner N, Geissler A. Versatility of the mitochondrial protein import machinery. Nat Rev Mol Cell Biol. 2001;2: 339–349. doi:10.1038/35073006
190. Makki A, Rada P, Žárský V, Kereïche S, Kováčik L, Novotný M, et al. Triplet-pore structure of a highly divergent TOM complex of hydrogenosomes in *Trichomonas vaginalis*. PLOS Biol. 2019;17: e3000098.
191. Eilers M, Schatz G. Binding of a specific ligand inhibits import of a purified precursor protein into mitochondria. Nature. 1986;322: 228–232. doi:10.1038/322228a0
192. Rada P, Makki A, Zimorski V, Garg S, Hampl V, Hrdý I, et al. N-terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*. J Eukaryot Cell. 2015;14: 1264–1275. doi:10.1128/EC.00104-15.Address
193. Geissler A, Krimmer T, Bömer U, Guiard B, Rassow J, Pfanner N. Membrane potential-driven protein import into mitochondria. The sorting sequence of cytochrome b(2) modulates the deltapsi-dependence of translocation of the matrix-targeting sequence. Mol Biol Cell. The American Society for Cell Biology; 2000;11: 3977–3991. doi:10.1091/mbc.11.11.3977
194. Rada P, Makki A, Žárský V, Tachezy J. Targeting of tail-anchored proteins to *Trichomonas vaginalis* hydrogenosomes. Mol Microbiol. John Wiley & Sons, Ltd (10.1111); 2018;0. doi:10.1111/mmi.14175

PLOS BIOLOGY

RESEARCH ARTICLE

Triplet-pore structure of a highly divergent TOM complex of hydrogenosomes in *Trichomonas vaginalis*

Abhijith Makki[1], Petr Rada[1], Vojtěch Žárský[1], Sami Kereïche[2], Lubomír Kováčik[2], Marian Novotný[3], Tobias Jores[4], Doron Rapaport[4], Jan Tachezy[1]*

1 Department of Parasitology, Faculty of Science, Charles University, BIOCEV, Prague, Czech Republic, **2** Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, Czech Republic, **3** Department of Cell Biology, Faculty of Science, Charles University, Prague, Czech Republic, **4** Interfaculty Institute of Biochemistry, University of Tübingen, Tübingen, Germany

* tachezy@natur.cuni.cz

Citation: Makki A, Rada P, Žárský V, Kereïche S, Kováčik L, Novotný M, et al. (2019) Triplet-pore structure of a highly divergent TOM complex of hydrogenosomes in *Trichomonas vaginalis*. PLoS Biol 17(1): e3000098. https://doi.org/10.1371/journal.pbio.3000098

Academic Editor: André Schneider, Universitat Bern, SWITZERLAND

Received: October 24, 2018

Accepted: December 11, 2018

Published: January 4, 2019

Data Availability Statement: The mass spectrometry data are available via PRIDE repository under the project accession PXD010850.

Funding: This work was supported by the programs KONTAKT II (LH15254), NPU II (LQ1604) provided by the Ministry of Education, Youth and Sport (MEYS), CePaViP (CZ.02.1.01/0.0/0.0/16_019/0000759) (to JT) supported by ERD Funds, Charles University projects - GAUK 268715 and GAUK 250937 (to AM), UNCE 204022

Abstract

Mitochondria originated from proteobacterial endosymbionts, and their transition to organelles was tightly linked to establishment of the protein import pathways. The initial import of most proteins is mediated by the translocase of the outer membrane (TOM). Although TOM is common to all forms of mitochondria, an unexpected diversity of subunits between eukaryotic lineages has been predicted. However, experimental knowledge is limited to a few organisms, and so far, it remains unsettled whether the triplet-pore or the twin-pore structure is the generic form of TOM complex. Here, we analysed the TOM complex in hydrogenosomes, a metabolically specialised anaerobic form of mitochondria found in the excavate *Trichomonas vaginalis*. We demonstrate that the highly divergent β-barrel *T. vaginalis* TOM (TvTom)40-2 forms a translocation channel to conduct hydrogenosomal protein import. TvTom40-2 is present in high molecular weight complexes, and their analysis revealed the presence of four tail-anchored (TA) proteins. Two of them, Tom36 and Tom46, with heat shock protein (Hsp)20 and tetratricopeptide repeat (TPR) domains, can bind hydrogenosomal preproteins and most likely function as receptors. A third subunit, Tom22-like protein, has a short *cis* domain and a conserved Tom22 transmembrane segment but lacks a *trans* domain. The fourth protein, hydrogenosomal outer membrane protein 19 (Homp19) has no known homology. Furthermore, our data indicate that TvTOM is associated with sorting and assembly machinery (Sam)50 that is involved in β-barrel assembly. Visualisation of TvTOM by electron microscopy revealed that it forms three pores and has an unconventional skull-like shape. Although TvTOM seems to lack Tom7, our phylogenetic profiling predicted Tom7 in free-living excavates. Collectively, our results suggest that the triplet-pore TOM complex, composed of three conserved subunits, was present in the last common eukaryotic ancestor (LECA), while receptors responsible for substrate binding evolved independently in different eukaryotic lineages.

(to SK), and GAČR Centre of Excellence project P302/12/G157 (to LK). We thank Karel Harant and Pavel Talacko from the Laboratory of Mass Spectrometry, BIOCEV. We acknowledge the core facility Cryo-Electron Microscopy and Tomography, CEITEC supported by the CIISB (LM2015043 funded by MEYS CR), and the IMCF, BIOCEV, supported by the Czech-BioImaging RI project LM2015062. 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.

Abbreviations: αSCS, α-subunit of succinyl CoA synthetase; 2D, two-dimensional; 3D, three-dimensional; ATOM, archaic translocase of the outer membrane; BN-PAGE, blue native PAGE; CLANS, cluster analysis of sequences; CoA, coenzyme A; COG, clusters of orthologous groups; coIP, co-immunoprecipitation; Cryo-EM, Cryo electron microscopy; CTF, Contrast Transfer Function; cytME, cytoplasmic malic enzyme; DHFR, dihydrofolate reductase; DIC, differential interference contrast; Dox, doxycycline; ERAD, endoplasmic reticulum-associated protein degradation; FASP, filter-aided sample preparation; Fdx, ferredoxin; Fis1, mitochondrial fission 1; HA, human influenza hemagglutinin; HMM, hidden Markov model; Hmp, hydrogenosomal membrane protein; Homp, hydrogenosomal outer membrane protein; Hsp, heat shock protein; IMS, intermembrane space; IPTG, Isopropyl β-D-1-thiogalactopyranoside; ITS, internal-targeting sequence; LECA, last common eukaryotic ancestor; LFQ-MS, label-free quantitative mass spectrometry; MAFFT, multiple sequence alignment based on fast Fourier transform; MS, mass spectrometry; NCBI, National Center for Biotechnology Information; Ni-NTA, Ni-nitrilotriacetic acid; NTS, N-terminal targeting sequence; OD, optical density; OMM, outer mitochondrial membrane; PDB, Protein Data Bank; Pfam, Protein families; SAM, sorting and assembly machinery; SAR, Stramenopiles, Alveolata and Rhizaria; SD-Leu, synthetic drop-out medium without leucine; SDS-PAGE, sodium dodecyl sulphate-PAGE; STED, Stimulated Emission Depletion; TA, tail-anchored; TCA, tricarboxylic acid; TEM, transmission electron microscopy; TIM, translocase of the inner membrane; TMD, transmembrane domain; TMHMM, transmembrane helices HMM; TOM, translocase of the outer membrane; TPR, tetratricopeptide repeat; TrichDB, Trichomonas Genome Resource; TvTOM, *T. vaginalis* TOM; VDAC, voltage-dependent anion channel; YPG, yeast extract-peptone-glycerol.

Author summary

Mitochondria carry out many vital functions in the eukaryotic cells, from energy metabolism to programmed cell death. These organelles descended from bacterial endosymbionts, and during their evolution, the cell established a mechanism to transport nuclear-encoded proteins into mitochondria. Embedded in the mitochondrial outer membrane is a molecular machine, known as the translocase of the outer membrane (TOM) complex, that plays a key role in protein import and biogenesis of the organelle. Here, we provide evidence that the TOM complex of hydrogenosomes, a metabolically specialised anaerobic form of mitochondria in *Trichomonas vaginalis*, is composed of highly divergent core subunits and lineage-specific peripheral subunits. Despite the evolutionary distance, the *T. vaginalis* TOM (TvTOM) complex has a conserved triplet-pore structure but with a unique skull-like shape suggesting that the TOM in the early mitochondrion could have formed three pores. Our results contribute to a better understanding of the evolution and adaptation of protein import machinery in anaerobic forms of mitochondria.

Introduction

Mitochondria originated from proteobacterial endosymbionts [1], and over time, massive endosymbiotic gene transfer to the host nucleus or gene deletion forged the development of a mechanism for retargeting of nuclear-encoded proteins to the evolving organelle [2]. To cross the double membrane of the mitochondrion, the proteins had to pass through the translocase of the outer (TOM) and inner (TIM) membranes. It has been inferred that most modules of the import machinery were created de novo and the ancient TOM complex comprised at least three components, the β-barrel translocation channel-forming Tom40 and two tail-anchored (TA) proteins, Tom22 and Tom7 [3,4].

The TOM complex in yeast consists of Tom40 and six α-helical proteins: two that are anchored to the outer mitochondrial membrane (OMM) by an N-terminal transmembrane domain (TMD; Tom20 and Tom70) and four that are anchored by a C-terminal TMD (Tom22, Tom5, Tom6, and Tom7). Tom20 and Tom70, both carrying tetratricopeptide repeat (TPR) domains, serve as primary receptors recognising proteins with N-terminal targeting sequence (NTS) and internal-targeting sequences (ITSs), respectively [5,6]. A prominent feature of the TOM complex is the variation in receptors across different eukaryotic lineages. A signal-anchored Tom20 is present in animals and fungi, whereas plant Tom20 evolved independently with a C-terminal anchor [7]. Lineage-specific Tom20 and Tom60 without any TMD are present in amoebozoans [8,9]. Tom20 and Tom70 are essentially absent in the eukaryotic supergroup Excavata [10–12]. In the excavate *Trypanosoma brucei*, the TOM complex (named the archaic translocase of the outer membrane [ATOM]) has only two orthologues, a highly divergent Tom40 (ATOM40) and a Tom22-like protein (ATOM14) [11,13]. Instead of Tom70 and Tom20, two unique receptors were identified, a TA protein ATOM69 and a signal-anchored ATOM46 [11].

Structural studies of the contemporary TOM complex are exclusively based on fungi, *Saccharomyces cerevisiae* and *Neurospora crassa* [14,15]. The yeast TOM complex is highly dynamic, with the mature trimeric complex formed by three pores, alternately switching with a dimeric form containing two pores, which serves as a platform for the integration of a new Tom40 into the complex [16]. The assembly of the Tom40 precursor in the OMM is mediated by the sorting and assembly machinery (SAM) that consists of a central β-barrel subunit Sam50 and two peripheral subunits Sam35 and Sam37 in yeast. To promote β-barrel

biogenesis, TOM and SAM form a transient supercomplex [17,18]. The dimeric and trimeric TOM structures are stabilised by the highly conserved TMD of Tom22 [19]. This specific function of Tom22 and its conservation in most eukaryotes led to speculation that the ancient TOM complex may have been a trimeric form [12]. However, this concept remains unsettled as it has not been clarified whether *N. crassa* TOM complex forms a three-pore or a two-pore structure [15,20], and so far, the information on TOM structure from other organisms is unavailable. Thus, to understand what subunits contributed to the formation of the earliest translocases and to reconstruct the evolutionary steps, it is important to study the composition and the structure of the translocases in organisms harbouring different variants of mitochondria as well as in organisms from different eukaryotic supergroups. Highly reduced mitochondria known as hydrogenosomes and mitosomes are found in certain organisms adapted to an anaerobic lifestyle [21] with simplified import machinery. The most studied hydrogenosomes are those found in the Parabasalia group of excavates, which includes the human parasite *Trichomonas vaginalis*. *T. vaginalis* hydrogenosomes have lost the tricarboxylic acid (TCA) cycle, and the oxidative phosphorylation has been replaced by substrate-level ATP synthesis, with the concomitant production of hydrogen [22]. Hydrogenosomes have lost the organellar genome entirely [23], and consequently, all hydrogenosomal proteins are imported from the cytosol. Like mitochondria, the import of proteins into hydrogenosomes is dependent on the hydrogenosomal NTS [24]. However, some matrix proteins are imported into hydrogenosomes independent of an NTS, and therefore the NTS-independent route was proposed to represent an ancestral mode of protein import [25,26]. Previous proteomic analysis of *T. vaginalis* hydrogenosomes revealed the presence of several β-barrel proteins of the mitochondrial porin 3 superfamily that were designated as putative Tom40. However, the protein sequences were highly divergent from known homologues, making it difficult to unequivocally distinguish between Tom40 and voltage-dependent anion channel (VDAC) [10]. Other hydrogenosomal β-barrel proteins include Sam50 and paralogues of two proteins of unknown function, hydrogenosomal membrane protein 35 (Hmp35) and Hmp36 [10,27]. Neither genomic nor proteomic analyses indicated the presence of other TOM components [10,28]. Hydrogenosomes also lack Tim50 and its regulatory subunit Tim21 that links the TOM complex with TIM in the intermembrane space (IMS) [10,28,29]. Furthermore, five paralogues of the Tim17/22/23 family that constitute the TIM channel have been detected. However, limited similarity of these hydrogenosomal proteins to Tim17, Tim22, and Tim23 subfamilies prevented determining whether they form a single multifunctional channel or distinct TIM23 and TIM22 channels for the import of matrix and inner membrane proteins, respectively [10]. Thus, structure and function of the hydrogenosomal protein import machineries remains elusive.

In the present study, we focus on the *T. vaginalis* TOM complex (TvTOM) and demonstrate that this highly divergent translocase mediates protein import into hydrogenosomes. Despite remarkable divergence in both primary structure and evolutionary distance, electron microscopy revealed some structural similarity between TvTOM and yeast three-pore TOM complex. However, the presence of an extra density provides a unique skull-like shape to TvTOM. Mass spectrometry (MS) of TvTOM and bioinformatic analysis identified two conserved and three lineage-specific TOM subunits, including two receptors, and revealed an association of TvTOM with Sam50. Although we did not identify Tom7 in TvTOM, our phylogenetic profiling predicted Tom7 in free-living representatives of Excavata. We propose that Tom40, Tom22, and probably Tom7 were present in the last common eukaryotic ancestor (LECA) and constituted a triplet-pore TOM complex, whereas the receptor subunits evolved independently in different eukaryotic lineages.

Results

Bioinformatic analyses of Tom40-like proteins

Seven Tom40-like proteins, named TvTom40-1 to TvTom40-7, identified in the hydrogenosomal proteome [10] displayed remarkably low sequence identity with fungal Tom40 sequences (e.g., 10%–14% identity compared with *N. crassa*). All TvTom40 proteins carry a conserved β-motif, PxGxxHxH (P = polar; G = glycine; H = hydrophobic; x = any amino acid) in the last β-strand similar to Tom40s and VDACs of other eukaryotes except TvTom40-3, where the last hydrophobic amino acid has been replaced by a polar hydroxylic residue, serine (S1 Fig). Bioinformatic analyses for all the seven proteins using HHpred tool identified TvTom40-2 (TVAG_332970) as the closest homologue to Tom40 (S1 and S2 Tables). Next, we built a local Tom40 hidden Markov model (HMM), based on 24 well-annotated Tom40 sequences (S1 Data) that was employed to scan the *T. vaginalis* proteome with HMMER jackhmmer tool, and again, TvTom40-2 was identified as the best Tom40 candidate.

A homology model of TvTom40-2 was constructed based on the *N. crassa* Tom40 template. TvTom40-2 forms a typical 19-strand β-barrel structure, but with only one N-terminal helix instead of two helices observed in Tom40 of other eukaryotes. Furthermore, TvTom40-2 contains a unique loop between β-strands five and six that is positively charged (Fig 1A). Most of the positions responsible for the interactions with other TOM proteins in yeast [16] are not conserved in TvTom40-2 (S2 Fig). A comparison of the electrostatic potential revealed that TvTom40-2 and *N. crassa* Tom40 share both positively and negatively charged patches inside the barrel, whereas mouse VDAC is almost uniformly positively charged (Fig 1B). Hence, based on homology searches and modeling, TvTom40-2 was chosen for further experimental studies.

TvTom40-2 forms a high molecular weight complex in the hydrogenosomal outer membrane

To verify the cellular localisation of TvTom40-2, a strain expressing C-terminally human influenza hemagglutinin (HA)-tagged TvTom40-2 was prepared. Immunofluorescence microscopy visualised TvTom40-2 as a ring, staining the membrane of hydrogenosomes. Malic enzyme, a hydrogenosomal marker enzyme, stained the organellar matrix (Fig 2A). Cell fractionation

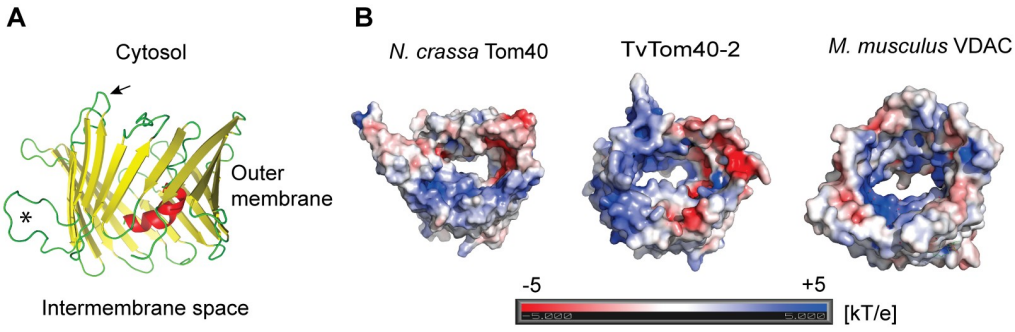


Fig 1. Homology model of TvTom40-2. (A) Model of TvTom40-2 was built using the *N. crassa* Tom40 structure (PDB ID 5o8o) as a template. The asterisk shows the extra loop between β-strands five and six, and the arrow shows the loop between β-strands four and five. (B) Comparison of 3D structures of *N. crassa* Tom40 (5o8o), TvTom40-2, and *Mus musculus* VDAC (3emn). Mouse VDAC is almost uniformly positively charged inside the barrel to bind negatively charged small molecules (ATP), while TvTom40-2 and *N. crassa* Tom40 share both positively and negatively charged patches inside the barrel. The scale of the electrostatic potential ranges from −5 to +5 kT/e. 3D, three-dimensional; PDB, Protein Data Bank; TOM, translocase of the outer membrane; TvTom, *T. vaginalis* TOM; VDAC, voltage-dependent anion channel.

https://doi.org/10.1371/journal.pbio.3000098.g001

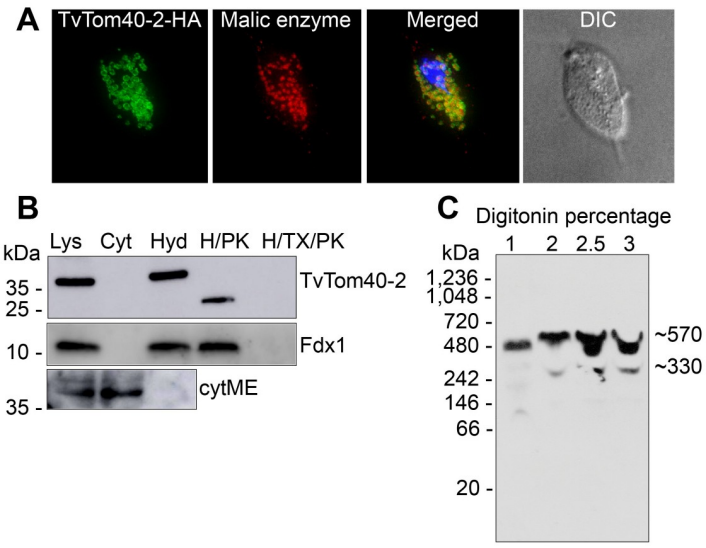


Fig 2. Localisation of TvTom40-2 in the hydrogenosomal outer membrane. (A) HA-tagged TvTom40-2 and malic enzyme (hydrogenosomal matrix protein) were visualised using mouse α-HA (green) and rabbit α-malic enzyme (red) antibodies, respectively. The nucleus was stained with DAPI (blue). (B) Localisation and topology of TvTom40-2 in *T. vaginalis* subcellular fractions. Immunoblot analysis of the whole cell lysate, cytoplasm, hydrogenosomes, hydrogenosomes treated with proteinase K, and hydrogenosomes treated with proteinase K in the presence of Triton X-100 using antibodies against HA, Fdx1 (hydrogenosomal matrix protein), and cytosolic malic enzyme. (C) BN-PAGE immunoblots of digitonin-lysed hydrogenosomal extract from the strain expressing HA-tagged TvTom40-2. The samples were probed with α-HA antibody. BN-PAGE, blue native PAGE; Cyt, cytoplasm; cytME, cytoplasmic malic enzyme; DIC, differential interference contrast; Fdx, ferredoxin; H/PK, hydrogenosomes treated with proteinase K; H/TX/PK, hydrogenosomes treated with proteinase K in the presence of Triton X-100; HA, human influenza hemagglutinin; Hyd, hydrogenosomes; Lys, lysate; TOM, translocase of the outer membrane; TvTom, *T. vaginalis* TOM.

https://doi.org/10.1371/journal.pbio.3000098.g002

and immunoblotting revealed the presence of TvTom40-2 exclusively in the hydrogenosomal fraction (Fig 2B). Treatment of hydrogenosomes carrying HA-tagged TvTom40-2 with proteinase K resulted in a shift of the molecular weight from 37 kDa to 28 kDa, indicating that the protein was likely cleaved within the loop between the fourth and fifth β-strands that is oriented towards the cytosol (Figs 2B and 1A). Then, the isolated hydrogenosomes were solubilised with varying concentrations of digitonin (1%–3%), and the samples were subjected to blue native-PAGE (BN-PAGE). TvTom40-2 was observed to be present in two high molecular weight complexes of 570 kDa and 330 kDa (Fig 2C). These experiments demonstrate that TvTom40-2 is present in a high molecular weight complex embedded in the hydrogenosomal outer membrane.

TvTom40-2 was inserted into the OMM in *S. cerevisiae*

The striking divergence of hydrogenosomal TvTom40-2 from Tom40 orthologues prompted us to test whether biogenesis of TvTom40-2 is specific to the hydrogenosomal machinery or whether, despite the variance in the sequence, it could be integrated into the OMM of distant eukaryotes from Opisthokonta lineage. We expressed TvTom40-2 with a C-terminal HA tag in *S. cerevisiae*. TvTom40-2 appeared in the mitochondrial fraction together with the mitochondrial marker, aconitase (Fig 3A). Alkaline extraction showed that most of the TvTom40-2 was present, similar to the OMM protein, mitochondrial fission 1 (Fis1), in the membrane fraction (Fig 3B). Finally, treatment of isolated mitochondria with proteinase K resulted in the formation of a proteolytic fragment of TvTom40-2 that resembled the one observed with

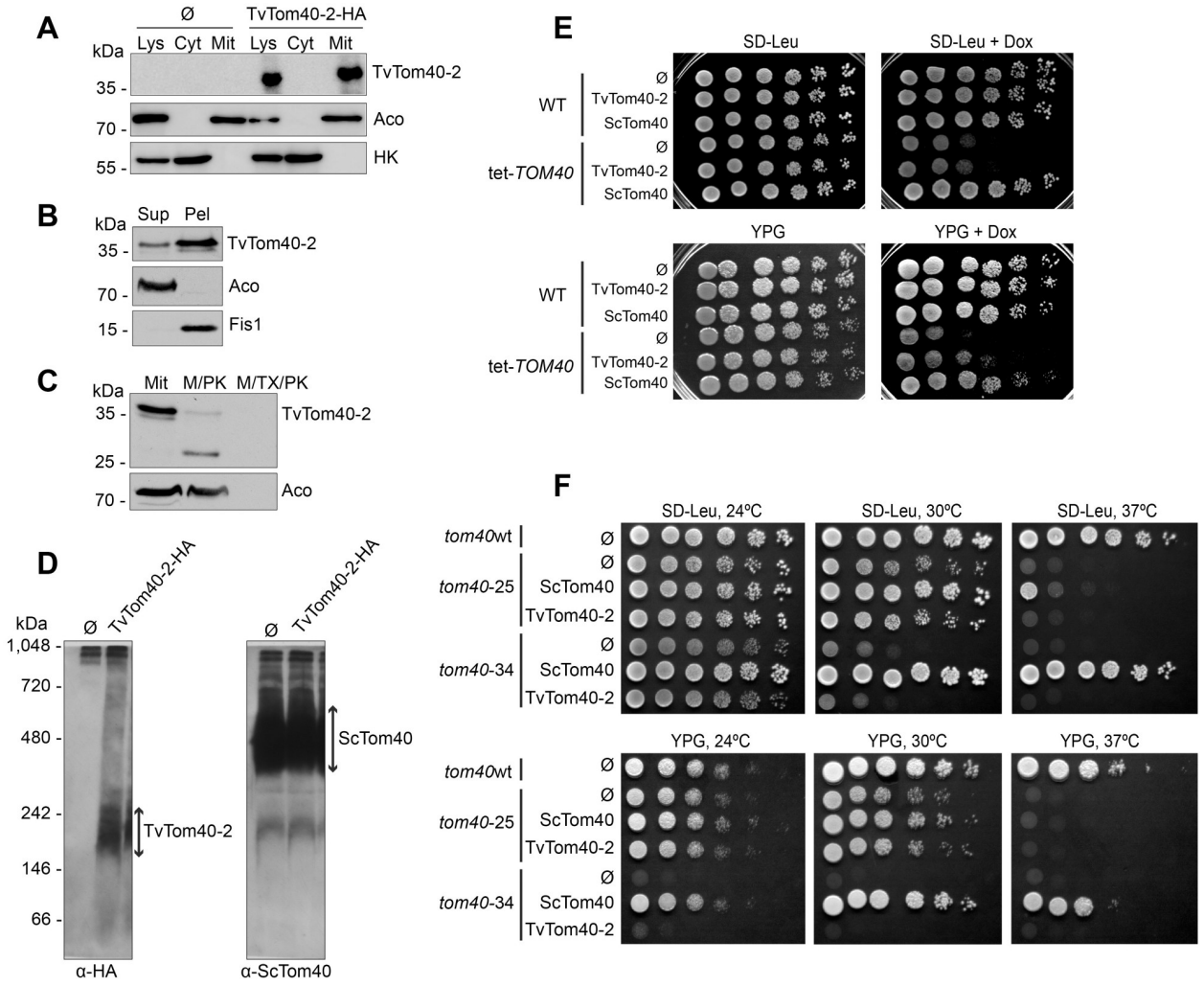


Fig 3. TvTom40-2 was assembled in the mitochondrial outer membrane in *S. cerevisiae*, and it could partially rescue the growth phenotype of *TOM40* mutants. (A) Whole cell lysate and fractions corresponding to cytoplasm and mitochondria were obtained from the WT strain transformed with an empty plasmid (Ø) or a plasmid encoding HA-tagged TvTom40-2. Proteins were analysed by SDS-PAGE and immunodecorated with antibodies against HA, aconitase (mitochondrial matrix protein) and hexokinase (cytosolic protein). (B) The mitochondrial fraction of cells expressing HA-tagged TvTom40-2 were subjected to alkaline extraction. Samples corresponding to supernatant and pellet fractions were analysed by western blotting using antibodies against HA, Aco, and Fis1. (C) Mitochondria as in panel B were treated with proteinase K or with proteinase K in the presence of Triton X-100. Further analysis was as in panel A. (D) Mitochondria were isolated from the strains described in panel A and solubilised in digitonin-containing buffer. Samples were analysed by BN-PAGE and immunodecorated with the indicated antibodies. (E) WT and tet-*TOM40* cells transformed with empty plasmid (Ø) or with plasmid encoding either TvTom40-2 or ScTom40 were grown to an OD_{600} of 1.0 and spotted in a 1:5 dilution series on synthetic glucose-containing medium lacking Leucine, SD-Leu supplemented with Dox, rich glycerol-containing medium (YPG), or YPG supplemented with Dox. The plates were then incubated at 30 ˚C for 2 to 3 days. (F) WT strain transformed with empty plasmid (Ø), or *tom40*-25 and *tom40*-34 strains transformed with empty plasmid (Ø) or with a plasmid encoding either TvTom40-2 or ScTom40, were grown to an OD_{600} of 1.0 and spotted in a 1:5 dilution series on SD-Leu or YPG. The plates were then incubated at 24 ˚C, 30 ˚C, or 37 ˚C for 2 to 4 days. Aco, aconitase; BN-PAGE, blue native PAGE; Cyt, cytoplasm; Dox, doxycycline; Fis1, mitochondrial fission 1; HA, human influenza hemagglutinin; HK, hexokinase; Lys, lysate; M/PK, mitochondria treated with proteinase K; M/TX/PK, mitochondria treated with proteinase K in the presence of Triton X-100; Mit, mitochondria; OD, optical density; Pel, pellet; SD-Leu, synthetic drop-out medium without leucine; SDS-PAGE, sodium dodecyl sulphate-PAGE; Sup, supernatant; TOM, translocase of the outer membrane; TvTom, *T. vaginalis* TOM; WT, wild-type; YPG, yeast extract-peptone-glycerol.

https://doi.org/10.1371/journal.pbio.3000098.g003

isolated hydrogenosomes (Fig 3C). As expected, this fragment was completely degraded upon solubilisation of the organelles with the detergent. Collectively, these observations indicate that TvTom40-2 is localised in the OMM in yeast. In addition, to check whether TvTom40-2 could form an oligomeric complex in yeast mitochondria, we performed BN-PAGE. TvTom40-2

migrated in a 230 kDa complex, while ScTom40 migrated in a 480 kDa complex (Fig 3D). This suggests that TvTom40-2 can form in yeast mitochondria a high molecular weight complex, although of smaller size than that in hydrogenosomes.

TvTom40-2 partially suppresses the growth phenotype of yeast *TOM40* mutants

Because TvTom40-2 was integrated into the OMM of yeast, we wanted to test whether it could functionally replace ScTom40. First, we prepared a yeast mutant, tet-*TOM40*, such that the *TOM40* promoter was replaced by a tetracycline promoter via homologous recombination, which would deplete ScTom40 in the presence of doxycycline (Dox). As expected, the addition of Dox to the growth medium resulted in a growth retardation of the tet-*TOM40* mutant. When TvTom40-2 was overexpressed, it could not rescue the growth defect of the tet-*TOM40* strain on fermentable medium (synthetic drop-out medium without leucine, SD-Leu) but could do so on nonfermentable medium (yeast extract-peptone-glycerol [YPG]) (Fig 3E). To substantiate this observation, we performed functional complementation studies using two yeast strains harbouring temperature-sensitive alleles of *TOM40*—*tom40*-25 and *tom40*-34. When grown at 30 ˚C, the overexpression of TvTom40-2 partially restored the growth phenotype of the *tom40*-25 strain both on fermentable and nonfermentable media (Fig 3F). Such an effect was not observed in the same strain grown at elevated temperature (37 ˚C, Fig 3F). The growth of *tom40*-34 was not restored even at lower temperatures (Fig 3F). Thus, it seems that TvTom40-2 can only partially replace yeast Tom40 function.

Identification of the TvTOM components

To identify interaction partners for TvTom40-2, we performed co-immunoprecipitations (coIPs) of HA-tagged TvTom40-2 under crosslinking and native conditions, and the eluted proteins were analysed using label-free quantitative MS (LFQ-MS). CoIPs using anti-HA antibody were performed with hydrogenosomes isolated from both the strain expressing HA-tagged TvTom40-2 and the wild-type (WT) strain, used as a negative control. The analysis revealed that 50 and 36 proteins were enriched with HA-tagged TvTom40-2 under crosslinking and native conditions, respectively (S2 Data). As TOM proteins are embedded in the hydrogenosomal outer membrane, we searched for proteins with TMDs in the data sets using TMHMM and found 19 and 13 proteins for crosslinking and native coIPs, respectively. The intersection between the two data sets and the hydrogenosomal membrane proteome [10] contained five TvTom40 isoforms, two TA proteins named Tom36 and hydrogenosomal outer membrane protein 19 (Homp19), two Sam50 paralogues, and Hmp35 (S2 Data and Fig 4A). In addition, the intersection between the coIP data set under crosslinking conditions and the membrane proteome contained two more TA proteins named Tom46 and Homp38. Based on our previous results [10], we selected Tom36 for the reciprocal coIPs.

Proteins enriched in the HA-tagged Tom36 coIPs under crosslinking conditions included three isoforms of TvTom40, Sam50, Hmp35, Homp38, Tom46, and Homp19, whereas under native conditions, three isoforms of TvTom40, Sam50, and Hmp35 were enriched (S2 Data and Fig 4B). Altogether, the coIP and MS data indicated four TA candidate proteins, Homp19, Tom36, Homp38, and Tom46. InterProScan [30] predicted that Tom36, Homp38, and Tom46 would carry an N-terminal heat shock protein (Hsp)20-like chaperone domain, three TPR-like domains, and a C-terminal TMD. This domain architecture resembles the recently reported ATOM69 in *T. brucei* [11] (Fig 4C). Indeed, HHpred searches using Tom36 and Homp38 as queries against the *T. brucei* proteome revealed ATOM69 as the first hit, with e-values of

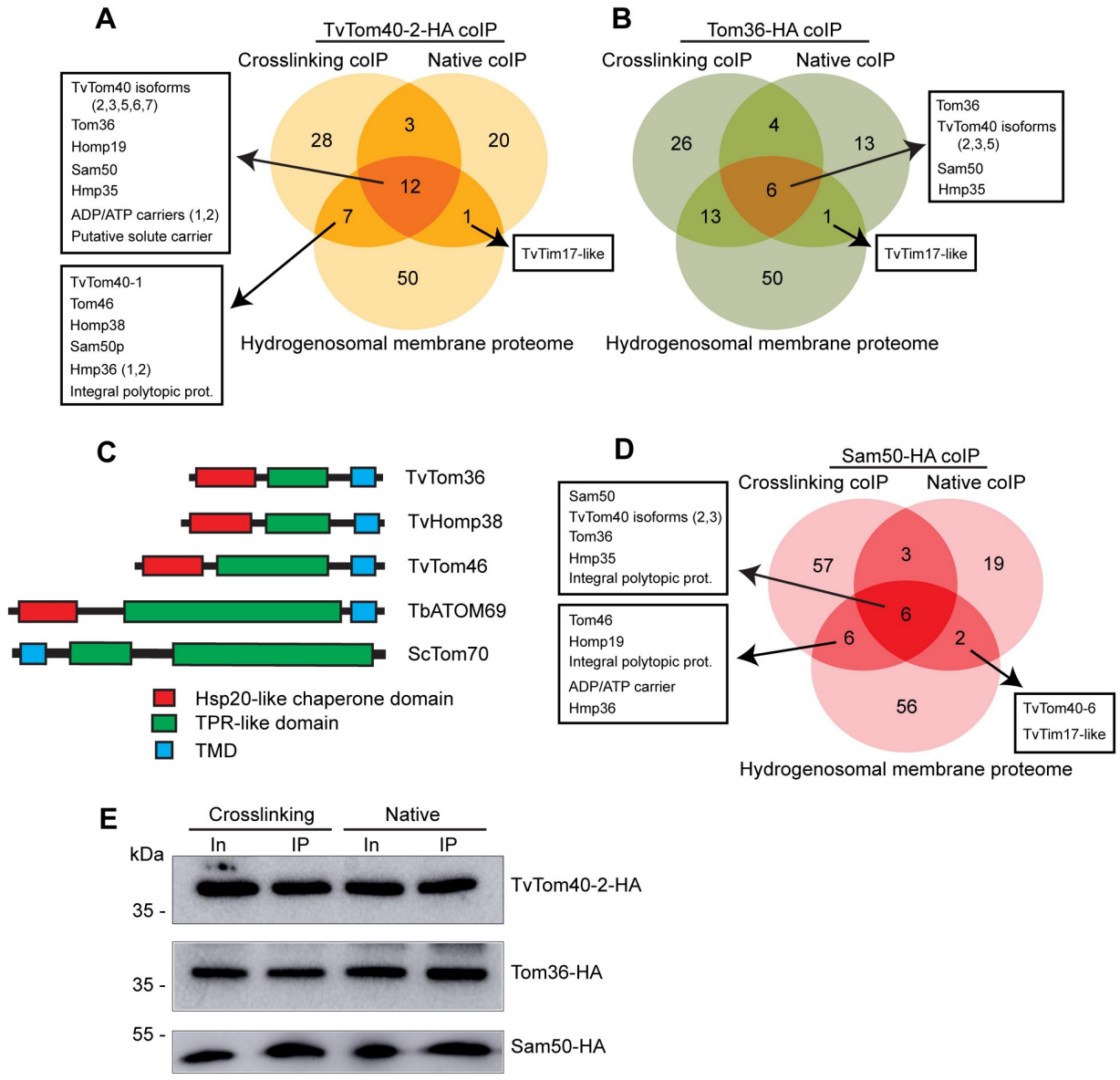


Fig 4. Identification of the components of the TvTOM complex. (A, B) Venn diagrams depicting the intersection between the hydrogenosomal membrane proteome and the proteins identified by LFQ-MS that were enriched in TvTom40-2-HA and Tom36-HA coIPs (under both crosslinking and native conditions), respectively. (C) Scheme of predicted domain architecture of Tom36, Homp38, and Tom46 in comparison with TbATOM69 and ScTom70. Hsp20-like chaperone domain, TPR-like domain, and TMD are represented by blue, green, and red, respectively. (D) Venn diagram depicting the intersection between the hydrogenosomal membrane proteome and the proteins identified by LFQ-MS that were enriched in Sam50-HA coIPs under both crosslinking and native conditions. (E) Immunoblots for the digitonin-lysed extract of hydrogenosomes (Input; 5%) and the IP eluates (2.5%) from TvTom40-2-HA, Tom36-HA, and Sam50-HA coIPs under crosslinking and native conditions decorated with α-HA antibody. ATOM, archaic translocase of the outer membrane; coIP, co-immunoprecipitation; HA, human influenza hemagglutinin; Homp, hydrogenosomal outer membrane protein; Hsp20, heat shock protein 20; In, Input; LFQ-MS, label-free quantitative mass spectrometry; Sam, sorting and assembly machinery; TMD, transmembrane domain; TOM, translocase of the outer membrane; TPR, tetratricopeptide repeat; TvTOM, *T. vaginalis* TOM.

https://doi.org/10.1371/journal.pbio.3000098.g004

4.9×10^{-17} and 2.3×10^{-11}, respectively. HHpred searches with Tom46 recognised various proteins with TPR domains, whereas no significant homology was observed for Homp19.

The coIP-MS data did not identify homologues of either Tom22 or Tom7. Thus, we used HMM to search for Tom22 and Tom7 sequences in the *T. vaginalis* protein database. The

searches for Tom22 identified a small protein with a predicted molecular weight of 6.4 kDa, containing a C-terminal TMD. It has a conserved Tom22 motif, including a tryptophan residue at the second position, followed by a few hydroxylated residues, with a serine at the +4 position and an invariant proline residue in the TMD; hence, we named it Tom22-like protein (TVAG_076160) (S3 Fig). In comparison to the fungal Tom22, Tom22-like protein is substantially shorter, similar to Tom22-like proteins in plants, apicomplexans, and kinetoplastids [4,31,32]. However, unlike Tom22, Tom22-like protein lacks a C-terminal IMS domain (S3 Fig). Searches for Tom7 in the *T. vaginalis* protein database did not identify a convincing orthologue.

Interestingly, Sam50 that only transiently associates with TOM in yeast [17] was copurified when both TvTom40-2 and Tom36 were pulled down both under crosslinking and native conditions, which may suggest a more stable association between TvTOM and Sam50. Therefore, we performed reciprocal coIPs using a strain expressing HA-tagged Sam50. LFQ-MS analysis revealed a similar spectrum of proteins as observed in the previous experiments that supports TvTOM-Sam50 association (S2 Data and Fig 4D). The presence of HA-tagged proteins in the eluates from TvTom40-2-HA, Tom36-HA, and Sam50-HA crosslinking and native coIPs were verified via immunoblotting (Fig 4E).

Hydrogenosomal localisation and the topology of TA proteins

To verify the localisation and topology of identified TA proteins, we prepared double transfectants that expressed TvTom40-2-HA together with one of the candidate proteins, all of which were C-terminally tagged with V5. In all cases, the TA protein was present in the hydrogenosomal fraction (Fig 5A). Treatment of isolated hydrogenosomes with proteinase K showed the presence of a truncated fragment that was protected from externally added proteinase K (Fig 5A). Next, we visualised V5-tagged candidate proteins, together with HA-tagged TvTom40-2, in the double transfectants using Stimulated Emission Depletion (STED) microscopy. All five candidates exhibited a ring-like pattern in the hydrogenosomal outer membrane similar to that observed with TvTom40-2 (Fig 5B). A Pearson correlation coefficient displayed the highest degrees of colocalisation with TvTom40-2 for Tom46 (77%) and Tom22-like protein (63%). Decreasing degrees of colocalisation with TvTom40-2 were observed for Tom36 (46%), Homp19 (26%), and Homp38 (17%). These experiments showed that all the selected TA proteins reside in the hydrogenosomal outer membrane.

TA proteins and Sam50 associated with TvTom40-2 are present in high molecular weight complex

To obtain further support for the association of identified TA proteins and Sam50 with the TvTOM complex, TvTom40-2-HA was pulled down from hydrogenosomes isolated from the double transfectants, and the samples were probed for V5-tagged proteins and Sam50 via immunoblotting using α-V5 and polyclonal α-Sam50 antibodies, respectively. Under crosslinking conditions, TvTom40-2 pulled down Tom36, Tom46, Homp19, and Tom22-like protein, while under native conditions, we observed a strong signal for Tom36, Homp19, and Tom22-like protein and a weaker signal for Tom46 (Fig 6A). Homp38 was not co-immunoprecipitated from the double transfectant under these conditions. On the other hand, Sam50 was detected in all samples analysed (Fig 6A). Furthermore, to validate whether the TvTom40-2-associated proteins are present in the high–molecular-weight complexes, hydrogenosomes isolated from the recombinant strains were subjected to BN-PAGE and immunoblotted with corresponding antibodies. Both Tom36 and Tom22-like protein migrated in 570 kDa and 330 kDa complexes. Tom46 and Homp19 migrated only in a 330 kDa complex, while Homp38 did

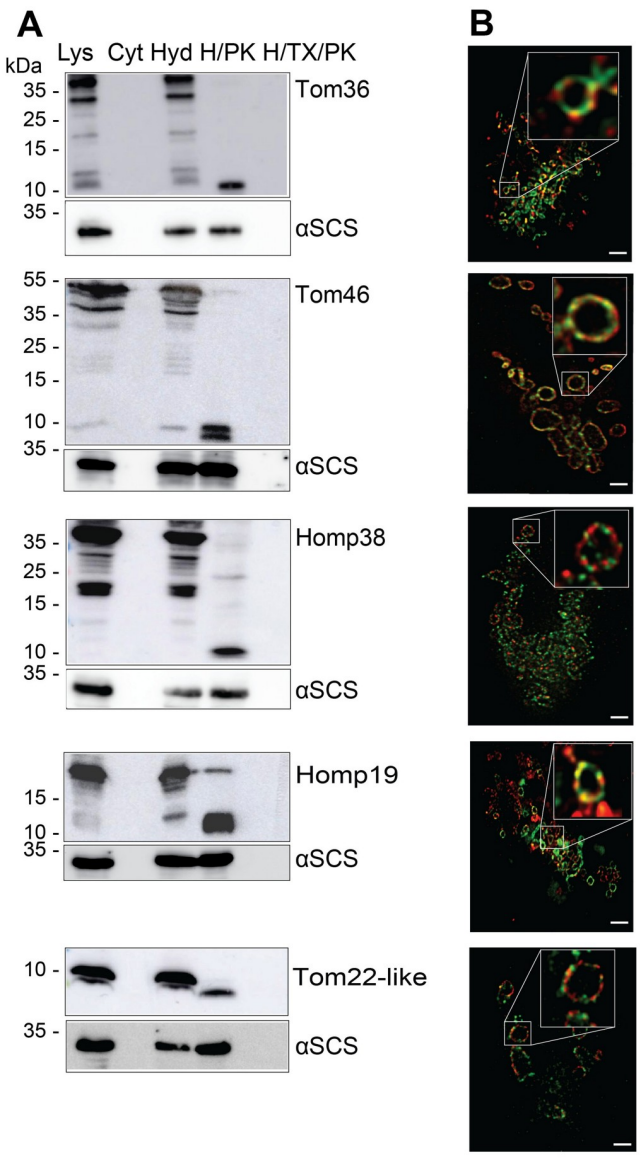


Fig 5. Localisation and topology of the TA proteins. (A) Immunoblot analysis of TA proteins in *T. vaginalis* subcellular fractions using α-V5 and α-αSCS (hydrogenosomal matrix protein) antibodies. Total cell lysates, cytoplasm, hydrogenosomes, hydrogenosomes treated with either proteinase K, or hydrogenosomes treated with proteinase K and Triton X-100 isolated from the strains expressing V5-tagged Tom36, Tom46, Homp38, Homp19, and Tom22-like protein. (B) Double transfectants expressing HA-tagged TvTom40-2 along with one of the V5-tagged proteins, Tom36, Tom46, Homp38, Homp19 or Tom22-like protein were visualised using mouse α-HA/α-mouse Abberior STAR 580 (green) and rabbit α-V5/α-rabbit Abberior STAR 635p (red) antibodies. Scale bar, 1 μm. αSCS, α-subunit of succinyl CoA synthetase; CoA, coenzyme A; Cyt, cytoplasm; H/PK, hydrogenosomes treated with proteinase K; H/TX/PK, hydrogenosomes treated with proteinase K in the presence of Triton X-100; HA, human influenza hemagglutinin; Homp, hydrogenosomal outer membrane protein; Hyd, hydrogenosomes; Lys, lysate; TA, tail-anchored; TOM, translocase of the outer membrane; TvTom, *T. vaginalis* TOM.

https://doi.org/10.1371/journal.pbio.3000098.g005

not appear to be present in any high molecular weight complex. TvTom40-2, used as a reference, migrated at 570 kDa and 330 kDa under the same conditions when immunodecorated with α-HA antibody (Fig 6B). HA-tagged Sam50 migrated at 570 kDa and 55 kDa, which corresponded to the high molecular weight of TvTOM complex and to Sam50 monomer, respectively (Fig 6B). These results confirmed the association of Tom36, Tom46, Homp19,

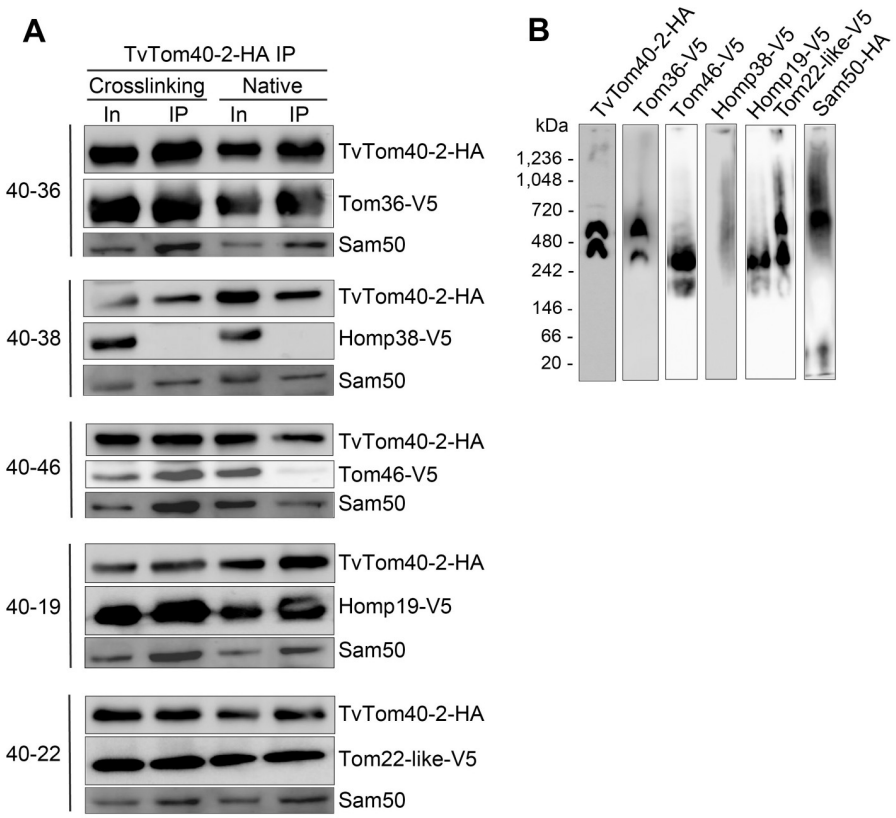


Fig 6. TA proteins and Sam50 associated with TvTom40-2 are present in a high molecular weight complex. (A) Digitonin-lysed extracts of hydrogenosomes isolated from the recombinant strains expressing both HA-tagged TvTom40-2 and one of the V5-tagged proteins, Tom36, Tom46, Homp38, Homp19, or Tom22-like protein were subjected to IP using α-HA antibody. Eluates from the IPs were probed for the presence of HA-tagged TvTom40-2, V5-tagged candidate proteins, and Sam50 under both crosslinking and native conditions using α-HA, α-V5, and polyclonal α-Sam50 antibodies, respectively. (B) BN-PAGE immunoblots of digitonin-lysed hydrogenosomal extracts from the strains expressing HA-tagged and V5-tagged proteins as indicated. BN-PAGE, blue native PAGE; HA, human influenza hemagglutinin; Homp, hydrogenosomal outer membrane protein; In, input; IP, immunoprecipitation; Sam, sorting and assembly machinery; TA, tail-anchored; TOM, translocase of the outer membrane; TvTom, *T. vaginalis* TOM.

https://doi.org/10.1371/journal.pbio.3000098.g006

Tom22-like, and Sam50 with TvTom40-2, and their ability to incorporate into high molecular complexes.

TvTom40-2 is involved in hydrogenosomal protein import

To demonstrate that the predicted TvTom40-2 participates in hydrogenosomal protein import, we performed an in vitro protein import and coIP assay. As an import substrate, we used the hydrogenosomal matrix protein ferredoxin (TvFdx1), which has an NTS fused to *Escherichia coli* dihydrofolate reductase (DHFR) at the C-terminus. TvFdx1-DHFR was synthesised in vitro in the presence of [^{35}S]-methionine. Under standard in vitro import conditions, using hydrogenosomes isolated from the double-transfected TvTom40-2-HA/Tom36-V5 strain, TvFdx1-DHFR was imported into hydrogenosomes, which was confirmed by a protease protection assay. The autoradiograph showed a time-dependent import of TvFdx1-DHFR (Fig 7A). Next, in vitro import assay was performed in the presence of methotrexate, which is known to cause the folding of DHFR and therefore arrests the

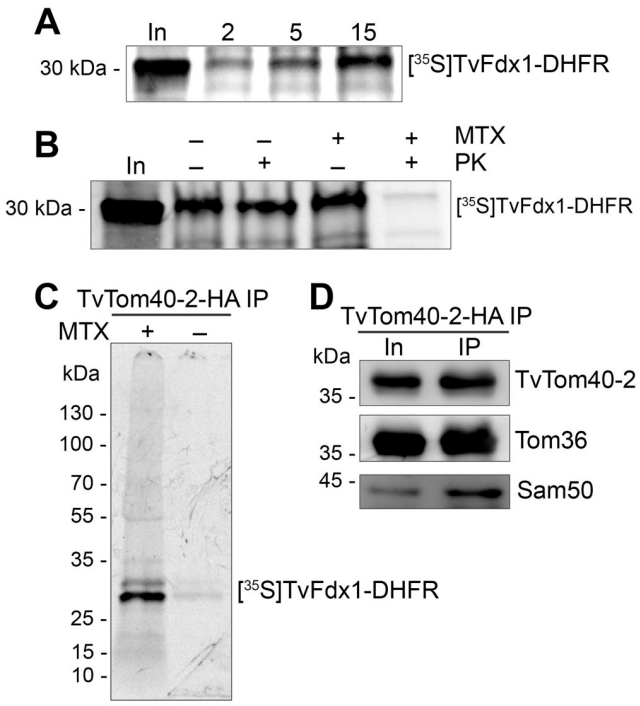


Fig 7. TvTom40-2 is involved in hydrogenosomal protein import. (A) Autoradiograph showing a time-dependent in vitro import of ^{35}S-Met-labeled TvFdx1-DHFR into hydrogenosomes. (B) Autoradiograph showing the in vitro import of ^{35}S-Met-labeled TvFdx1-DHFR into hydrogenosomes in either the absence (–) or the presence (+) of MTX, followed by proteinase K (+) treatment. (C) Autoradiograph showing the eluates for the TvTom40-2-HA coIP following the in vitro import of ^{35}S-Met-labeled TvFdx1-DHFR into hydrogenosomes isolated from a strain expressing both TvTom40-2-HA and Tom36-V5 either in the presence (+) or the absence (–) of MTX. (D) Immunoblot of the same eluates as in panel C using α-HA, α-V5, and α-Sam50 antibodies. coIP, co-immunoprecipitation; DHFR, dihydrofolate reductase; Fdx, ferredoxin; HA, human influenza hemagglutinin; In, input; MTX, methotrexate; PK, proteinase K; Sam, sorting and assembly machinery; TOM, translocase of the outer membrane; TvTom, *T. vaginalis* TOM.

https://doi.org/10.1371/journal.pbio.3000098.g007

translocating protein at the mitochondrial protein import site [33]. As expected, TvFdx1-DHFR was arrested at the hydrogenosomal outer membrane, and the exposed region was degraded when the hydrogenosomes were treated with proteinase K (Fig 7B). Finally, to prove that TvTom40-2, Tom36, and the substrate are present in the same complex, we performed in vitro import assay for TvFdx1-DHFR either in the presence or absence of methotrexate, crosslinked the interacting proteins, and immunoprecipitated the complex via TvTom40-2-HA. Autoradiography of the eluted sample revealed the presence of arrested TvFdx1-DHFR associated with the complex when methotrexate was added (Fig 7C). The two bands present on the autoradiograph (lane 1) correspond to TvFdx1-DHFR (30 kDa) and its proteolytically cleaved product (29 kDa) most likely. Immunoblot analysis of the complex confirmed the presence of TvTom40-2 and Tom36 in the same sample (Fig 7D). No substrate signal was observed when methotrexate was omitted from the reaction mixture (Fig 7C). These results demonstrate that TvFdx1-DHFR was imported into hydrogenosomes in an unfolded state and the arrested TvFdx1-DHFR was associated with TvTom40-2 and Tom36.

Tom36 and Tom46 can bind to hydrogenosomal preproteins

Because both Tom36 and Tom46 interact with TvTom40-2, are present in high–molecular-weight complexes, carry TPR-like domains and Hsp20-like chaperone domain that are

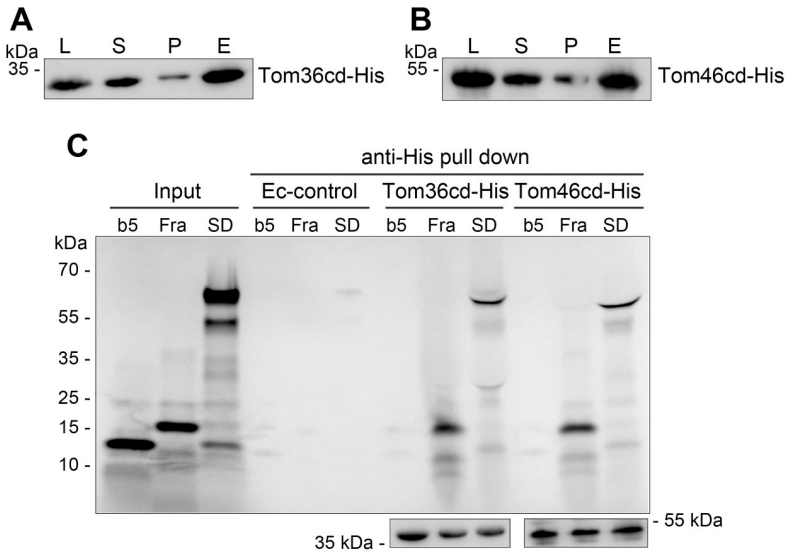


Fig 8. Tom36 and Tom46 can bind to hydrogenosomal preproteins. (A, B) Expression and coupling of His-tagged Tom36cd and Tom46cd to Ni-NTA agarose beads. *E. coli* cells expressing Tom36cd-His (panel A) or Tom46cd-His (panel B) were lysed (L; 2.5%), the lysate was centrifuged, and the supernatant with soluble proteins (S; 1%, input for the pull-down experiments) and pellet (P; 1%) fractions were obtained. The supernatant fraction was incubated with Ni-NTA agarose beads, and bound proteins were eluted (E; 5%) and probed on immunoblots using α-His antibody. (C) Binding assay. Proteins were pulled down from control *E. coli* or from cells expressing Tom36cd-His or Tom46cd-His using Ni-NTA agarose beads. The radiolabelled proteins cytochrome b5, frataxin, and αSCS-DHFR were incubated with various Ni-NTA agarose beads, and the His-tagged proteins were eluted using a buffer containing 500 mM imidazole. The samples were analysed by SDS-PAGE and autoradiography. The top panel shows an autoradiograph for the input radiolabelled proteins (Input; 10%) and the eluted fractions (20%). The bottom panel shows immunoblots using α-His antibody for Tom36cd-His and Tom46cd-His pull-down eluates (2.5%) from the binding assays. αSCS, α-subunit of succinyl CoA synthetase; b5, cytochrome b5; CoA, coenzyme A; DHFR, dihydrofolate reductase; Ec-control, control *E. coli*; Fra, frataxin; His, histidine; Ni-NTA, Ni-nitrilotriacetic acid; SD, αSCS-DHFR; SDS-PAGE, sodium dodecyl sulphate-PAGE; Tom, translocase of the outer membrane.

https://doi.org/10.1371/journal.pbio.3000098.g008

involved in protein–protein interactions, and are paralogues, we selected these proteins as receptor candidates. To test whether they can bind to hydrogenosomal proteins, we performed in vitro binding assay. The cytosolic domain of Tom36 (Tom36cd, residues 1–308) and Tom46 (Tom46cd, residues 1–402) were expressed with a C-terminal polyhistidine (His) tag in *E. coli* BL21 (DE3) strain, respectively, and coupled with Ni-nitrilotriacetic acid (Ni-NTA) agarose beads (S4 Fig and Fig 8A and 8B). Beads preincubated with untransformed *E. coli* lysate were used as a negative control. A cytosolic protein cytochrome b5 was used as a negative control. Radiolabelled precursors of two hydrogenosomal matrix proteins, frataxin and the α-subunit of succinyl coenzyme A (CoA) synthetase (αSCS), with the latter fused to DHFR at the C-terminus (αSCS-DHFR), were incubated with Tom36cd-His or Tom46cd-His coupled with or mock-treated beads for 1 hour. Then, the His-tagged proteins with the bound substrates were eluted with imidazole. The eluate from the Tom36cd-His and Tom46cd-His binding assay showed the presence of two radiolabeled proteins, frataxin and αSCS-DHFR (Fig 8C, top panel). The cytosolic cytochrome b5 was not observed to be bound to either Tom36cd-His or Tom46cd-His (Fig 8C, top panel). Furthermore, the eluates were immunoblotted with anti-His antibody to verify the presence of His-tagged proteins (Fig 8C, bottom panel). These experiments indicate that the cytosolic domain of Tom36 and Tom46 can bind hydrogenosomal preprotein substrates.

The TvTOM forms three protein translocation channels and has a unique skull-like structure

The diversity of TvTom40 paralogues and the presence of unusual components in the TvTOM complex prompted us to investigate the structure of the TvTOM complex via electron microscopy analysis. The hydrogenosomal TOM complex was purified from *T. vaginalis* expressing TvTom40-2-HA under native conditions. The isolated hydrogenosomes were solubilised with digitonin to release the complex, and then the TvTOM complex was purified by IP using α-HA antibody coupled to Dynabeads and negatively stained for electron microscopy. The identity of the HA-tagged TvTom40-2 in the IP eluate was verified by immunoblotting and silver staining (S5A and S5B Fig). The unprocessed electron micrographs mainly showed particles composed of ring-shaped structures with one, two, or three centers of stain accumulation (representative micrograph in S5C Fig). These stain-filled openings are interpreted as pores, each of which represents one channel of the protein translocase. A total of 10,038 particles were selected from 650 micrographs for further processing. Two-dimensional (2D) classification with 3,412 particles (34% of 10,038 particles) resulted in class averages representing TvTOM with one, two, or three pores of resolution between 21 and 34 Å (Fig 9A–9C). TvTOM with one or two pores were the most prominent, accounting for 35% (n = 1,175) and 40% (n = 1,377), respectively, while TvTOM with three pores accounted for 25% (n = 860). The single-pore particles were oval, 70 × 125 Å in size with an eccentric pore placement. Two-pore particles were oval or triangular and 140 × 100 Å in size. The particles with three pores were skull-shaped and measured 150 × 175 Å in size, although a fourth spot of stain accumulation with a low contrast was observed in one of the class averages (Fig 9C). A single translocation channel measured 70 Å in diameter, and the inner pore size of the channel measured 25–30 Å. The distance between two pore centers measured 50–60 Å. The most striking difference from the yeast TOM is the presence of an extra density, measuring 50 Å in diameter observed in most classes of single-, double-, and triple-pore TvTOM particles, suggestive of a subunit(s) interacting with the peripheral part of the channel formed by TvTom40.

Conserved core components and lineage-specific peripheral components of TOM complex in Excavata

Conservation of Tom40 and Tom22, and the identification of two novel peripheral components with Hsp20 and TPR domains (Tom36 and Tom46) suggest a peculiar evolutionary history for TvTOM complex. Therefore, we searched for orthologues of TOM components using a local HMM in selected genomes across different eukaryotic supergroups, with a focus on Excavata to estimate the conservation, gain, and loss of components (S3 Table and S3 Data). For our evolutionary scheme (Fig 10), we adapted a view that Excavata has two major sister groups: Metamonada, comprising anaerobic protists such as *T. vaginalis*, and Discoba, comprising *T. brucei* [34,35], although an alternative placement of Metamonada has been suggested [36]. Our phylogenomic profiling supported the current view that at least Tom40 and Tom22 are conserved in all eukaryotes and might have been present in the TOM complex of LECA (Fig 10). The only exception is *Monocercomonoides* sp., which has completely lost mitochondria including all genes coding for TOM and TIM components [37] (S3 Table and Fig 10). Support for Tom7 was less clear because neither *T. vaginalis* nor *T. brucei* seems to possess Tom7 (S3 Table and Fig 10). However, we took advantage of the available genome sequences of some free-living excavates [38–40] and identified putative Tom7 orthologues in *Carpediemonas membranifera* of Metamonada, and *Euglena gracilis* and *Stygiella incarcerata* of Discoba lineages (S3 Table and Fig 10). As expected, our searches showed that Tom20 and plant Tom20 were most likely gained independently in Opisthokonta and Viridiplantae,

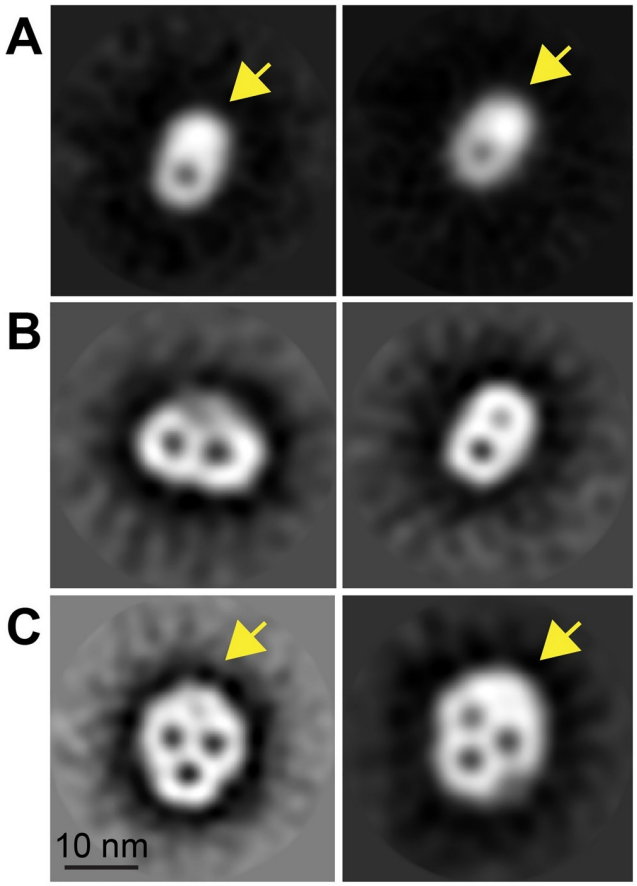


Fig 9. Electron microscopy analysis of the isolated TvTOM. Gallery of TvTOM class averages resulting from 2D classifications. Two class averages for TvTOM with (A) one pore at 24 Å resolution from 532 particles (left) and at 21 Å resolution from 517 particles (right); (B) two pores at 26 Å resolution from 348 particles (left) and at 31 Å resolution from 222 particles (right); and (C) three pores at 34 Å resolution from 298 particles (left) and at 21 Å resolution from 327 particles (right). Arrow indicates the additional mass. Scale bar, 10 nm. 2D, two-dimensional; TOM, translocase of the outer membrane; TvTom, *T. vaginalis* TOM.

https://doi.org/10.1371/journal.pbio.3000098.g009

respectively, and their orthologues are not present in other lineages, including Excavata (S3 Table and Fig 10). The evolutionary history of Tom70, Tom5, and Tom6 is more complex. All three components have been found in opisthokonts, while only Tom5 and Tom6 are present in Viridiplantae. Conversely, in the supergroup Stramenopiles, Alveolata and Rhizaria (SAR), which is related to Viridiplantae [34], Tom5 and Tom6 are absent, whereas Tom70 was reported in *Blastocystis*, other SAR species, and the haptophyte *Emiliania huxleyi* [41] (S3 Table and Fig 10). In our searches, none of these three components have been identified in both Excavata and Amoebozoa (S3 Table and Fig 10). The most puzzling aspect is the appearance of unique peripheral TOM components in the Excavata group. The searches for proteins with the same domain structure as Tom36 (Hsp20-TPR-TMD) in the available genome of 11 excavates and in the genome of selected organisms from other eukaryotic supergroups revealed the presence of homologous proteins only in *Tritrichomonas foetus*, a close relative of *T. vaginalis* (Parabasalia lineage), in kinetoplastids, and interestingly, in a fungus *Neocallimastix californiae* (S3 Table and Fig 10).

Next, we performed homology searches using Tom36 or ATOM69 as queries against the National Center for Biotechnology Information (NCBI) nonredundant protein database

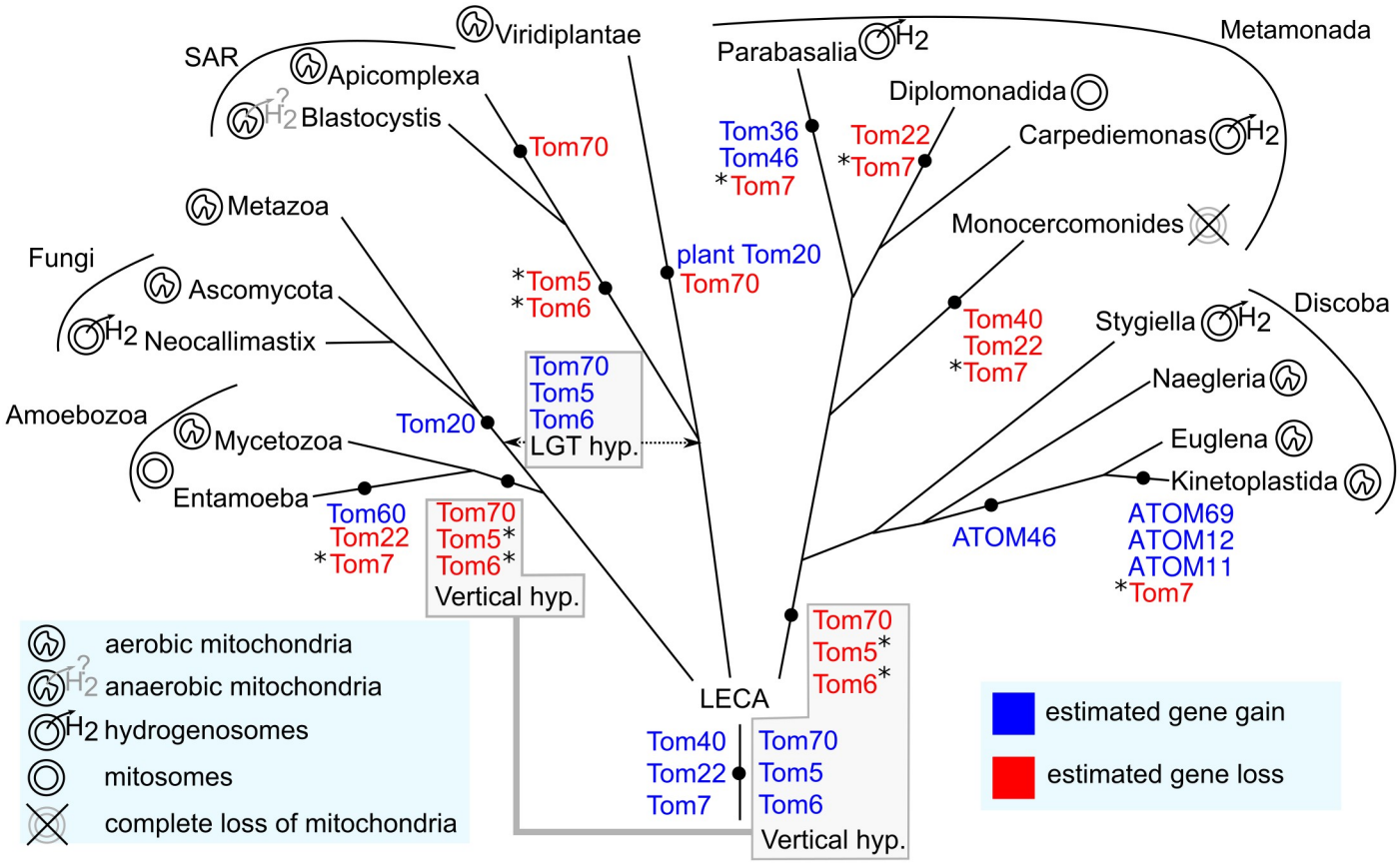


Fig 10. Phylogenetic scheme showing the gain and loss of TOM components across eukaryotic groups. Vertical gene transfer and LGT hypotheses for Tom5, Tom6, and Tom70 are in the boxes. Asterisk indicates small TOM subunits that were not identified; however, failure to identify them needs to be considered with caution. The relationships between the eukaryotic lineages are based on the recent phylogenetic results that employed concatenated gene data sets [42]. ATOM, archaic translocase of the outer membrane; LGT, lateral gene transfer; SAR, Stramenopiles, Alveolata and Rhizaria; TOM, translocase of the outer membrane.

https://doi.org/10.1371/journal.pbio.3000098.g010

regardless of the domain composition that resulted in a data set of 299 eukaryotic, 810 bacterial, and 5 archaeal sequences that were analysed using CLuster ANalysis of Sequences (CLANS) algorithm [43] (Fig 11A and S4 Data). Tom36 and Tom46 formed a cluster together with 10 other *T. vaginalis* and four *T. foetus* homologues (Fig 11A). All these homologues share Hsp20-TPR domains, two of them without any predicted TMD. A distinct cluster included seven ATOM69 homologues found in kinetoplastids that included dixenic, monoxenic, and free-living species (Fig 11A). The other clusters were formed by various TPR proteins, including elongation factor 2 kinase and endoplasmic reticulum-associated protein degradation (ERAD)-associated E3 ubiquitin-protein ligase (Fig 11A). The largest cluster predominantly contained bacterial proteins (Fig 11A). The formation of distinct clusters for Hsp20-TPR-TMD proteins of trichomonads and kinetoplastids suggests that Tom36/Tom46 and ATOM69 may have evolved independently in their respective lineages (Fig 11A). This view is supported by our phylogenetic analysis, in which Tom36/Tom46 and ATOM69 form two separate branches that are interleaved by a large bacterial group (Fig 11B).

Discussion

In spite of the fundamental role of mitochondrial translocases for the function and evolution of the eukaryotic cell, our experimental knowledge of the TOM complex is limited to a few

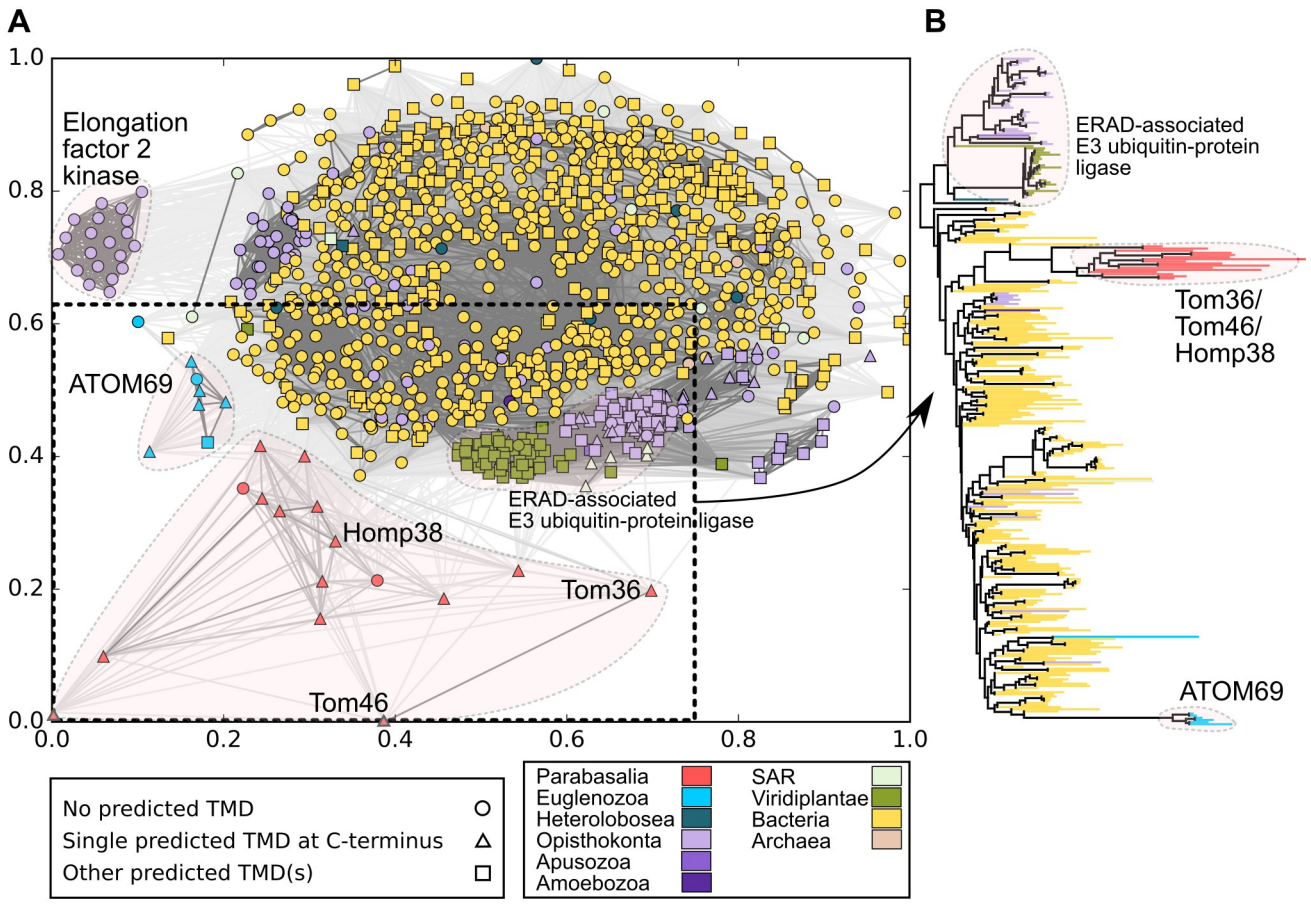


Fig 11. Relationship between Tom36/Tom46 and ATOM69. (A) CLANS similarity network for 1,114 homologues of Tom36 and ATOM69. The proteins from different eukaryotic and prokaryotic lineages are color coded. The prediction of the TMD using TMHMM is indicated by point shapes. For clarity, only 20% of the strongest connections between the proteins are shown in grey lines. The sequences and their coordinates for all the 1,114 proteins are given in S4 Data. Sequences within the marked rectangle were selected for the phylogeny. (B) Phylogeny of the TPR domains of Tom36, Tom46, ATOM69, and other related TPR proteins. The tree was constructed with IQ-TREE version 1.6.7 using the LG + I + G4 model and 10,000 ultra-fast bootstrap replicates. The sequences from different eukaryotic lineages and bacteria are color coded (418 taxa and 179 sites) (S5 Data). An enlarged version of the phylogenetic tree with accession numbers of taxa is shown in S6 Fig. ATOM, archai translocase of the outer membrane; CLANS, cluster analysis of sequences; ERAD, endoplasmic reticulum-associated protein degradation; Homp, hydrogenosomal outer membrane protein; SAR, Stramenopiles, Alveolata and Rhizaria; TMD, transmembrane domain; TMHMM, transmembrane helices hidden Markov model; Tom, translocase of the outer membrane; TPR, tetratricopeptide repeat.

https://doi.org/10.1371/journal.pbio.3000098.g011

model organisms, and direct visualisation of the TOM complex has only been achieved in two fungi, *S. cerevisiae* and *N. crassa* [14,15]. To extend our knowledge on TOM diversity in eukaryotes, we isolated and characterised the TOM complex from hydrogenosomes, an anaerobic form of mitochondria in *T. vaginalis*. In the present study, we have demonstrated the function of a highly divergent pore-forming TvTom40-2 and identified a protein that has limited homology with Tom22. The other components of TvTOM include three TA proteins with no orthologues in the fungal TOM complex. Furthermore, TvTOM seems to be tightly associated with Sam50 for a more efficient β-barrel biogenesis.

Electron microscopic visualisation of the TvTOM complex revealed interesting similarities and differences when compared with the TOM complex in fungi. Most observed TvTOM particles displayed two pores, which in fungi represent the TOM core complex, or particles with three pores, corresponding to the holo complex. The distance between two pore centers, the inner pore diameter, the single translocation channel diameter, and the size of the particles

with two pores are similar to those determined for the TOM complex in fungi [15,20]. The appearance of single-pore particles could more likely be either a result of the dissociation of holo complexes during experimental procedures [20,44] or stable assembly intermediates. A striking deviation from known TOM models is the presence of an extra density in the single-, double-, and triple-pore particles, providing a skull-like shape to the TvTOM holo complex. Based on coIP-MS analysis, it can be speculated that the extra mass may contain the identified β-barrel proteins Sam50 or Hmp35. In yeast, the TOM and SAM complexes form a labile supercomplex that allows coupling of the translocation of the Tom40 precursor through TOM and its insertion into the OMM via SAM [17]. It has been suggested that Sam50 may account for the third pore in the yeast triplet-pore complex [15]. Cryo electron microscopy (Cryo-EM) has shown that the Sam50 monomer measures 50 Å [45], which is consistent with the size of the additional mass observed in TvTOM. BN-PAGE analysis showed that HA-tagged Sam50 migrated with the high–molecular-weight complex of TvTOM or as a monomer. The enrichment of TOM subunits, as well as Sam50 in the reciprocal coIPs, supports a tight TOM-Sam50 association in hydrogenosomes. Formation of the supercomplex in yeast is mediated by the N-terminal cytosolic domain of Tom22 and Sam37 [17,18]. In trichomonads, Sam37 has not been identified [28], and Tom22 has a short cytosolic domain. Therefore, if the observed association of TvTOM and Sam50 represents a functional complex, different protein–protein interactions are to be expected. Hmp35 is a β-barrel protein in the hydrogenosomal membrane with an unknown function that exists in a stable 300 kDa complex of Hmp35 oligomers [27]. This complex is too large to imply the formation of a complex with TvTOM.

The presence of a TOM complex with three pores observed in *T. vaginalis* strongly indicates that triplet-pore complex is the generic form of TOM in eukaryotes that was inherited from LECA. It has been proposed that the ancient TOM complex contained—in addition to Tom40—Tom22, which tethers Tom40s using its TMD, and a regulatory subunit Tom7 [4,12,16,46]. The Excavata group includes two major lineages, Metamonada and Discoba, represented by *T. vaginalis* and *T. brucei*, respectively. Investigations of *T. brucei* TOM complex initially suggested that Tom40 in kinetoplastids (ATOM40) might be a homologue of the bacterial Omp85-like protein [13]. However, profile-sequence searches found that ATOM40 belongs to the eukaryotic porin family [12,47]. Our analysis, with an extended sampling of Excavata—which included a Tom40 orthologue in *E. gracilis*, which shares a common ancestry with kinetoplastids—confirmed this view.

Previous sequence searches implied the absence of Tom22 in some excavates with reduced forms of mitochondria, including the hydrogenosomes of *T. vaginalis* [12]. However, due to its short sequence and low conservation [4,12,32], the identification of Tom22 might have been beyond the sensitivity of most search tools. Our sensitive, structure-based HMM search identified a short 6 kDa Tom22-like protein as a potential candidate. This protein is tightly associated with TvTom40-2 in the hydrogenosomal outer membrane and is present in both high molecular weight complexes (570 and 330 kDa). Tom22-like protein contains a conserved TMD motif, including invariable tryptophan and proline residues, and a short cytosolic N-terminal (*cis*) domain similar to the 9 kDa Tom22 orthologue, Tom9 in higher plants, the 8 kDa apicomplexan Tom22, and the kinetoplastid Tom22 orthologue, ATOM14 [4,31,32]. The long acidic extension of the *cis* domain evolved only in opisthokonts that interacts with lineage-specific Tom20 and Tom70 [4], and therfore its absence in Tom22-like protein is not surprising. Most Tom22s contain an IMS-localised acidic (*trans*) domain that interacts with the substrate and enhances its transfer to Tim50 in the TIM23 complex [19]. Tom22-like protein identified here lacks the *trans* domain, which may reflect the absence of Tim50 in *T. vaginalis* [28]. In addition to *T. vaginalis* and *T. brucei*, we retrieved Tom22 orthologues from members of both Metamonada and Discoba in support of its presence in Excavata common ancestor.

Tom7 has not been identified in parabasalids, diplomonads, and in kinetoplastids. A fusion protein with limited sequence similarity to Tom7 and Tom22 has been reported in *Naegleria* species [12]. Importantly, Tom7 orthologues appears to be present in free-living members of both Excavata lineages, in *C. membranifera* (Metamonada), and *E. gracilis* and *S. incarcerata* (Discoba). This suggests that the absence of Tom7 might be a result of a secondary loss, and if so, it happened independently in certain lineages of both Metamonada and Discoba. However, failure to identify small Toms—Tom7 as well as Tom5 and Tom6—needs to be tread with caution. Their sequences are very short and might be highly divergent, particularly in parasitic lineages and those with reduced forms of mitochondria, which can hamper their identification. Collectively, our results suggest that the triplet-pore form of the TOM complex constituted the ancestral form of TOM in LECA.

Functional studies of TvTom40-2 using a DHFR-methotrexate system demonstrated that hydrogenosomal preprotein binds to TvTom40-2 and subsequently is imported into the hydrogenosomal matrix in an unfolded or loosely folded state, a feature that is conserved in mitochondria [33]. Of note, *T. vaginalis* has at least seven TvTom40 paralogues that are all expressed [10]. CoIP-MS analysis revealed that TvTom40-2 is associated with five other paralogues, and therefore various combinations of TvTom40 paralogues appear to be present in a single TOM complex, as observed in the rat TOM complex, in which two Tom40 isoforms interact with each other [48]. Further, we asked whether the hydrogenosomal TvTom40-2 could be integrated and can function in the yeast OMM. Despite low amino acid sequence conservation between TvTom40-2 and yeast orthologue, heterologous expression of TvTom40-2 in yeast resulted in its localisation in the OMM and the formation of a 230 kDa complex. This finding is consistent with the recent investigation of the targeting signal in β-barrel proteins, wherein the signal appears not to be encoded in a conserved linear amino acid sequence but is embedded in the structure of a β-hairpin motif [49]. Such a targeting signal was likely inherited from bacterial β-barrel proteins and remains conserved across all eukaryotic lineages, as supported by our experiment. As observed via protease protection assay, the topology of TvTom40-2 both in hydrogenosomes and mitochondria was similar. Interestingly, TvTom40-2 was able to very partially substitute yeast Tom40, indicating that at least some proteins were imported into yeast mitochondria through TvTom40-2. It is of note that some yeast mitochondrial proteins were imported into hydrogenosomes of *T. vaginalis* regardless of the presence or absence of NTS [26]. Based on this, it was proposed that the hydrogenosomal Tom40 is able to recognise unspecified ITSs conserved in the proteins of mitochondrial ancestry [26].

The key question is whether the TvTOM complex in hydrogenosomes consists of only core subunits or whether there any peripheral TOM subunit(s) that contribute to the import of proteins. This is expected because both NTS- and ITS-dependent protein targeting to hydrogenosomes have been demonstrated [24–26]. However, our HMM searches confirmed the absence of known TOM receptors Tom20 and Tom70 in excavates. These receptors either evolved only in certain eukaryotic lineages (Tom20) or were present in LECA (Tom70) as hypothesised here and by others [41]. To identify yet unknown peripheral TvTOM subunits, we performed proteomic analyses of the isolated TvTOM complex that indicated the presence of three TA proteins, in addition to Tom22-like protein. Two of them, Tom36 and Tom46, possess Hsp20-TPR-TMD architecture, which is similar to *T. brucei* receptor ATOM69. Indeed, we observed that Tom36 and Tom46 could bind to two hydrogenosomal preproteins, frataxin and αSCS, through binding assay. Tom36, Tom46, and ATOM69 are similar to yeast Tom70 with respect to the presence of TPR domains. The proximal TPR set of Tom70 interacts with Hsp90 [50] and may have an analogous function with the Hsp20 domain in Tom36, Tom46, and ATOM69 [11]. Of note, only Tom36 was tightly associated with TvTom40-2 and was detected

in both high–molecular-weight complexes, whereas Tom46 appears to be loosely associated because it appeared only in the 330 kDa complex. This is similar to the loose association of Tom70 with the TOM complex that was reported in *N. crassa* [20] and the absence of Tom70 in the 550 kDa TOM complex in *S. cerevisiae* [14]. The third protein, Homp19, is unique to *T. vaginalis*, and neither HHpred nor PfamA searches identified any known functional domains.

It is tempting to speculate that the subunits with similar Hsp20-TPR-TMD architecture in both *T. vaginalis* and *T. brucei* evolved from a common excavate ancestor. However, our phylogenetic profiling of Hsp20-TPR-TMD proteins revealed that they were present exclusively in parabasalids and kinetoplastids but absent in the basal lineages, *S. incarcerata* (Discoba), *Naegleria gruberi* (Discoba), and *C. membranifera* (Metamonada). Therefore, such a distribution is more consistent with independent gains in parabasalid and kinetoplastid lineages. This is also supported by our cluster analysis and phylogeny of TPR domains, in which Tom36/Tom46 and ATOM69 displayed a polyphyletic origin. This finding is interesting considering the recent phylogenetic studies that challenged the monophyletic origin of Excavata [35,36]. Although the phylogenetic analysis of Excavata—including long-branch members such as trichomonads—placed Metamonada as a sister group of Discoba, when long-branch representatives were excluded, these two groups separated [35]. Regardless of whether the origin of Excavata is monophyletic or polyphyletic, Tom36/Tom46 and ATOM69 most likely represent an example of convergent evolution rather than a diversification of a common ancestor.

In spite of the presence of Tom40 and Tom22 homologues, the hydrogenosomal TvTOM complex revealed considerable differences compared with the mitochondrial TOM complex. There are several constraints to be considered for the specific shaping of TvTOM. Hydrogenosomes are adapted to operate under anaerobic conditions, which resulted in a vast reduction of mitochondrial functions and, consequently, a reduction in the proteome from 1,000–1,500 proteins in mitochondria [51–53] to approximately 600 proteins in *T. vaginalis* hydrogenosomes [10,54]. In yeast, the positively charged NTS, forming an amphipathic α-helix, interacts with Tom20, the *cis* and *trans* domains of Tom22, and the presequence-binding groove of the Tim50 receptor during translocation across the OMM [55]. The positive charge of the NTS contributes to the membrane potential ($\Delta\psi$)-driven import step through TIM23 [56]. However, hydrogenosomes have lost the inner-membrane–associated respiratory chain that generates $\Delta\psi$, and this loss has possibly triggered the positive net charge of NTS to become dispensable. Indeed, most hydrogenosomal NTSs possess only a single positively charged residue [57], are considerably shorter, are not essential for preprotein import, and—in a number of matrix proteins—are not present. Thus, the import of these proteins is based on recognition of poorly understood ITSs [25,26,57]. These changes in the targeting signals are likely reflected by the modifications in TOM receptors, the loss of both Tom22 *trans* domain and Tim50, and the divergence of downstream import machinery [10]. Collectively, the adaptation to anaerobiosis and the loss of $\Delta\psi$ were critical constraints that may have allowed mutation, leading to the divergence of the TvTOM complex. Another reason for the divergence of TvTOM could be different evolutionary history of the lineage. Our finding of trichomonad Tom36 and Tom46 in Parabasalia and the phylogenomic profiling of TOM components supports the notion that the peripheral TOM subunits were added to the core components after the separation of the main eukaryotic lineages.

Materials and methods

Cell cultivation

T. vaginalis strain T1 (J. H. Tai, Institute of Biomedical Sciences, Taipei, Taiwan) and the recombinant strains were grown in Tryptone-Yeast extract-Maltose medium (TYM; pH 6.2)

with 10% (v/v) heat-inactivated horse serum, without or with 200 μg/mL Geneticin 418 (Single transfectant), or with both 200 μg/mL Geneticin 418 and 40 μg/mL Puromycin (Double transfectant) at 37 ˚C. Recombinant *E. coli* strains were grown on Luria-Bertani medium with 100 μg/mL of Ampicillin at 37 ˚C. The yeast strains were grown either in liquid medium (SD-Leucine or SLac-Leucine) or on solid medium (SD-Leucine or YPG) at 30 ˚C. For drop dilution assays, cells were cultured to an OD_{600} of 1.0 and diluted 5-fold, followed by spotting 5 μL of each dilution on SD-Leu, SD-Leu supplemented with 2 μg/mL Dox, YPG, or YPG supplemented with 2 μg/mL Dox.

Preparation of recombinant strains

The genes encoding TvTom40-2 (TVAG_332970) and Sam50 (TVAG_178100) were cloned into a pTagVag2 vector fused to a 2×HA tag at the C-terminus [58]. The genes encoding Tom36 (TVAG_277930), Tom46 (TVAG_137270), Homp38 (TVAG_190830), Homp19 (TVAG_283120), and Tom22-like protein (TVAG_076160) were cloned into a pTagVagV5 vector fused to a 2×V5 tag at the C-terminus [59]. The plasmids were transfected by electroporation [58] into either the WT strain or the strain expressing HA-tagged TvTom40-2. For studies in yeast, TvTom40-2 was cloned into a pYX142 vector (Novagen) fused to an HA tag at the C-terminus. The plasmid with no insert or plasmid encoding either HA-tagged TvTom40-2 or ScTom40 was transformed into yeast cells (WT strain W303α, tet-*TOM40*, *tom40*-25, and *tom40*-34) by lithium acetate method. The tet-*TOM40* yeast strain was constructed by inserting the tetracycline operator into the genome of WT strain, YMK120, upstream of *TOM40* ORF by homologous recombination, using an insertion cassette amplified from the plasmid pMK632 as described previously [60]. Yeast strains carrying temperature-sensitive alleles of *TOM40*, *tom40*-25, and *tom40*-34 were obtained from elsewhere [61]. The oligonucleotides used are listed in S4 Table.

Bioinformatics

Tom40-like protein sequences from *T. vaginalis* were searched against the NCBI Conserved Domains database and the *S. cerevisiae* proteome or against Protein Data Bank (PDB) using the HHpred tool [62]. A Tom40-specific HMM was built using the HMMER3 hmmbuild module [63], with a set of 24 well-annotated Tom40 sequences (S1 Data) and was scanned against the *T. vaginalis* protein database on the HMMER3 jackhmmer tool with the default settings [64]. Human Tom22 and Tom7 sequences were searched against the NCBI nonredundant protein database using three PSI–Basic Local Alignment Search Tool (BLAST) iterations from different eukaryotic organisms. The alignments for Tom22 and Tom7 were constructed using MAFFT [65] with 447 (S6 Data) and 349 (S7 Data) sequences, which were used to build Tom22-specific and Tom7-specific HMMs, respectively, and were searched against the Trichomonas proteome database (www.trichdb.org) using HMMER3 [64].

The homologues of 14 TOM subunits were searched against the predicted proteomes of selected eukaryotes using HHsearch. The query alignments and their sources are given in S8 Data. The best hits were then checked for conserved domains using HHpred (https://toolkit.tuebingen.mpg.de/#/tools/hhpred) and were searched against the NCBI nonredundant protein database using BLAST. The transmembrane helices were predicted using TMHMM server version 2.0 (http://www.cbs.dtu.dk/services/TMHMM/) with a relaxed cutoff of 0.3. For CLANS [43], an extensive data set of Tom36 and ATOM69 homologues was prepared. Tom36 and ATOM69 protein sequences were used as queries to search against the NCBI nonredundant protein database using PSI–BLAST with two iterations, and the sequences with an e-value less than 0.1 were selected. Altogether, 1,114 sequences were used for CLANS, which was run with

10,000 iterations. The obtained 2D clustering data were processed to color-code taxonomies. The TMD was predicted using TMHMM with a relaxed cutoff of 0.3. A subset of 418 sequences from the data set was selected for the phylogenetic analysis of their TPR domains. The TPR domains were detected using HHsearch with TPR domains from the COG database (COG0790) as a query. Multiple sequence alignment was created with MAFFT [65], and the alignment was trimmed with BMGE [66], which resulted in 179 sites. The phylogenetic tree was constructed with IQ-TREE [67] using the LG + I + G4 model and 10,000 ultra-fast bootstrap replicates.

Structural modeling

The model of TvTom40-2 was built using the *N. crassa* Tom40 structure (PDB ID 5o8o) as a template. The alignment was based on 140 Tom40 and VDAC sequences from a wide spectrum of eukaryotic organisms (S9 Data). The alignment was constructed by MAFFT, using the local pair alignment settings and 100 iterations [65] and later manually edited to reflect the secondary structure prediction of TvTom40-2 made by PSIPRED [68]. The three-dimensional (3D) structure model of TvTom40-2 was built using MODELLER 9v17 [69]. The quality of the final model was verified using ModFOLD 6 [70,71]. The electrostatic potential on the solvent-accessible surface of TvTom40-2 was calculated using APBS tool2 [72].

Subcellular fractionation, protease protection assay, alkaline carbonate extraction, and immunoblotting

Trichomonas cells from a 1 liter culture were harvested and homogenised by sonication, and the subcellular fractions were isolated by differential centrifugation, as described previously [10]. Isolated hydrogenosomes (protein concentration 1 mg/mL) carrying either HA-tagged or V5-tagged proteins were washed to remove protease inhibitors and incubated for 30 minutes at 37 ˚C in isolation buffer (225 mM sucrose, 10 mM KH_2PO_4, 20 mM HEPES, 0.5 mM KCl, 5 mM $MgCl_2$, and 1 mM EDTA [pH 7.2]) supplemented with either 100 μg/mL proteinase K enzyme (Roche Holding AG, Basel, Switzerland) or proteinase K with 0.5% Triton X-100. The incubation was terminated using 1 mM of phenylmethylsulfonyl fluoride (PMSF, Sigma Aldrich). Then, samples were analysed by immunoblotting using α-HA, α-V5, α-Fdx1, α-cytosolic malic enzyme, or α-αSCS antibody, followed by either α-mouse or α-rabbit antibody conjugated to peroxidase. The blot was developed using Amersham imager 600. Subcellular fractionation for yeast strains, and alkaline carbonate extraction and protease protection assay with isolated mitochondria were performed as described previously [73]. Proteins were separated by SDS-PAGE; immunoblotted with α-HA, α-HK, α-Fis1, or α-Aco antibody; and developed using an ECL system.

Immunofluorescence and STED microscopy

The cells for immunofluorescence microscopy were processed as previously described [74]. Recombinant proteins were visualised using mouse α-HA and rabbit α-V5 antibodies, and Alexa Fluor 488 donkey α-mouse and Alexa Fluor 594 donkey α-rabbit antibodies (Thermo Fisher Scientific). The hydrogenosomal marker malic enzyme was detected by rabbit polyclonal antibody. The slides were mounted using Vectashield containing DAPI (4',6-diamidino-2-phenylindole) (Vector laboratories). The cells were examined with an Olympus Cell-R IX-81 microscope, and the images were processed using ImageJ. For STED, Abberior STAR 580 α-mouse and Abberior STAR 635p α-rabbit antibodies, along with Abberior TDE mounting medium, were used. STED images were acquired on a commercial Abberior STED 775 QUAD Scanning microscope (Abberior Instruments) equipped with a Nikon CFI Plan Apo

Lambda objective (60× Oil, NA 1.40). Abberior STAR580- and STAR 635P-labeled proteins were illuminated by pulsed 561 nm and 640 nm lasers and depleted by a pulsed 775 nm STED depletion laser of the 2D donut. Fluorescence signal was filtered (Emission bandpasses: 605–625 nm and 650–720 nm; pinhole 40 μm) and detected on single photon counting modules, with time gates set to 0.8–8.8 ns. Images were scanned with a pixel size of 20 nm × 20 nm, with a 10 μs dwell time and in-line interleaved acquisition mode using the Imspector software. All images were deconvolved with Huygens Professional version software 17.04 using the Classic Maximum Likelihood Estimation algorithm.

BN-PAGE

Isolated hydrogenosomes from the recombinant strains expressing tagged proteins were lysed with the native sample buffer (Life Technologies) containing either varying concentrations (1%–3%) of digitonin or 1% digitonin. The clarified extracts were electrophoresed on 3%–12% or 4%–16% NativePAGE bis-tris gel (Thermo Fisher Scientific), immunoblotted with either α-HA or α-V5 antibody, and developed by chemiluminescence. For BN-PAGE with yeast cells, isolated mitochondria from the strain with empty plasmid, or from strain expressing HA-tagged TvTom40-2, were lysed with lysis buffer containing 1% digitonin, and the clarified samples were electrophoresed on a 6%–13% native gel, immunoblotted with either α-HA or α-ScTom40 antibody, and developed using an ECL system.

Crosslinking and native coIP

CoIPs were performed for the HA-tagged TvTom40-2 either with or without crosslinker using isolated hydrogenosomes from both WT and recombinant strains. For crosslinking, interacting proteins in hydrogenosomes (protein concentration 1 mg/mL) were crosslinked with 1 mM DSP (dithiobis(succinimidyl propionate); Thermo Scientific) for 30 minutes at 25 ˚C, excess DSP was quenched with 50 mM Tris (pH 7.5), and the hydrogenosomes were washed twice with isolation buffer. For coIP, the hydrogenosomes (protein concentration 1 mg/mL) were solubilised in MKG buffer (10 mM MOPS [3-(N-morpholino)propanesulfonic acid; pH 7], 50 mM potassium acetate, 10% glycerol, and EDTA-free cOmplete protease inhibitor cocktail [Roche]) containing 1% digitonin (Merck Millipore), and the clarified extract was incubated with Dynabeads (Thermo Fisher Scientific) coupled with α-HA antibody for 90 minutes on an overhead rotator at room temperature. The beads were washed thrice before elution with either SDS-PAGE buffer for crosslinking coIPs or elution buffer (MKG buffer with 0.25% digitonin and 1 mg/mL HA peptide, Thermo Fisher Scientific) for native coIPs. The coupling of α-HA antibody to the Dynabeads was performed according to the manufacturer's instructions.

LFQ-MS analysis

LFQ-MS was performed according to standard procedures as described previously [59]. To remove SDS from the crosslinking coIP eluates and to remove HA peptides from the native coIP eluates, samples were resuspended in 8 M urea and processed using a Filter Aided Sample Preparation (FASP) protocol, according to Wisniewski et al. [75]. The samples were digested with trypsin and the peptides obtained were subjected to liquid chromatography-MS. The MS/MS spectra obtained were searched against the *T*. *vaginalis* database (downloaded from Trichomonas Genome Resource [TrichDB; www.trichdb.org] containing 59,862 entries), the quantifications were performed with the label-free algorithms, and the data analysis was performed using Perseus 1.5.2.4 software. The MS data have been deposited to the ProteomeXchange

consortium via the PRIDE [76] partner repository. The MS data were obtained from four independent coIP experiments for each immunoprecipitated protein.

Isolation of the TvTOM complex, transmission electron microscopy, and data analysis

The TvTOM complex was purified under native conditions from hydrogenosomes isolated from the recombinant strain expressing C-terminal HA-tagged TvTom40-2 as described earlier. Five microliters of purified TvTOM complexes in solution was applied to copper electron microscopy grids (EMS200-Cu) covered with a 20 nm carbon film, which were glow discharged for 40 seconds with a 5 mA current prior to specimen application. Excess sample was removed after 1 minute by blotting (Whatman no. 1 filter paper) for 1 to 2 seconds, and the grid was immediately stained with 5 μL of 2% phosphotungstic acid for 1 minute 40 seconds and blotted to remove excess stain. A large data set of optimised, negatively stained specimen grids was acquired with a Tecnai F20 microscope (Thermo Fisher Scientific) operating at an accelerating voltage of 200 kV, with a FEI Eagle 4K CCD camera, at a magnification of 78,000× and a pixel size of 1.79 Å. Altogether, 1,000 images were acquired with defocus ranging from 2 to 5 μm. After quality inspection and determination of Contrast Transfer Function (CTF) parameters with the GCTF program [77], 650 micrographs were subjected to particle picking. Approximately 6,000 particles were manually picked from the first 200 micrographs with the e2boxer.py routine of the EMAN2 program [78] and subjected to three rounds of class averaging in Relion 1.4 [79], with 200, 150, and 100 classes, respectively. The box size was set to 192 pixels to accommodate higher-order multimers. This analysis resulted in a set of three representative class averages, which were low-pass filtered to 30 Å and used as templates for automated particle selection of the preselected set of 650 micrographs with the Gautomatch program. Altogether, 71,834 identified particles were subjected to five rounds of 2D classification in Relion with 200 classes, which reduced the data set to 10,038 particles. All 2D classifications comprised 40 iterations. The presented resolution of the class averages corresponds to the lowest SSNR value $\geq$1 indicated in the *model.star file resulting from the last iteration of the final 2D classification. The number of particles contributing to the class averages was also found in the *model.star files.

In vitro protein import assay

The gene encoding Ferredoxin1 (TVAG_003900) was cloned into NEB PURExpress control vector fused to the DHFR gene (*E. coli*) at the C-terminus. Radiolabeled TvFdx1-DHFR was synthesised in vitro in the presence of L-[^{35}S] methionine (MGP spol sro) according to the manufacturer's instructions (NEB PURExpress in vitro protein synthesis kit). Cytoplasmic extract was prepared from the *T. vaginalis* strain T1 as described elsewhere [24]. For the time course experiment, the import assay was conducted in a 500 μL reaction volume, and the mixture contained 500 μg of hydrogenosomes (protein concentration) carrying both TvTom40-2-HA and Tom36-V5, import buffer (250 mM sucrose, 10 mM MOPS-KOH [pH 7.2], 3% BSA, 80 mM KCl, 7 mM $MgCl_2$, and 10 mM ATP), 125 μL cytosolic extract, and 25 μL radiolabeled precursors at 37 ˚C. At each time point, 100 μL was removed and shifted to ice, and the hydrogenosomes were re-isolated and washed twice with import buffer. For the import-arrest experiment, the import assay was performed either in the presence or absence of 10 μM methotrexate (Sigma Aldrich) and 1 mM NADPH. Wherever indicated, the hydrogenosomes were treated with 50 μg/mL of proteinase K. For the import-arrest and coIP assay, the import assay was performed either in the presence or absence of 10 μM methotrexate, the hydrogenosomes obtained were subjected to crosslinking, and the HA-tagged protein was immunoprecipitated as described earlier except that 0.5% Triton X-100 was used to lyse the organelles instead of

digitonin. The samples were electrophoresed, and the gel was vacuum dried. The gel was exposed for 4 to 5 days prior to phosphorimaging with Typhoon TLA 7000 scanner.

Protein expression, pull-down, and binding assay

The gene encoding for the cytosolic domain of Tom36 and Tom46 (Tom36cd and Tom46cd) were cloned into pET42b vector tagged to polyhistidine at the C-terminus. The genes encoding for cytochrome b5 (TVAG_063210), frataxin (TVAG_182610), and αSCS (TVAG_165340; αSCS was fused to DHFR to the C-terminus) were subcloned into NEB PURExpress control plasmid, and the radiolabeled precursors were synthesised in the presence of L-[^{35}S] methionine as described earlier. The recombinant His-tagged proteins were expressed in *E. coli* BL21 (DE3) strain at 37 ˚C for 3 hours following the induction with 0.5 mM IPTG. The cells from a 10 mL culture of *E. coli* (negative control) and strains expressing His-tagged proteins were harvested, resuspended in 4.5 mL lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, 1 mg/mL lysozyme, and EDTA-free cOmplete protease inhibitor cocktail), incubated on ice for 45 minutes, and lysed using QSonica sonicator. The homogenised extract was clarified at 9,000 rcf for 30 minutes at 4 ˚C. Aliquots of supernatant and pellet were used for immunoblotting to test the solubility of the proteins. The supernatant obtained was split into three equal parts and was incubated with 50 μL of Ni-NTA agarose resin (Qiagen) on an overhead rotator for 2 hours at room temperature. The resin collected was washed five times using 10 volumes of wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, and EDTA-free cOmplete protease inhibitor cocktail). To block, the beads were washed thrice with wash buffer II (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, 3% BSA, and EDTA-free cOmplete protease inhibitor cocktail). To the mock-treated beads or beads bound with His-tagged protein, binding buffer (50 mM Tris, 150 mM NaCl [pH 7.4]), 50 μL of *Trichomonas* cytosolic extract, and 10 μL of radiolabeled precursors were added and incubated for 1 hour at 37 ˚C with gentle shaking. The beads were washed three times with the binding buffer, and the proteins were eluted with the elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 500 mM imidazole, and EDTA-free cOmplete protease inhibitor cocktail). The samples were electrophoresed, and the gel was vacuum dried. The gel was exposed for 4 to 5 days prior to phosphorimaging with Typhoon TLA 7000. The oligonucleotides used for cloning are listed in S4 Table.

Antibody production

The gene encoding Sam50 was cloned into pET42b fused to a C-terminal His tag. The protein was expressed in *E. coli* BL21 (DE3) strain following an induction with 1 mM IPTG, and the His-tagged Sam50 was purified using affinity chromatography under denaturing conditions. The purified antigen was separated via SDS-PAGE, and the Coomassie-stained band was used to generate polyclonal antibody in rat.

Supporting information

S1 Fig. Comparison of the conserved β-motif of TvTom40-like proteins (TvTom40-1-7) with Tom40s and VDACs of other eukaryotes. The conserved residues of the β-motif, PxGxxHxH, are highlighted: P is polar (fluorescent green), x is any amino acid, G is glycine (fluorescent yellow), and H is hydrophobic (turquoise). All TvTom40 isoforms have the conserved β-motif except TvTom40-3, where the last hydrophobic residue has been replaced by serine. TOM, translocase of the outer membrane; TvTOM, *T. vaginalis* TOM; VDAC, voltage-dependent anion channel.
(PDF)

S2 Fig. Conservation of TOM complex-forming residues. Highlighted residues mark anchoring positions for possible interactions between the Tom40 β-barrel and essential subunits of the TOM complex in *S. cerevisiae*. The selected 21 sequences were chosen out of the multiple alignment of 140 sequences to demonstrate the potential conservation of key residues and to highlight the differences between VDAC and Tom40 proteins. TOM, translocase of the outer membrane; TvTOM, *T. vaginalis* TOM; VDAC, voltage-dependent anion channel.
(PDF)

S3 Fig. Sequence alignment of Tom22-like protein from *T. vaginalis* against Tom22 from other eukaryotes. Names of the organisms are as follows: *T. vaginalis*, *S. cerevisiae*, *Arabidopsis thaliana*, *Rattus rattus*, *Mus musculus*, *Homo sapiens*, and *Bos taurus*. The TMD is marked by a box, and the conserved residues are highlighted—tryptophan (yellow), hydroxylated residues (turquoise), and proline (green). TMD, transmembrane domain; Tom, translocase of the outer membrane.
(PDF)

S4 Fig. Expression of His-tagged Tom36cd and Tom46cd in *E. coli* BL21 (DE3) strains. (A, B) Expression of Tom36cd-His and Tom46cd-His. SDS-PAGE gels stained with Coomassie and immunoblots probed with α-His antibody for the whole cell lysate from a 250 μL culture of *E. coli* strain expressing Tom36cd-His (panel A) and Tom46cd-His (panel B), respectively, before (Control) and 1 hour and 3 hours after induction with 0.5 mM IPTG. IPTG, Isopropyl β-D-1-thiogalactopyranoside; SDS-PAGE, sodium dodecyl sulphate-PAGE; Tom, translocase of the outer membrane.
(TIF)

S5 Fig. EM analysis of the isolated TvTOM complex. (A, B) Preparation of purified TvTOM for EM analysis. (A) Immunoblot of digitonin-lysed extract of hydrogenosomes (Input; 5%) and the eluate (IP, 2.5%) from TvTom40-2-HA IP under native conditions using α-HA antibody. (B) Silver stained-gel showing the α-HA IP eluates from TvT1 WT strain and *Trichomonas* strain expressing HA-tagged TvTom40-2. Two bands marked were identified by MS as TvTom40-2. The common contaminant was identified as Cpn60. (C) Purified TvTOM complexes were applied on EM grids and negatively stained with phosphotungstic acid. Electron micrograph of negatively stained TvTOM complexes recorded at a magnification of 78,000×. Scale bar, 40 nm. Bottom panel: magnified view of selected particles with three, two, and one pore(s) (left to right). Scale bar, 10 nm. EM, electron microscopy; HA, human influenza hemagglutinin; In, Input; IP, immunoprecipitation; MS, mass spectrometry; TOM, translocase of the outer membrane; TvTOM, *T. vaginalis* TOM; WT, wild-type.
(PDF)

S6 Fig. Enlarged version of the phylogenetic tree shown in Fig 11B.
(PDF)

S1 Table. HHpred search with each TvTom40 homologue against the NCBI conserved domains database (version 3.16) and *S. cerevisiae* proteome. NCBI, National Center for Biotechnology Information; TOM, translocase of the outer membrane; TvTom, *T. vaginalis* TOM.
(PDF)

S2 Table. Pairwise comparison of HMM profiles for the seven TvTom40 homologues against PDB database using the HHpred tool. HMM, hidden Markov model; PDB, Protein Data Bank; TOM, translocase of the outer membrane; TvTOM, *T. vaginalis* TOM.
(PDF)

S3 Table. TOM subunit orthologues identified in selected eukaryotic lineages. TOM, translocase of the outer membrane.
(XLSX)

S4 Table. List of oligonucleotides.
(PDF)

S1 Data. A list of 24 well-annotated Tom40 sequences that were used to build Tom40 HMM. HMM, hidden Markov model; TOM, translocase of the outer membrane.
(TXT)

S2 Data. A data set of proteins identified from TvTom40-2-HA, Tom36-HA, and Sam50-HA coIPs both under crosslinking and native conditions using LFQ-MS analysis. The data sets shown were obtained were four independent coIP experiments indicated by columns A, B, C, and D. A protein was considered enriched either if the protein was present only in the test sample and absent in the control or if the protein was enriched by a fold change of >1 in the test sample. Following are the column headings: accession number (protein ID on NCBI protein database or TrichDB), protein name, molecular weight of the protein, sequence coverage (percentage coverage of the peptide sequence to the full length protein sequence), peptides (number of peptides identified for a particular protein), unique peptides (number of unique peptides identified for a particular protein), score from the MS identification, intensity of the MS, MS/MS count. (A–D) Intensity from four independent IP experiments in binary logarithmic values; mean: arithmetic mean of intensity from four independent (A–D) IP experiments in binary logarithmic values; n: difference between mean of the test and the control samples; and fold change: actual change in the protein levels between the test and the control samples. coIP, co-immunoprecipitation; HA, human influenza hemagglutinin; LFQ-MS, label-free quantitative mass spectrometry; NCBI, National Center for Biotechnology Information; Sam, sorting and assembly machinery; TOM, translocase of the outer membrane; TrichDB, Trichomonas Genome Resource.
(XLSX)

S3 Data. Protein sequences of the TOM subunit orthologues listed in S3 Table. TOM, translocase of the outer membrane.
(FASTA)

S4 Data. A set of 1,114 proteins with their coordinates used for CLANS that were obtained from two iterations of PSI–BLAST with Tom36 and ATOM69 as queries. ATOM, archaic translocase of the outer membrane; CLANS, cluster analysis of sequences; TOM, translocase of the outer membrane.
(FASTA)

S5 Data. An alignment of 418 TPR proteins from CLANS that were selected for the phylogenetic analysis. CLANS, cluster analysis of sequences; TPR, tetratricopeptide repeat.
(FASTA)

S6 Data. A list of 447 Tom22 sequences that were used to build Tom22 HMM. HMM, hidden Markov model; Tom, translocase of the outer membrane.
(TXT)

S7 Data. A list of 349 Tom7 sequences that were used to build Tom7 HMM. HMM, hidden Markov model; Tom, translocase of the outer membrane.
(TXT)

S8 Data. Sequence alignments for TOM subunits that were used to identify orthologues in different eukaryotic lineages. Alignments of ATOM11, ATOM12, ATOM46, and ATOM69 homologues from kinetoplastids, Tom60 homologues from *Entamoeba* sp., and Tom36 homologues from parabasalids using MAFFT; Tom40 and VDAC (Porin_3) homologues, fungal Tom5, metazoan Tom5, plant Tom5, metazoan Tom6, fungal Tom6, Tom7, Tom20, plant Tom20, and Tom22 homologues from the Pfam database; plant Tom6 homologues from the Eggnog database; and Tom70 homologues from the COG database. ATOM, archaic TOM; COG, clusters of orthologous groups; MAFFT, multiple sequence alignment based on fast Fourier transform; Pfam, Protein families; Tom, translocase of the outer membrane.
(TXT)

S9 Data. A list of Tom40 and VDAC sequences that were used for TvTom40-2 modelling. Tom, translocase of the outer membrane; TvTom, *T. vaginalis* TOM; VDAC, voltage-dependent anion channel.
(TXT)

Acknowledgments

We thank Karel Harant and Pavel Talacko from the Laboratory of Mass Spectrometry, BIOCEV. We acknowledge the core facility Cryo-Electron Microscopy and Tomography, CEITEC and the IMCF, BIOCEV.

Author Contributions

Conceptualization: Abhijith Makki, Petr Rada, Jan Tachezy.

Data curation: Abhijith Makki, Sami Kereïche.

Formal analysis: Vojtěch Žárský, Sami Kereïche, Lubomír Kováčik, Marian Novotný, Tobias Jores.

Funding acquisition: Abhijith Makki, Jan Tachezy.

Investigation: Abhijith Makki, Petr Rada, Tobias Jores.

Methodology: Abhijith Makki, Petr Rada.

Resources: Doron Rapaport, Jan Tachezy.

Software: Vojtěch Žárský.

Supervision: Doron Rapaport, Jan Tachezy.

Visualization: Abhijith Makki, Vojtěch Žárský.

Writing – original draft: Abhijith Makki, Jan Tachezy.

Writing – review & editing: Abhijith Makki, Jan Tachezy.

References

1. Martijn J, Vosseberg J, Guy L, Offre P, Ettema TJG. Deep mitochondrial origin outside the sampled alphaproteobacteria. Nature. 2018; 557: 101–105. https://doi.org/10.1038/s41586-018-0059-5 PMID: 29695865
2. Timmis JN, Ayliff MA, Huang CY, Martin W. Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. Nat Rev Genet. 2004; 5: 123–135. https://doi.org/10.1038/nrg1271 PMID: 14735123
3. Dolezal P, Likic V, Tachezy J, Lithgow T. Evolution of the molecular machines for protein import into mitochondria. Science (80-). 2006; 313: 314–318. https://doi.org/10.1126/science.1127895 PMID: 16857931

4. Maćašev D, Whelan J, Newbigin E, Silva-Filho MC, Mulhern TD, Lithgow T. Tom22′, an 8-kDa transsite receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes. Mol Biol Evol. 2004; 21: 1557–1564. https://doi.org/10.1093/molbev/msh166 PMID: 15155803
5. Söllner T, Griffiths G, Pfaller R, Pfanner N, Neupert W. MOM19, an import receptor for mitochondrial precursor proteins. Cell. 1989; 59: 1061–1070. PMID: 2557158
6. Söllner T, Pfaller R, Griffiths G, Pfanner N, Neupert W. A mitochondrial import receptor for the ADP/ATP carrier. Cell. 1990; 62: 107–115. https://doi.org/10.1016/0092-8674(90)90244-9 PMID: 2163763
7. Perry AJ, Hulett JM, Likić VA, Lithgow T, Gooley PR. Convergent evolution of receptors for protein import into mitochondria. Curr Biol. 2006; 16: 221–229. https://doi.org/10.1016/j.cub.2005.12.034 PMID: 16461275
8. Wojtkowska M, Buczek D, Stobienia O, Karachitos A, Antoniewicz M, Slocinska M, et al. The TOM complex of Amoebozoans: The cases of the amoeba *Acanthamoeba castellanii* and the slime mold *Dictyostelium discoideum*. Protist. Elsevier GmbH.; 2015; 166: 349–362. https://doi.org/10.1016/j.protis.2015.05.005 PMID: 26074248
9. Makiuchi T, Mi-ichi F, Nakada-Tsukui K, Nozaki T. Novel TPR-containing subunit of TOM complex functions as cytosolic receptor for Entamoeba mitosomal transport. Sci Rep. 2013; 3: 1–7. https://doi.org/10.1038/srep01129 PMID: 23350036
10. Rada P, Doležal P, Jedelský PL, Bursac D, Perry AJ, Šedinová M, et al. The core components of organelle biogenesis and membrane transport in the hydrogenosomes of *Trichomonas vaginalis*. PLoS ONE. 2011; 6(9). e24428. https://doi.org/10.1371/journal.pone.0024428 PMID: 21935410
11. Mani J, Desy S, Niemann M, Chanfon A, Oeljeklaus S, Pusnik M, et al. Mitochondrial protein import receptors in kinetoplastids reveal convergent evolution over large phylogenetic distances. Nat Commun. 2015; 6: 6646. https://doi.org/10.1038/ncomms7646 PMID: 25808593
12. Fukasawa Y, Oda T, Tomii K, Imai K. Origin and evolutionary alteration of the mitochondrial import system in eukaryotic lineages. Mol Biol Evol. 2017; 34: 1574–1586. https://doi.org/10.1093/molbev/msx096 PMID: 28369657
13. Pusnik M, Schmidt O, Perry AJ, Oeljeklaus S, Niemann M, Warscheid B, et al. Mitochondrial preprotein translocase of trypanosomatids has a bacterial origin. Curr Biol. 2011; 21: 1738–1743. https://doi.org/10.1016/j.cub.2011.08.060 PMID: 22000100
14. Model K, Meisinger C, Kühlbrandt W. Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase. J Mol Biol. Elsevier Ltd; 2008; 383: 1049–1057. https://doi.org/10.1016/j.jmb.2008.07.087 PMID: 18706915
15. Bausewein T, Mills DJ, Langer JD, Nitschke B, Nussberger S, Kühlbrandt W. Cryo-EM structure of the TOM core complex from *Neurospora crassa*. Cell. Elsevier; 2017; 170: 693–700. https://doi.org/10.1016/j.cell.2017.07.012 PMID: 28802041
16. Shiota T, Imai K, Qiu J, Hewitt VL, Tan K, Shen H-H, et al. Molecular architecture of the active mitochondrial protein gate. Science (80-). 2015; 349: 1544–1548. https://doi.org/10.1126/science.aac6428 PMID: 26404837
17. Qiu J, Wenz L-S, Zerbes RM, Oeljeklaus S, Bohnert M, Stroud DA, et al. Coupling of mitochondrial import and export translocases by receptor-mediated supercomplex formation. Cell. Elsevier; 2013; 154: 596–608. https://doi.org/10.1016/j.cell.2013.06.033 PMID: 23911324
18. Wenz L-S, Ellenrieder L, Qiu J, Bohnert M, Zufall N, van der Laan M, et al. Sam37 is crucial for formation of the mitochondrial TOM–SAM supercomplex, thereby promoting β-barrel biogenesis. J Cell Biol. 2015; 210: 1047–1054. Available from: http://jcb.rupress.org/content/210/7/1047.abstract https://doi.org/10.1083/jcb.201504119 PMID: 26416958
19. Shiota T, Mabuchi H, Tanaka-Yamano S, Yamano K, Endo T. In vivo protein-interaction mapping of a mitochondrial translocator protein Tom22 at work. Proc Natl Acad Sci. 2011; 108: 15179–15183. https://doi.org/10.1073/pnas.1105921108 PMID: 21896724
20. Künkele K, Heins S, Dembowski M, Nargang FE, Benz R, Thieffry M, et al. The preprotein translocation channel of the outer membrane of mitochondria. Cell. 1998; 93: 1009–1019. https://doi.org/10.1016/S0092-8674(00)81206-4 PMID: 9635430
21. Embley TM, Martin W. Eukaryotic evolution, changes and challenges. Nature. 2006; 440: 623–630. https://doi.org/10.1038/nature04546 PMID: 16572163
22. Hrdy I, Tachezy J, Muller M. Metabolism of trichomonad hydrogenosomes. In: Tachezy J, editor. Hydrogenosomes and Mitosomes:Mitochondria of Anaerobic Euakryotes. Berlin, Heidelberg: Springer-Verlag; 2008. pp. 114–145.

23. Clemens DL, Johnson PJ. Failure to detect DNA in hydrogenosomes of *Trichomonas vaginalis* by nick translation and immunomicroscopy. Mol Biochem Parasitol. 2000; 106: 307–313. https://doi.org/10.1016/S0166-6851(99)00220-0 PMID: 10699261

24. Bradley PJ, Lahti CJ, Plumper E, Johnson PJ. Targeting and translocation of proteins into the hydrogenosome of the protist Trichomonas: Similarities with mitochondrial protein import. EMBO J. 1997; 16: 3484–3493. https://doi.org/10.1093/emboj/16.12.3484 PMID: 9218791

25. Rada P, Makki A, Zimorski V, Garg S, Hampl V, Hrdý I, et al. N-terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*. J Eukaryot Cell. 2015; 14: 1264–1275.

26. Garg S, Stölting J, Zimorski V, Rada P, Tachezy J, Martin WF, et al. Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. Genome Biol Evol. 2015; 7: 2716–2726. https://doi.org/10.1093/gbe/evv175 PMID: 26338186

27. Dyall SD, Lester DC, Schneider RE, Delgadillo-Correa MG, Plümper E, Martinez A, et al. *Trichomonas vaginalis* Hmp35, a putative pore-forming hydrogenosomal membrane protein, can form a complex in yeast mitochondria. J Biol Chem. 2003; 278: 30548–30561. https://doi.org/10.1074/jbc.M304032200 PMID: 12766161

28. Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, et al. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science (80-). 2007; 315: 207–212. https://doi.org/10.1126/science.1132894 PMID: 17218520

29. Wiedemann N, Pfanner N. Mitochondrial machineries for protein import and assembly. Annu Rev Biochem. Annual Reviews; 2017; 86: 685–714. https://doi.org/10.1146/annurev-biochem-060815-014352 PMID: 28301740

30. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterProScan 5: Genome-scale protein function classification. Bioinformatics. 2014; 30: 1236–1240. https://doi.org/10.1093/bioinformatics/btu031 PMID: 24451626

31. van Dooren GG, Yeoh LM, Striepen B, McFadden GI. The import of proteins into the mitochondrion of *Toxoplasma gondii*. J Biol Chem. 2016; 291: 19335–19350. https://doi.org/10.1074/jbc.M116.725069 PMID: 27458014

32. Mani J, Rout S, Desy S, Schneider A. Mitochondrial protein import—Functional analysis of the highly diverged Tom22 orthologue of *Trypanosoma brucei*. Sci Rep. Nature Publishing Group; 2017; 7: 40738. https://doi.org/10.1038/srep40738 PMID: 28094338

33. Eilers M, Schatz G. Binding of a specific ligand inhibits import of a purified precursor protein into mitochondria. Nature. 1986; 322: 228–232. https://doi.org/10.1038/322228a0 PMID: 3016548

34. Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AGB, et al. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups." Proc Natl Acad Sci. 2009; 106: 3859–3864. Available from: http://www.pnas.org/content/106/10/3859.abstract https://doi.org/10.1073/pnas.0807880106 PMID: 19237557

35. Heiss AA, Kolisko M, Ekelund F, Brown MW, Roger AJ, Simpson AGB. Combined morphological and phylogenomic re-examination of malawimonads, a critical taxon for inferring the evolutionary history of eukaryotes. R Soc Open Sci. The Royal Society Publishing; 2018; 5: 171707. https://doi.org/10.1098/rsos.171707 PMID: 29765641

36. Derelle R, Torruella G, Klimeš V, Brinkmann H, Kim E, Vlček Č, et al. Bacterial proteins pinpoint a single eukaryotic root. Proc Natl Acad Sci. 2015; 112: E693–E699. Available from: http://www.pnas.org/content/112/7/E693.abstract https://doi.org/10.1073/pnas.1420657112 PMID: 25646484

37. Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petrželková R, Eme L, et al. A eukaryote without a mitochondrial organelle. Curr Biol. Elsevier; 2016; 26: 1274–1284. https://doi.org/10.1016/j.cub.2016.03.053 PMID: 27185558

38. Leger MM, Kolisko M, Kamikawa R, Stairs CW, Kume K, Čepička I, et al. Organelles that illuminate the origins of Trichomonas hydrogenosomes and Giardia mitosomes. Nat Ecol & Evol. Macmillan Publishers Limited, part of Springer Nature.; 2017; 1: 92. Available from: http://dx.doi.org/10.1038/s41559-017-0092

39. Ebenezer TE, Carrington M, Lebert M, Kelly S, Field MC. Euglena gracilis genome and transcriptome: Organelles, nuclear genome assembly strategies and initial features. In: Schwartzbach SD, Shigeoka S, editors. Euglena: Biochemistry, Cell and Molecular Biology. Cham: Springer International Publishing; 2017. pp. 125–140. https://doi.org/10.1007/978-3-319-54910-1_7 PMID: 28429320

40. Leger MM, Eme L, Hug LA, Roger AJ. Novel hydrogenosomes in the microaerophilic jakobid *Stygiella incarcerata*. Mol Biol Evol. Oxford University Press; 2016; 33: 2318–2336. https://doi.org/10.1093/molbev/msw103 PMID: 27280585

41. Tsaousis AD, Gaston D, Stechmann A, Walker PB, Lithgow T, Roger AJ. A functional Tom70 in the human parasite Blastocystis sp.: Implications for the evolution of the mitochondrial import apparatus. Mol Biol Evol. 2011; 28: 781–791. https://doi.org/10.1093/molbev/msq252 PMID: 20871025

42. Brown MW, Sharpe SC, Silberman JD, Heiss AA, Lang BF, Simpson AGB, et al. Phylogenomics demonstrates that breviate flagellates are related to opisthokonts and apusomonads. Proc R Soc B Biol Sci. 2013; 280. Available from: http://rspb.royalsocietypublishing.org/content/280/1769/20131755.abstract

43. Frickey T, Lupas A. CLANS: a Java application for visualizing protein families based on pairwise similarity. Bioinformatics. 2004; 20: 3702–3704. Available from: http://dx.doi.org/10.1093/bioinformatics/bth444 PMID: 15284097

44. Ahting U, Thun C, Hegerl R, Typke D, Nargang FE, Neupert W, et al. The TOM core complex: The general protein import pore of the outer membrane of mitochondria. J Cell Biol. 1999; 147: 959–968. https://doi.org/10.1083/jcb.147.5.959 PMID: 10579717

45. Klein A, Israel L, Lackey SWK, Nargang FE, Imhof A, Baumeister W, et al. Characterization of the insertase for β-barrel proteins of the outer mitochondrial membrane. J Cell Biol. 2012; 199: 599–611. Available from: http://jcb.rupress.org/content/199/4/599.abstract https://doi.org/10.1083/jcb.201207161 PMID: 23128244

46. Becker T, Wenz L-S, Thornton N, Stroud D, Meisinger C, Wiedemann N, et al. Biogenesis of mitochondria: Dual dole of Tom7 in modulating assembly of the preprotein translocase of the outer membrane. J Mol Biol. 2011; 405: 113–124. https://doi.org/10.1016/j.jmb.2010.11.002 PMID: 21059357

47. Zarsky V, Tachezy J, Dolezal P. Tom40 is likely common to all mitochondria. Curr Biol. 2012; 22: R479–R481. https://doi.org/10.1016/j.cub.2012.03.057 PMID: 22720677

48. Kinoshita J, Mihara K, Oka T. Identification and characterization of a new tom40 isoform, a central component of mitochondrial outer membrane translocase. J Biochem. 2007; 141: 897–906. Available from: http://dx.doi.org/10.1093/jb/mvm097 PMID: 17437969

49. Jores T, Klinger A, Groß LE, Kawano S, Flinner N, Duchardt-Ferner E, et al. Characterization of the targeting signal in mitochondrial β-barrel proteins. Nat Commun. The Author(s); 2016; 7: 12036. Available from: http://dx.doi.org/10.1038/ncomms12036

50. Young JC, Hoogenraad NJ, Hartl FU. Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell. Elsevier; 2003; 112: 41–50. https://doi.org/10.1016/S0092-8674(02)01250-3

51. Sickmann A, Reinders J, Wagner Y, Joppich C, Zahedi R, Meyer HE, et al. The proteome of *Saccharomyces cerevisiae* mitochondria. Proc Natl Acad Sci. 2003; 100: 13207–13212. Available from: http://www.pnas.org/content/100/23/13207.abstract https://doi.org/10.1073/pnas.2135385100 PMID: 14576278

52. Panigrahi AK, Ogata Y, Zíková A, Anupama A, Dalley RA, Acestor N, et al. A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome. Proteomics. WILEY-VCH Verlag; 2009; 9: 434–450. https://doi.org/10.1002/pmic.200800477 PMID: 19105172

53. Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, Ong S-E, et al. A mitochondrial protein compendium elucidates complex I disease biology. Cell. Elsevier; 2008; 134: 112–123. https://doi.org/10.1016/j.cell.2008.06.016 PMID: 18614015

54. Schneider RE, Brown MT, Shiflett AM, Dyall SD, Hayes RD, Xie Y, et al. The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. Int J Parasitol. 2011; 41: 1421–1434. https://doi.org/10.1016/j.ijpara.2011.10.001 PMID: 22079833

55. Li J, Sha B. The structure of Tim50(164–361) suggests the mechanism by which Tim50 receives mitochondrial presequences. Acta Crystallogr Sect F, Struct Biol Commun. United States; 2015; 71: 1146–1151. https://doi.org/10.1107/S2053230X15013102 PMID: 26323300

56. Geissler A, Krimmer T, Bömer U, Guiard B, Rassow J, Pfanner N. Membrane potential-driven protein import into mitochondria: The sorting sequence of cytochrome b(2) modulates the Δψ-dependence of translocation of the matrix-targeting sequence. Fox TD, editor. Mol Biol Cell. The American Society for Cell Biology; 2000; 11: 3977–3991. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC15051/

57. Šmíd O, Matušková A, Harris SR, Kučera T, Novotný M, Horváthová L, et al. Reductive evolution of the mitochondrial processing peptidases of the unicellular parasites *Trichomonas vaginalis* and *Giardia intestinalis*. PLoS Pathog. 2008; 4(12): e1000243. Available from: https://doi.org/10.1371/journal.ppat.1000243

58. Hrdy I, Hirt RP, Dolezal P, Bardonová L, Foster PG, Tachezy J, et al. Trichomonas hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. Nature. 2004; 432: 618–622. https://doi.org/10.1038/nature03149 PMID: 15577909

59. Štáfková J, Rada P, Meloni D, Žárský V, Smutná T, Zimmann N, et al. Dynamic secretome of *Trichomonas vaginalis*: Case study of β-amylases. Mol Cell Proteomics. 2017; https://doi.org/10.1074/mcp.RA117.000434 PMID: 29233912

60. Gnanasundram SV, Koš M. Fast protein-depletion system utilizing tetracycline repressible promoter and N-end rule in yeast. Boone C, editor. Mol Biol Cell. The American Society for Cell Biology; 2015; 26: 762–768. https://doi.org/10.1091/mbc.E14-07-1186 PMID: 25540433

61. Becker T, Wenz L-S, Krüger V, Lehmann W, Müller JM, Goroncy L, et al. The mitochondrial import protein Mim1 promotes biogenesis of multispanning outer membrane proteins. J Cell Biol. 2011; 194: 387–395. Available from: http://jcb.rupress.org/content/194/3/387.abstract https://doi.org/10.1083/jcb.201102044 PMID: 21825073

62. Alva V, Nam S-Z, Söding J, Lupas AN. The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res. 2016; 44: W410–W415. https://doi.org/10.1093/nar/gkw348 PMID: 27131380

63. Finn RD, Clements J, Eddy SR. HMMER web server: Interactive sequence similarity searching. Nucleic Acids Res. 2011; 39: 29–37. https://doi.org/10.1093/nar/gkr367 PMID: 21593126

64. Finn RD, Clements J, Arndt W, Miller BL, Wheeler TJ, Schreiber F, et al. HMMER web server: 2015 Update. Nucleic Acids Res. 2015; 43: W30–W38. https://doi.org/10.1093/nar/gkv397 PMID: 25943547

65. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol. 2013; 30: 772–780. https://doi.org/10.1093/molbev/mst010 PMID: 23329690

66. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol. 2010; 10: 210. https://doi.org/10.1186/1471-2148-10-210 PMID: 20626897

67. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015; 32: 268–274. Available from: http://dx.doi.org/10.1093/molbev/msu300 PMID: 25371430

68. Buchan DWA, Minneci F, Nugent TCO, Bryson K, Jones DT. Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic Acids Res. 2013; 41: 349–357. https://doi.org/10.1093/nar/gkt381 PMID: 23748958

69. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen M, et al. Comparative protein structure modeling using modeller. Current Protocols in Bioinformatics. John Wiley & Sons, Inc.; 2002. https://doi.org/10.1002/0471250953.bi0506s15 PMID: 18428767

70. Maghrabi AHA, Mcguffin LJ. ModFOLD6: An accurate web server for the global and local quality estimation of 3D protein models. Nucleic Acids Res. 2017; 45: W416–W421. https://doi.org/10.1093/nar/gkx332 PMID: 28460136

71. McGuffin LJ, Shuid AN, Kempster R, Maghrabi AHA, Nealon JO, Salehe BR, et al. Accurate template-based modeling in CASP12 using the IntFOLD4-TS, ModFOLD6, and ReFOLD methods. Proteins Struct Funct Bioinforma. 2017; https://doi.org/10.1002/prot.25360 PMID: 28748648

72. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: Application to microtubules and the ribosome. Proc Natl Acad Sci. 2001; 98: 10037–10041. https://doi.org/10.1073/pnas.181342398 PMID: 11517324

73. Walther DM, Papic D, Bos MP, Tommassen J, Rapaport D. Signals in bacterial beta-barrel proteins are functional in eukaryotic cells for targeting to and assembly in mitochondria. Proc Natl Acad Sci U S A. 2009; 106: 2531–2536. https://doi.org/10.1073/pnas.0807830106 PMID: 19181862

74. Dawson SC, Sagolla MS, Cande WZ. The cenH3 histone variant defines centromeres in *Giardia intestinalis*. Chromosoma. 2007; 116: 175–184. https://doi.org/10.1007/s00412-006-0091-3 PMID: 17180675

75. Wiśniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Methods. Nature Publishing Group; 2009; 6: 359. Available from: http://dx.doi.org/10.1038/nmeth.1322

76. Vizcaíno JA, Csordas A, del-Toro N, Dianes JA, Griss J, Lavidas I, et al. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. 2016; 44: D447–D456. Available from: http://dx.doi.org/10.1093/nar/gkv1145 PMID: 26527722

77. Zhang K. Gctf: Real-time CTF determination and correction. J Struct Biol. Academic Press; 2016; 193: 1–12. https://doi.org/10.1016/j.jsb.2015.11.003 PMID: 26592709

78. Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, et al. EMAN2: An extensible image processing suite for electron microscopy. J Struct Biol. 2007; 157: 38–46. https://doi.org/10.1016/j.jsb.2006.05.009 PMID: 16859925

79. Scheres SHW. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. 2012; 180: 519–530. https://doi.org/10.1016/j.jsb.2012.09.006 PMID: 23000701

EC Journals.ASM.org

N-Terminal Presequence-Independent Import of Phosphofructokinase into Hydrogenosomes of *Trichomonas vaginalis*

Petr Rada,[a] Abhijith Radhakrishna Makki,[a] Verena Zimorski,[b] Sriram Garg,[b] Vladimír Hampl,[a] Ivan Hrdý,[a] Sven B. Gould,[b] Jan Tachezy[a]

Department of Parasitology, Charles University in Prague, Faculty of Science, Prague, Czech Republic[a]; Institute for Molecular Evolution, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany[b]

Mitochondrial evolution entailed the origin of protein import machinery that allows nuclear-encoded proteins to be targeted to the organelle, as well as the origin of cleavable N-terminal targeting sequences (NTS) that allow efficient sorting and import of matrix proteins. In hydrogenosomes and mitosomes, reduced forms of mitochondria with reduced proteomes, NTS-independent targeting of matrix proteins is known. Here, we studied the cellular localization of two glycolytic enzymes in the anaerobic pathogen *Trichomonas vaginalis*: PP_i-dependent phosphofructokinase (*Tv*PP_i-PFK), which is the main glycolytic PFK activity of the protist, and ATP-dependent PFK (*Tv*ATP-PFK), the function of which is less clear. *Tv*PP_i-PFK was detected predominantly in the cytosol, as expected, while all four *Tv*ATP-PFK paralogues were imported into *T. vaginalis* hydrogenosomes, although none of them possesses an NTS. The heterologous expression of *Tv*ATP-PFK in *Saccharomyces cerevisiae* revealed an intrinsic capability of the protein to be recognized and imported into yeast mitochondria, whereas yeast ATP-PFK resides in the cytosol. *Tv*ATP-PFK consists of only a catalytic domain, similarly to "short" bacterial enzymes, while *Sc*ATP-PFK includes an N-terminal extension, a catalytic domain, and a C-terminal regulatory domain. Expression of the catalytic domain of *Sc*ATP-PFK and short *Escherichia coli* ATP-PFK in *T. vaginalis* resulted in their partial delivery to hydrogenosomes. These results indicate that *Tv*ATP-PFK and the homologous ATP-PFKs possess internal structural targeting information that is recognized by the hydrogenosomal import machinery. From an evolutionary perspective, the predisposition of ancient ATP-PFK to be recognized and imported into hydrogenosomes might be a relict from the early phases of organelle evolution.

The transition of the mitochondrion into an ATP-producing organelle was the crucial event at the eukaryote origin (1). ATP synthesis in eukaryotes is typically compartmentalized, with glycolysis in the cytosol and pyruvate oxidation in the mitochondria, which is linked to highly efficient oxidative phosphorylation (1, 2). In protists, however, there are notable exceptions to the usual scheme regarding both glycolysis and pyruvate oxidation. In *Trichomonas vaginalis* and other eukaryotes that possess an anaerobic form of mitochonria called hydrogenosomes, pyruvate is oxidized within the organelle via less efficient anaerobic fermentation (3). *Giardia intestinalis*, *Entamoeba histolytica*, and other eukaryotes possess a reduced form of mitochondria called mitosomes that do not produce ATP at all (4). In these organisms, pyruvate oxidation takes place exclusively in the cytosol (1). In kinetoplastids, glycolysis is compartmentalized in specialized microbodies called glycosomes (5). In some green algae, the first half of the glycolytic pathway is localized in the chloroplast (6, 7), while in the diatom *Phaeodactylum tricornutum* and other stramenopiles, several glycolytic enzymes are targeted to multiple compartments, such as the cytosol, plastids, and mitochondria (8, 9).

A particularly vexing case of compartmentalization involves *T. vaginalis* phosphofructokinase (PFK). In *Trichomonas*, glycolysis proceeds via a pyrophosphate (PP_i)-dependent phosphofructokinase (PP_i-PFK) (10), an enzyme that is generally rare in eukaryotes, albeit typical in plants (11). Therefore, it was surprising that genes for ATP-dependent phosphofructokinase (ATP-PFK) turned up in the *Trichomonas* genome (12). Furthermore, peptides of the expressed protein were found in the hydrogenosomal proteome (13–15), although the exact topology of hydrogenosome-associated *T. vaginalis* ATP-PFK (*Tv*ATP-PFK) remains unclear (13, 15). PP_i-PFK and ATP-PFK share an evolutionary origin (16, 17). In bacteria, ATP-PFK is a homo-oligomeric enzyme that is formed by ~35-kDa subunits (18). In opisthokonts, ATP-PFK underwent gene duplication and fusion events, resulting in an ~90-kDa protein with an N-terminal catalytic domain and a C-terminal regulatory domain (19). The PP_i-PFK protein forms homo- or, in plants, heterotetramers of ~40- to 60-kDa subunits, and in Apicomplexa, the two subunits are fused to a protein of ~140 kDa (20). The advantage of using PP_i-PFK rather than ATP-PFK in glycolysis lies in the increased yield of ATP due to the replacement of ATP with PP_i as a phosphate donor in the phosphorylation of fructose-6-phosphate (3). This is particularly important for *T. vaginalis* and other anaerobes with energy metabolism based mainly on glycolysis (10).

In most eukaryotes, the N-terminal targeting sequences (NTS) are required for the delivery of nuclear-encoded proteins into the mitochondrial matrix, whereas the NTS-independent pathway is mainly involved in the routing of proteins into the outer and inner mitochondrial membranes and the intermembrane space. NTS are typically 15 to 55 residues in length and form a positively

Received 4 July 2015 **Accepted** 8 October 2015

Accepted manuscript posted online 16 October 2015

Citation Rada P, Makki AR, Zimorski V, Garg S, Hampl V, Hrdý I, Gould SB, Tachezy J. 2015. N-terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*. Eukaryot Cell 14:1264–1275. doi:10.1128/EC.00104-15.

Address correspondence to Jan Tachezy, tachezy@natur.cuni.cz.

Supplemental material for this article may be found at http://dx.doi.org/10.1128/EC.00104-15.

charged amphipathic α-helix (21). Upon preprotein delivery into the matrix by the outer (TOM) and inner (TIM) membrane translocases, the NTS is removed by a heterodimeric zinc-dependent mitochondrial processing peptidase (MPP) (22). Proteins routed by the NTS-independent pathway possess either a single or multiple internal targeting signals (ITS) (23). In *Saccharomyces cerevisiae* and human mitochondria, the components and mechanisms of protein import via the NTS-dependent pathway are well characterized (23), whereas less is known about protein import in hydrogenosomes. The NTS-dependent mechanism is present in hydrogenosomes and mitosomes (4, 24, 25), but a few studies have also reported NTS-independent import into the hydrogenosomes of *T. vaginalis* (26, 27, 58).

Interestingly, there are four ~35-kDa *Tv*ATP-PFK proteins encoded in the *T. vaginalis* genome, none of which possesses an NTS. The multiple copies preclude the generation of *Tv*ATP-PFK knockouts with current *Trichomonas* tools to study their functions, which remain mysterious. To clarify the localization and exact organellar topology of *Tv*ATP-PFK, we investigated the targeting of products encoded by *Tv*ATP-PFK genes when expressed in transformed *T. vaginalis* cells using immunofluorescence microscopy and cell fractionation, characterized the ATP dependence of *Tv*ATP-PFK import into isolated hydrogenosomes, and tested whether *Tv*ATP-PFK could be recognized as a substrate for NTS-independent import into yeast mitochondria. Conversely, we assessed whether the homologous catalytic domain of yeast ATP-PFK, as well as ~35-kDa *Escherichia coli* ATP-PFK (*Ec*ATP-PFK), showed a tendency to be imported into hydrogenosomes when expressed in *T. vaginalis*.

MATERIALS AND METHODS

T. vaginalis strain T1 (provided by J.-H. Tai, Institute of Biomedical Sciences, Taipei, Taiwan) was grown in Diamond's tryptone-yeast extract-maltose (TYM) medium supplemented with 10% (vol/vol) heat-inactivated horse serum. *S. cerevisiae* strain INVSc1 (Invitrogen) was grown in yeast extract-peptone-dextrose (YPD) medium or minimal medium devoid of uracil when transfected.

Phylogenetic analyses. The sequences of ATP-PFK and PP_i-PFK in a wide diversity of prokaryotes and eukaryotes were downloaded from the protein and EST database of GenBank release 200.0 and aligned with the *T. vaginalis* sequences with MAFFT (28; http://mafft.cbrc.jp/alignment/server/) using an L-INS-i strategy. The alignment was manually edited using BioEdit 7.0.9.0 (29), and 340 well-aligned positions were used for the subsequent analyses. The phylogenetic tree was constructed by the maximum-likelihood method in RAxML version 7.2.8 (30) using the PROTGAMMALGF model on the RAxML black box server (31). The statistical support was assessed by bootstrapping with 100 repetitions in RAxML. Bayesian posterior probabilities were calculated in Phylobayes (32) on the CIPRES Science Gateway v. 3.3 (http://www.phylo.org/index.php/). Two chains of Markov chain Monte Carlo were run under the CAT GTR model with a sampling frequency of 1,800. The run was terminated when the discrepancy observed across all bipartitions (maxdiff) dropped below 0.3 and effective sizes were larger than 50. The first 500 trees were discarded as burn in, and a consensus tree with posterior probabilities was calculated from the sample of 14,080 trees.

Gene cloning and transformation. Selected genes (*Tv*ATP-PFK1, TVAG_293770; *Tv*PP_i-PFK1, TVAG_430830; *T. vaginalis* ferredoxin 1 [Fdx1], TVAG_003900; *S. cerevisiae*ATP-PFK [*Sc*ATP-PFK], DAA08331; and *E. coli Ec*ATP-PFK, EFJ85506.1) were amplified by PCR from *T. vaginalis* and *S. cerevisiae* genomic DNA and cloned into the plasmids (i) pTagVag2, enabling the expression of the inserted genes with a C-terminal dihemagglutinin (di-HA) tag in trichomonads (33), and (ii) a self-modified version of plasmid pYES2/CT that allows the expression of the inserted genes with C-terminal green fluorescent protein (GFP) in yeasts. Transformed trichomonads and *S. cerevisiae* cells were selected as previously described (33, 34). The primers that were used for amplification and cloning of the selected genes into the pTagVag2 and pYES2/CT plasmids are shown in the supplemental material.

The pTagVag2 plasmid allows expression of the inserted genes under the control of the *T. vaginalis* hydrogenosomal α-subunit succinyl-coenzyme A (CoA) synthetase (SCSα) gene promoter (33). Alternatively, we used native promoters of selected genes instead of the SCSα promoter. The selected genes were amplified by PCR with 300 bp of upstream noncoding sequences and inserted into the pTagVag2 plasmid with a deleted SCSα promoter (pTagVagN). The primers used to amplify and clone the selected genes with their native promoters are shown in the supplemental material.

Immunofluorescence microscopy. Episomally expressed recombinant proteins were detected in trichomonads using a monoclonal mouse anti-HA antibody (35). In double-labeling experiments, hydrogenosomal malic enzyme was detected using a rabbit polyclonal antibody (36). A secondary Alexa Fluor 488 (green) donkey anti-mouse antibody and Alexa Fluor 594 (red) donkey anti-rabbit antibody were used for visualization of target proteins. The cells were examined using an Olympus Cell-R IX81 microscope system. The acquired images were processed using ImageJ software (version 1.4d) (http://rsbweb.nih.gov/ij/). In *S. cerevisiae* cells, episomally expressed recombinant proteins with GFP were detected and examined as described above. In double-labeling experiment, mitochondria were detected with MitoTracker dye (Invitrogen).

Enzyme assays. ATP-PFK activity was determined in the glycolytic direction using a continuous spectrophotometric assay according to the method of Chi et al. (37) with some modifications. The assay mixture for ATP-PFK consisted of 2 ml of 100 mM HEPES, 50 mM KCl, 3 mM $MgCl_2$, 1 mM EDTA, pH 7.0, buffer; 1 mM ATP; 20 mM fructose-6-phosphate; 0.15 to 0.20 mM NADH; 2 to 3 U each of aldolase, triosephosphate isomerase, and glycerol-3-phosphate dehydrogenase (Sigma-Aldrich); and 0.05% (vol/vol) Triton X-100 (ATP-PFK assay buffer). The assay was performed in 1-cm anaerobic cuvettes. The reaction was started by alternatively adding ATP, fructose-6-phosphate, auxiliary enzymes, or protein sample to the assay mixture, and the reaction was monitored as a decrease in the absorbance of NADH at 340 nm using a Shimadzu UV-2600 spectrophotometer. PP_i-PFK activity was determined as previously described (38). The protein concentrations in the subcellular fractions of *T. vaginalis* were determined by the Lowry protein assay.

Preparation of cellular fractions. Highly purified hydrogenosomes were obtained from *T. vaginalis* total cell lysates by differential and Percoll gradient centrifugation as described previously (35). The cytosolic fraction was isolated according to the method of Sutak et al. (35) and subsequently centrifuged at 190,000 × *g* (the high-speed cytosolic fraction). Mitochondria of *S. cerevisiae* were isolated from the yeast according to the method of Gregg et al. (39).

Protease protection assay. Aliquots of intact hydrogenosomes (3 mg) were resuspended in 1 ml of 1× ST buffer (250 mM sucrose, 10 mM Tris, pH 7.8, 0.5 mM KCl) supplemented with protease inhibitor cocktail tablets (Roche Complete, EDTA free). Trypsin (Sigma) was added to a final concentration of 200 μg/ml, and the samples were incubated at 37°C for 30 min. After incubation, the trypsin activity was stopped by the addition of soybean inhibitors (5 mg/ml), and the samples were analyzed by immunoblotting with a monoclonal mouse anti-HA antibody.

Aliquots of intact mitochondria (1 mg) were resuspended in 1 ml of SEM buffer (1 mM MOPS [morpholinepropanesulfonic acid]-KOH, pH 7.2, 250 mM sucrose, 1 mM EDTA). Proteinase K (Sigma) was added to a final concentration of 50 μg/ml, and the samples were incubated at 37°C for 30 min. After incubation, the proteinase K activity was stopped by the addition of 250 μl of trichloroacetic acid. The samples were analyzed by immunoblotting with a monoclonal anti-GFP antibody (Pierce).

Preparation of radiolabeled precursor proteins. The *Tv*ATP-PFK1 gene was cloned into the modified psp64 poly(A) plasmid, which enables

in vitro mRNA synthesis from the inserted genes (Promega). The primers designed for PCR and cloning into the psp64 plasmid are described in the supplemental material. *In vitro* transcription was performed using the mMachine kit (Ambion). [^{35}S]methionine-radiolabeled precursor protein was synthesized *in vitro* using the Flexi Rabbit Reticulocyte Lysate System (Promega).

In vitro import. Each *in vitro* import assay was performed in a reaction mixture that included 100 μl of import buffer (10 mM HEPES, pH 7.4, 250 mM sucrose, 2 mM KP_i, pH 7.4, 25 mM KCl, 10 mM $MgCl_2$, 0.5 mM EDTA, pH 8.0, 1 mM dithiothreitol [DTT], 10 mM ATP), 50 μl of cytosolic extract, 5 μl of radiolabeled precursor protein, and 5 mg of isolated hydrogenosomes. Apyrase (20 U/ml) was used for the import assay, which was conducted in the absence of ATP. The organelles were preincubated for 10 min at 25°C in import buffer with cytosolic extract, after which radiolabeled precursor protein was added to the assay mixture, and the mixture was incubated for 1, 10, and 60 min at 25°C. At each time point, the *in vitro* import was stopped by the addition of 100 μg/ml of proteinase K and placed on ice for 20 min. After incubation, the activity of proteinase K was inhibited by adding 2 mM phenylmethylsulfonyl fluoride (PMSF) (Sigma). The hydrogenosomes were then washed in import buffer and solubilized in SDS loading buffer. To test the activity of proteinase K, after a 60-min incubation of the protein import reaction mixture, the hydrogenosomes were dissolved with 0.5% (vol/vol) Triton X-100, followed by the addition of 100 μg/ml of proteinase K. Proteins in the supernatant were precipitated with methanol-chloroform and solubilized in SDS loading buffer. All of the samples were subjected to SDS-PAGE in a 13.5% separating gel. The gels were vacuum dried and exposed to X-ray films.

RESULTS

Phylogenetic analysis reveals the presence of PP_i-PFK and the short type of ATP-PFK in *T. vaginalis* and other parabasalids. The *T. vaginalis* genome possesses 11 genes encoding phosphofructokinases, four of which encode "short" (~35-kDa)-type ATP-dependent PFKs (*Tv*ATP-PFK1 to -4 [TVAG_293770, TVAG_496160, TVAG_462920, and TVAG_391760]) and seven of which encode PP_i-dependent PFKs (*Tv*PP_i-PFK1 to -7 [TVAG_430830, TVAG_077440, TVAG_281070, TVAG_364620, TVAG_079260, TVAG_263690, and TVAG_335880]). A phylogenetic analysis of ATP-PFKs and PP_i-PFKs revealed that *Trichomonas Tv*ATP-PFK1 to -4 fall into the single robust clade T2, together with PFKs from other parabasalids (Fig. 1). The closest eukaryotic relatives of this clade are tandem-fusion PFKs from opisthokonts and amoebozoans (clade E), as well as enzymes from prokaryotes (clades B1 and B2). The *Trichomonas* homologues *Tv*PP_i-PFK1 to -7 also form a clade with parabasalian sequences (Fig. 1). This parabasalian clade (clade T1) branches with enzymes from jakobids, heteroloboseans, and prokaryotes. The presence of both versions of the enzyme in other parabasalids suggests that both PP_i-PFK and ATP-PFK were present in the parabasalid ancestor. The branching of the *T. vaginalis* sequences in several unrelated positions in both clades T1 and T2 indicates that the genes have undergone gene duplications and possibly gene losses within parabasalids. The specificity of both types of PFKs for either ATP or PP_i has been ascribed to the amino acid residues at positions 104 and 124 (according to the numbering of the *E. coli Ec*ATP-PFK [40]). The G_{104} ($GGDG_{104}$ motif) and G/K_{124} residues are important for ATP binding, whereas PP_i binding requires residues D_{104} ($GGDD_{104}$ motif) and K_{124} (10, 17). *Tv*ATP-PFK1, -3, and -4 contain glycine at position 104, and *Tv*ATP-PFK1 and -3 contain glycine at position 124, whereas *Tv*ATP-PFK4 contains an alanine residue at the latter position (Fig. 2; see Fig. S2 in the supplemental material). The interchange of the glycine residue with alanine should not affect the interaction with the ATP molecule. The alanine residue possesses a small side chain, and it is unlikely that the residue creates steric hindrance to prevent binding of the ATP molecule. However, *Tv*ATP-PFK2 contains threonine and serine residues at positions 104 and 124, respectively. Therefore, the ability of *Tv*ATP-PFK2 to bind ATP is uncertain. The expected amino acid residues (D_{104} and K_{124}) are present in *Tv*PP_i-PFK1 and -3 to -6, whereas *Tv*PP_i-PFK2 and -7 contain glutamic acid and alanine residues at position 104, respectively (Fig. 2; see Fig. S2 in the supplemental material). Interestingly, scanning of the alignment of a broad range of sequences that were used for the phylogenetic analysis (Fig. 1) revealed the presence of paralogous genes with canonical G/D_{104} and G/K_{124} amino acid residues and with different residues at these positions in other parabasalids of clade T2 and in members of the Embryophyta, clade P. For example, serine residues at position 124 are also present in the putative ATP-PFKs of *Tritrichomonas foetus* and *Histomonas meleagridis* (see Fig. S2 in the supplemental material). Moreover, the *H. meleagridis* protein contains asparagine at position 104. These sequences, together with *Tv*ATP-PFK2 and -4, form the upper branch of clade T2 (Fig. 1). The unusual paralogues of Embryophyta PP_i-PFK-like sequences contain threonine/isoleucine and valine at positions 104 and 124, respectively (see Fig. S2 in the supplemental material), and they are grouped in the upper Embryophyta branch of clade P (Fig. 1). The functions of plant PP_i-PFK-like proteins are unknown (41).

Cellular localization of *Tv*ATP-PFK paralogues. The analysis of *Tv*ATP-PFK1 to -4 revealed an absence of sequence motifs thought to target precursors to hydrogenosomes. The *Tv*ATP-PFK sequences are colinear with their bacterial orthologues, lacking a predictable NTS and the cleavage site for the processing peptidase (Fig. 2). We found no internal motifs for subcellular targeting, and PSORT II predicted *Tv*ATP-PFKs to localize to the cytosol.

The subcellular localization of *Tv*ATP-PFK1 to -4 was investigated by the transient expression of C-terminally HA-tagged proteins in *T. vaginalis*. Immunofluorescence microscopy revealed that recombinant *Tv*ATP-PFK1, -2, and -4 colocalized with malic enzyme, the hydrogenosomal marker protein (Fig. 3; see Fig. S1 in the supplemental material), which suggested that these three proteins were transported into the hydrogenosomal matrix (we were unable to detect any expression of *Tv*ATP-PFK3 after several independent rounds of transfection). The topology of *Tv*ATP-PFK1 was further tested by protease protection assays. The treatment of isolated organelles with trypsin had no effect on the *Tv*ATP-PFK1 signal in the Western blot analysis, and the signal disappeared only in response to treatment with detergent (Fig. 3B). This finding indicates that *Tv*ATP-PFK is imported into *T. vaginalis* hydrogenosomes and is not associated with the organelle surface.

Although the bioinformatics analysis did not predict the presence of a cleavable NTS, we cannot exclude the possibility that a noncleavable "cryptic" NTS signal might direct *Tv*ATP-PFK1 to hydrogenosomes. Therefore, we expressed a truncated version of *Tv*ATP-PFK1 that lacked the first 16 amino acid residues (aa) (double the size of the known NTS in Fdx1). The truncated *Tv*ATP-PFK1 was delivered to the hydrogenosomes as its complete form (Fig. 3). This result confirmed that import of *Tv*ATP-PFK1 into hydrogenosomes is NTS independent. The expression of *Tv*PP_i-PFK revealed a cytosolic localization of the enzyme, as expected (Fig. 3).

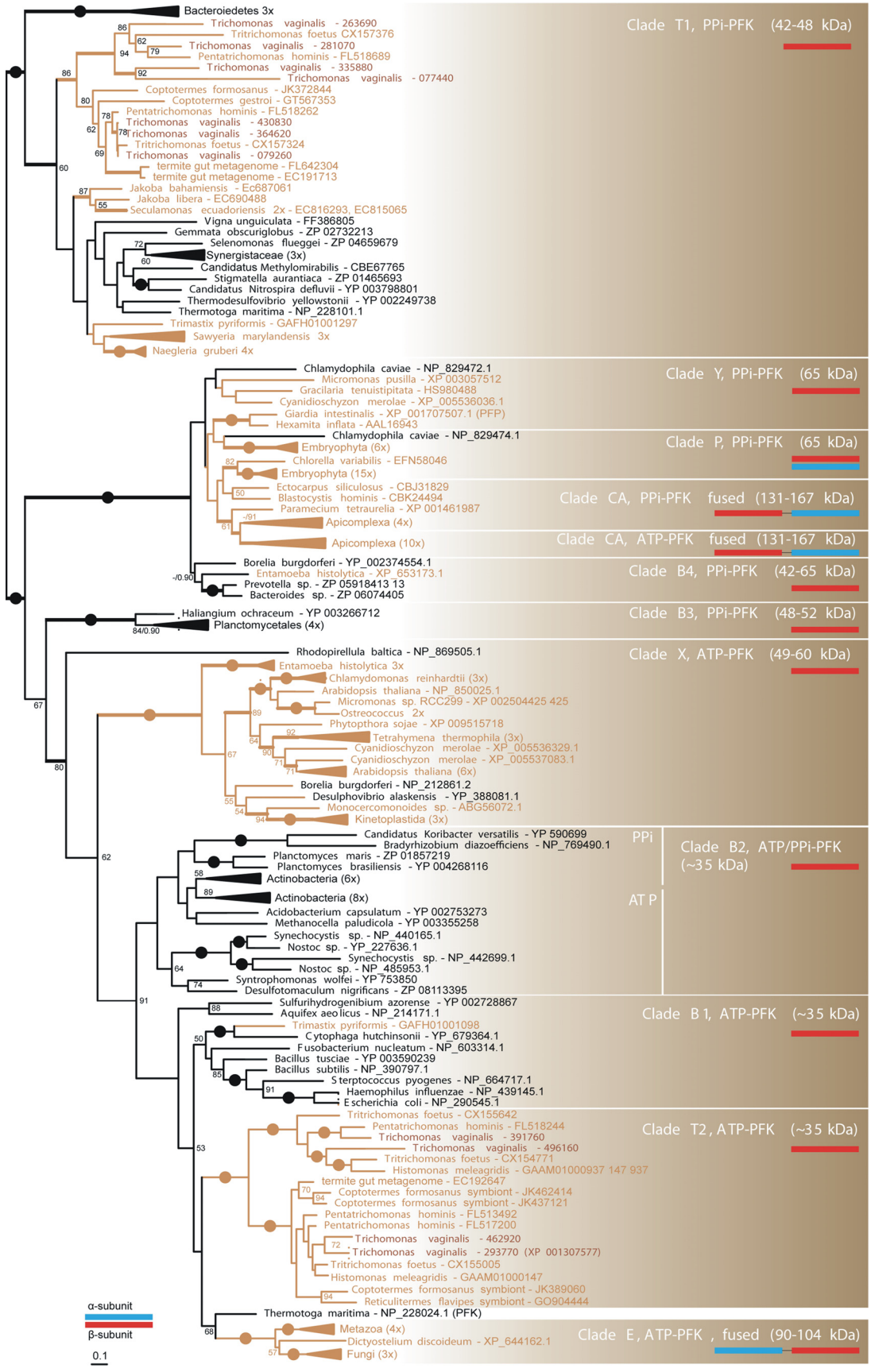


FIG 1 Phylogeny of ATP- and PP_i-dependent PFKs. Shown s a maximum-likelihood (ML) tree of PFK (191 taxa and 340 sites). The numbers at the nodes indicate bootstrap values (BV)/posterior probabilities (PP). Only BV and PP greater than 50% and 0.9, respectively, are shown. Branches with BV of >95% are marked by black circles, and branches with PP of >0.95 are marked by thick lines. Substrate specificity, molecular mass, and subunit composition for clades are indicated. The names of eukaryotes are in brown, and those of prokaryotes are in black.

```
TvATP-1                          MSLKNIAV LTSGGDNAGL  18-101 IGGNGSLSGA SLLAKDG--- -FPVIGMPGS IDDDVMG--T EVCVG 140-326
TvATP-2                            MKNIAI LSSGSDNSGI  16-100 IGGYTSLTQS KKFVDAG--- -IPTVAIPST IQDDIVG--T DICLG 139-324
TvATP-3                            MKSIGI LTSGGDSAGL  16-99  VGGNGSLAGA NLLQKDG--- -FPVIGLPGS IDDDVYG--T DVCIG 138-324
TvATP-4                            MKRIAV LSSGRDVSGA  16-99  VGGGGSFAHS RVLADKG--- -VPIIGIPAS IQDDVVG--T DICLG 138-323
EcATP                             MIKKIGV LTSGGDAPGM  17-100 IGGDGSYMGA MRLTEMG--- -FPCIGLPGT IDNDIKG--T DYTIG 159-340
ScATP  MQSQDSCYGVAFRSIITNDEΔSSQKKKKIAV MTSGGDSPGM 220-305 CGGDGSLTGA DLFRHEWPSK NLSIVGLVGS IDNDMSG--T DSTIG 367-987
TvPPi-1                        MSTEAPVLGI LCGGGPAPGL  20-110 IGGDLTASSA VSVASGMNGN EISVISCPKT IDNDLPLPAD QSTFG 155-426
TvPPi-2                         MSDAKTLCI VVTGGTSPGV  19-109 LSGNENVAMC HRIAEQFKND DIQVLVVAKT IDNDVPLPDF TSTFG 154-425
TvPPi-3                        MSTEAPVLGI IIGGAPAPGL  20-110 IGGNDKIATT HIITSGLDPA QMQVIAIPKT IDNDISLPYN TDTFG 155-429
TvPPi-4                        MSTEAPVLGI LCGGGPAPGL  20-110 IGGDLTASSA VSVAQGMDGN EISVISCPKT IDNDLPLPAD QSTFG 155-426
TvPPi-5                        MSAEAPVLGI LCGGGPAPGL  20-110 IGGDLTASSA VSVAQGMDGN EISVISCPKT IDNDLPLPSD QSTFG 155-425
TvPPi-6                 MFAQIEEPAKDAPILAI ICGGTPVPGL  27-117 IGGTDKVISA HIITQGIDPY SMSVLVIPKT IDNDVCLPYG QSTFG 162-434
TvPPi-7          MPQQYDYNLQSIEMGEPEILGI VVAGGTAPGL  32-122 IGGNAKLRMC HYISQGIDPT IMQVIAVPKT ISNDVQLPPE QTSLG 167-432
```

Δ= 180AA out

FIG 2 Multiple-protein-sequence alignment of the N-terminal portions and ATP/PP_i binding domains of *T. vaginalis* ATP- and PP_i-dependent PFKs. *T. vaginalis* TrichDB accession numbers: *Tv*ATP-PFK1 to -4, TVAG_293770, TVAG_496160, TVAG_462920, and TVAG_391760; *Tv*PP_i-PFK1 to -7, TVAG_430830, TVAG_077440, TVAG_281070, TVAG_364620, TVAG_079260, TVAG_263690, and TVAG_335880. NCBI accession numbers: *E. coli*, NP_418351; *S. cerevisiae*, DAA08331. A PSORT II-predicted NTS in *Sc*ATP-PFK is underlined; the arrow indicates the predicted cleavage site. The amino acid residues that are required for the interaction with ATP are shaded in green, and the residues that are crucial for the interaction with a PP_i molecule are shaded in red.

Next we investigated PP_i- and ATP-dependent PFK activities in cellular fractions of *T. vaginalis*. Under anaerobic conditions, we detected specific PP_i-PFK activity of 0.4 to 0.9 μmol min^{-1} mg $protein^{-1}$ in the high-speed cytosolic fraction. Percoll-purified hydrogenosomes contained a low specific activity (~0.008 to 0.020 μmol min^{-1} mg $protein^{-1}$) of ATP-PFK. PP_i-PFK activity was not associated with the organelles. These results indicate that PP_i- and ATP-dependent PFK activities are present in *T. vaginalis* in two distinct cellular compartments, in the cytosol and in hydrogenosomes, respectively. However, the hydrogenosomal (ATP-dependent) activity is dwarfed by the well-characterized cytosolic PP_i-dependent activity, raising questions about the role of the ATP-dependent activity, if any, in core energy metabolism.

Expression of *Tv*ATP-PFK1 and ferredoxin 1 under the control of native promoters. The *T. vaginalis* SCSα promoter is a strong endogenous promoter for transient expression (42). The unexpected localization of *Tv*ATP-PFK1 when transiently expressed under the control of the SCSα promoter prompted us to test whether the promoter itself could influence the localization of the product. First, we tested SCSα versus the native promoter (NP) by determining the cellular localization of Fdx1, a model hydrogenosomal matrix protein that possesses a typical NTS (24), as well as an ITS (26). Full-length Fdx1 expressed under the control of the SCSα promoter localized to hydrogenosomes (Fig. 4). However, the expression of the same protein with a deleted NTS (ΔFdx1, with deletion of the first 8 amino acids, MLSQVCRF) resulted in a dual localization: the majority of the ΔFdx1 was accumulated in the cytosol, whereas a portion of the ΔFdx1 was targeted to the organelle. The matrix localization of ΔFdx1 was verified by a protease protection assay (Fig. 4). When the SCSα promoter was replaced with the native Fdx1 promoter (300 bp upstream of the coding sequence of the Fdx1 gene), the complete Fdx1 protein was imported into hydrogenosomes; however, Fdx1 with a deleted NTS remained in the cytosol (Fig. 4). It thus appears that the nature of the promoter that is used for protein expression may affect protein localization. In the case of Fdx1, the ITS is apparently not sufficient to deliver the protein into the organelles when the protein is expressed without NTS (ΔFdx1) under the control of the native promoter. Therefore, we also assessed the localization of the recombinant *Tv*ATP-PFK1 expressed in *T. vaginalis* under the control of its native *Tv*ATP-PFK1 promoter (Fig. 3). Immunofluorescence microscopy and Western blot analysis confirmed that under these conditions, *Tv*ATP-PFK1 was targeted into the hydrogenosomal matrix (Fig. 3).

***In vitro* import of *Tv*ATP-PFK1 into hydrogenosomes.** *Tv*ATP-PFK1 import into hydrogenosomes was investigated using an *in vitro* import system. *Tv*ATP-PFK1 labeled with ^{35}S was incubated with hydrogenosomes in import buffer supplemented with ATP and cytosolic extract for 0 to 60 min. After the incubation, the hydrogenosomes were treated with proteinase K to remove labeled proteins that were not imported into the organelles. These experiments revealed the time-dependent accumulation of radiolabeled *Tv*ATP-PFK1 within isolated hydrogenosomes (Fig. 5). Furthermore, we investigated whether ATP was necessary for import. When the import assay was supplemented with apyrase (20 U/ml), which converts ATP to AMP and pyrophosphate, no import of *Tv*ATP-PFK1 was observed (Fig. 5). This result indicates that NTS-independent import of *Tv*ATP-PFK1 requires ATP.

TvATP-PFK is recognized and imported into yeast mitochondria. It has been demonstrated that mitochondria and hydrogenosomes employ a common mode of NTS-dependent protein import (24). Thus, we were curious whether *Tv*ATP-PFK1 possesses an NTS-independent signal that is recognized by the protein import machinery of yeast mitochondria. We expressed *Tv*ATP-PFK1 with a C-terminal GFP tag in *S. cerevisiae*. Immunofluorescence microscopy showed that the GFP fusion protein colocalized with the mitochondrial marker MitoTracker (Fig. 6). A protease protection assay using isolated yeast mitochondria revealed that *Tv*ATP-PFK1 was imported into the organelle and excluded the possibility that the protein was associated with the mitochondrial surface. Cytochrome oxidase subunit VI was used as a control inner membrane protein. *Sc*ATP-PFK consists of an N-terminal extension of 200 aa, a catalytic domain of 359 aa, and a C-terminal regulatory domain (423 aa). When we expressed a full-length *Sc*ATP-PFK and a truncated form that lacked the C-terminal regulatory domain (1/2*Sc*PFK) in yeast, both recombinant proteins remained in the cytosol after translation (Fig. 6). The unique N-terminal extension of *Sc*PFK is rich in negatively charged amino acid residues (pI 4.67), which might prevent the targeting of the protein to mitochondria (43). Thus, we also ex-

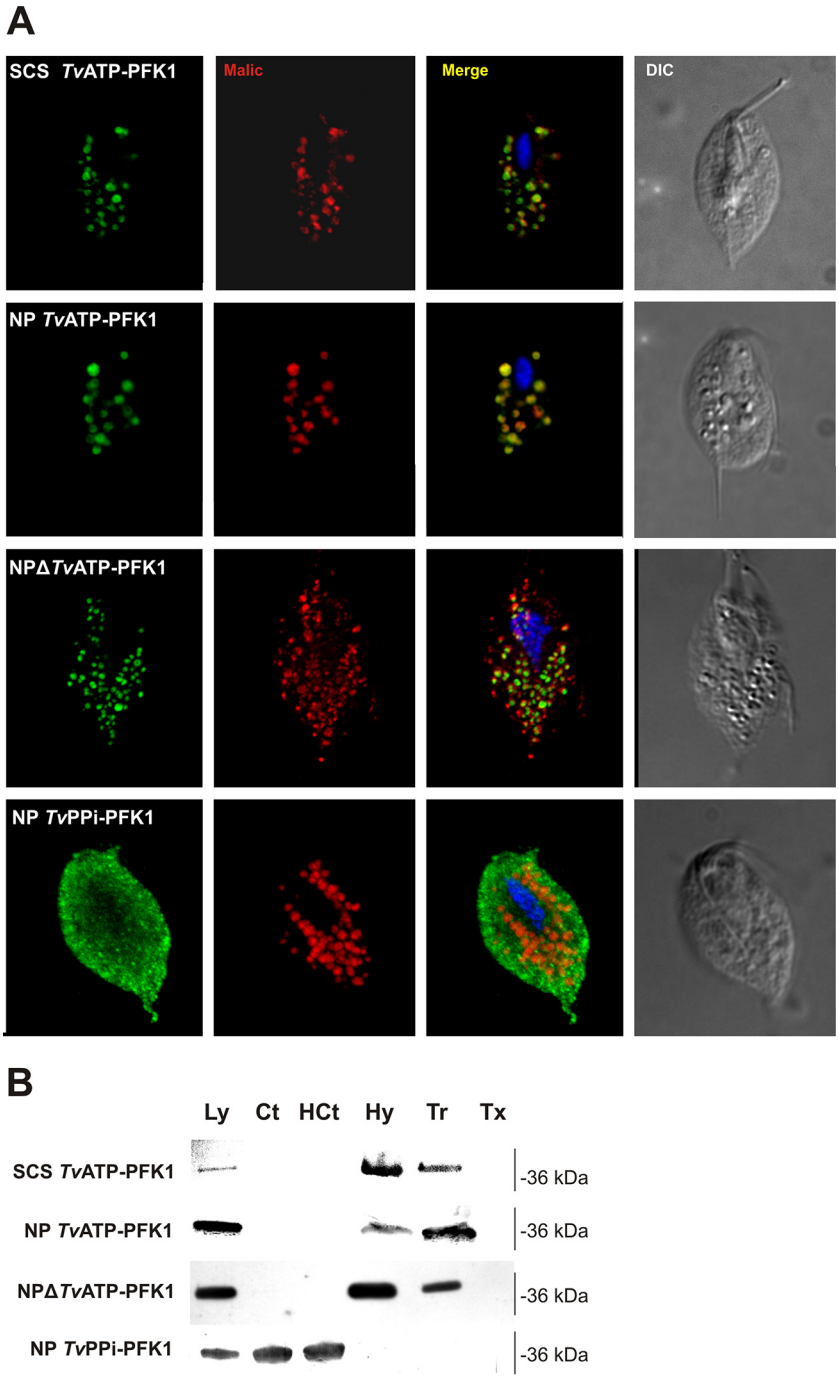


FIG 3 Cellular localization of ATP- and PP_i-dependent PFKs in *T. vaginalis*. (A) Immunofluorescence microscopy. Recombinant HA-tagged proteins were expressed in *T. vaginalis* cells and visualized using a monoclonal anti-HA antibody (green). *Tv*ATP-PFK1 and NP *Tv*ATP-PFK1 were expressed under the control of the strong SCSα promoter and the NP, respectively. NP Δ*Tv*ATP-PFK1 lacks 16 N-terminal amino acid residues. The hydrogenosomal marker protein malic enzyme was stained with a polyclonal rabbit antibody (red). The nucleus was stained using DAPI (4′,6-diamidino-2-phenylindole) (blue). DIC, differential interference contrast. (B) Protein protection assay. Hydrogenosomes were isolated from trichomonads expressing recombinant proteins with the C-terminal HA_2 tag and incubated with trypsin (Tr) or with trypsin and Triton X-100 (Tx). Samples were analyzed by immunoblotting using the monoclonal anti-HA tag antibody. Ly, total cell lysate; Ct, cytosol; HCt, high-speed cytosol; Hy, hydrogenosomes.

pressed the catalytic domain of *Sc*ATP-PFK, which is homologous to that of *Tv*ATP-PFK (ΔN1/2*Sc*PFK) alone. Interestingly, although some ΔN1/2*Sc*PFK signal was still observed in the cytosol, a significant portion was now also associated with the yeast mitochondrial membrane, as demonstrated by a protease protection assay (Fig. 6).

Collectively, these experiments show that *Tv*ATP-PFK1 possesses a targeting signal that is recognized by yeast mitochondria.

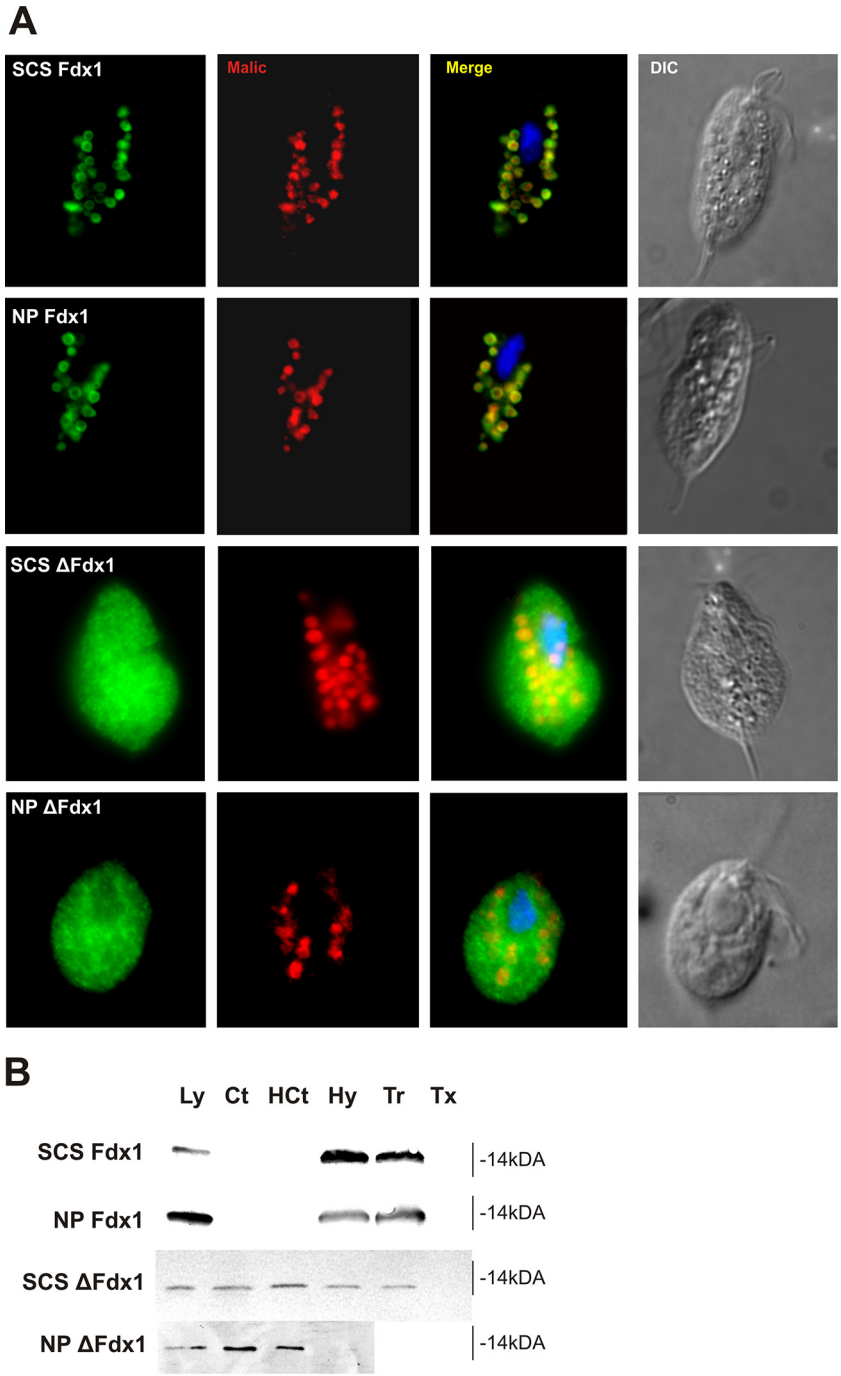


FIG 4 Effects of promoters on the cellular localization of ferredoxin. Fdx1 was used as a model protein with NTS-dependent targeting to test the effect of the SCSα promoter and the native promoter on Fdx1 localization. SCS Fdx1, Fdx1 (TVAG_003900) expressed under the control of the SCSα promoter; NP Fdx1, Fdx1 expressed under its native promoter; SCS ΔFdx1, Fdx1 with a deleted NTS that was expressed under the control of the SCSα promoter; NP ΔFdx1, ΔFdx1 expressed under the control of its native promoter. (A) Immunofluorescence microscopy. Recombinant HA-tagged proteins were expressed in *T. vaginalis* cells and visualized with monoclonal anti-HA antibody (green). The hydrogenosomal marker protein (malic enzyme) was detected using a polyclonal rabbit antibody (red). (B) Immunoblotting of subcellular fractions and protein protection assay. Ly, total cell lysate; Ct, cytosol; HCt, high-speed cytosol; Hy, hydrogenosomes; Tr, hydrogenosomes treated with trypsin; Tx, hydrogenosomal fraction treated with trypsin and Triton X-100.

The complete *Sc*ATP-PFK is retained in the cytosol, but the catalytic portion of *Sc*ATP-PFK displays mitochondrial membrane affinity.

Cellular localization of heterologous ATP-PFKs in *T. vaginalis*. We tested whether the hydrogenosomal protein import machinery can import heterologous ATP-PFKs. When we expressed complete *Sc*ATP-PFK in *T. vaginalis* under the control of the *Tv*ATP-PFK1 promoter, immunofluorescence microscopy revealed predominantly cytosolic localization of the protein, although the protein partially localized to hydrogenosomes (Fig. 7).

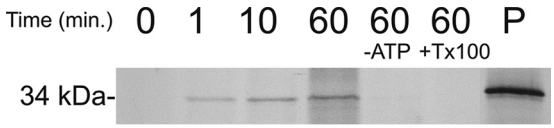


FIG 5 *In vitro* import of *Tv*ATP-PFK1 into hydrogenosomes. *In vitro*-synthesized ^{35}S-radiolabeled *Tv*ATP-PFK1 protein was incubated with isolated hydrogenosomes in import buffer at 25°C for 1, 10, and 60 min. At each time point, surface-associated proteins were degraded with proteinase K. Radiolabeled precursor was not imported in the absence of ATP (−ATP), depleted by addition of apyrase. A control for proteinase K activity was performed by the addition of Triton X-100 to the sample after 60 min of protein import (+Tx100). P, radiolabeled *Tv*ATP-PFK1 precursor protein. The samples were analyzed by SDS-PAGE and autoradiography.

The expression of 1/2*Sc*PFK revealed that the N-terminal half of *Sc*ATP-PFK was mainly associated with hydrogenosomes; however, the hydrogenosomal labeling was rather irregular in comparison to the labeling of malic enzyme, which was used as a control matrix protein. Western blot analysis of cellular fractions confirmed that both *Sc*ATP-PFK and 1/2*Sc*PFK were present in the cytosolic fractions (low- and high-speed cytosolic fractions). Parts of both proteins were also associated with the hydrogenosomal fractions; however, the signals disappeared after trypsin treatment. When we expressed only the catalytic part of the yeast enzyme lacking the negatively charged N-terminal sequence (ΔN1/2*Sc*PFK), a significant portion of the protein appeared inside the hydrogenosomes (Fig. 7). Next, we were interested in whether the targeting information is also present in short *E. coli* ATP-PFK orthologues that display 42% amino acid sequence identity with *Tv*ATP-PFKs. Thus, we expressed *Ec*ATP-PFK under the control of the *Tv*ATP-PFK1 promoter. Under these conditions, the *E. coli* protein was detected in the cytosol, and in part, it was associated with the hydrogenosomal surface (Fig. 7). However, when expressed under the SCSα promoter, a significant part of the protein was imported into the hydrogenosomes.

DISCUSSION

We investigated the cellular localization and NTS-independent import of *Tv*ATP-PFK into *T. vaginalis* hydrogenosomes. The parasite expresses both PP_i- and ATP-dependent enzymes, which are compartmentalized in the cytosol and hydrogenosomes, respectively. The classical PP_i-dependent activity of the parasite is about 50-fold higher than the newly characterized ATP-dependent activity, rendering the metabolic significance of the latter unclear. A phylogenetic analysis revealed that both types of PFKs are present across the parabasalids sampled so far. *Tv*ATP-PFK corresponds to a "short" ~35-kDa form of bacterial PFK that consists of only a catalytic domain, whereas the C-terminal regulatory domain typical of opisthokont ATP-PFKs is lacking. The targeting of *Tv*ATP-PFK1 to hydrogenosomes appears to be a

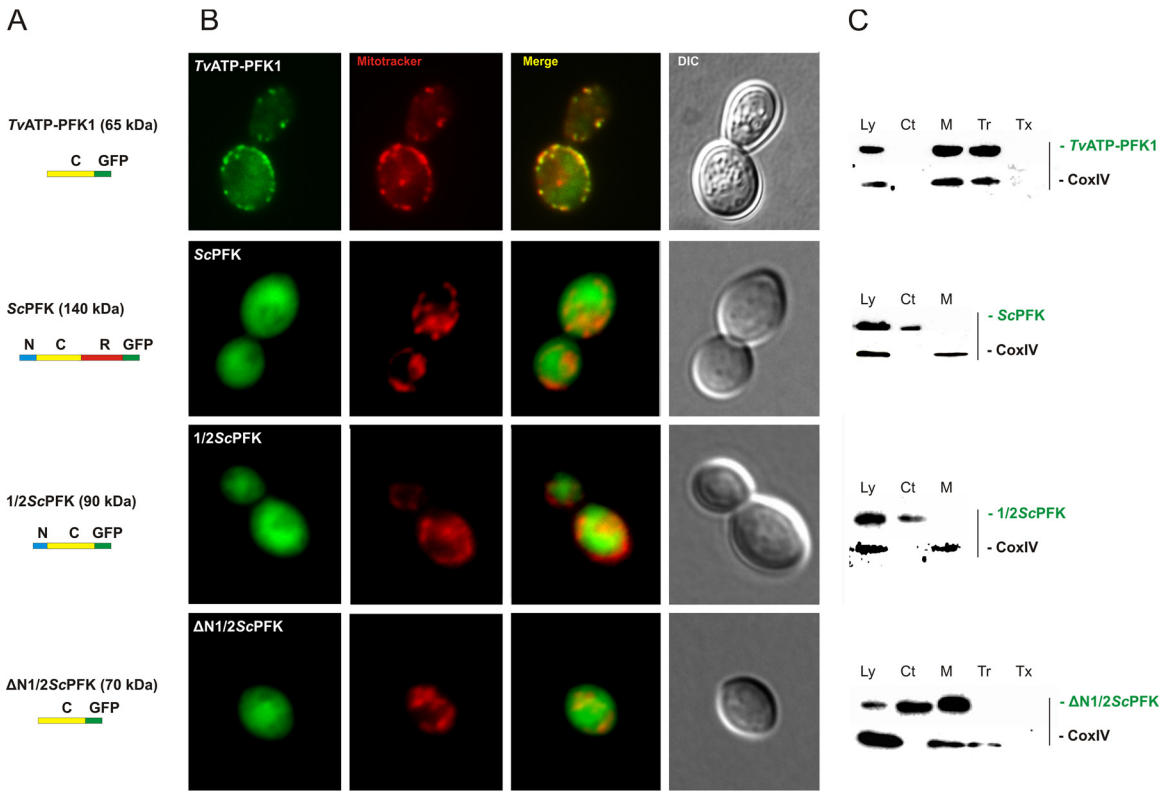


FIG 6 Cellular localization of *Tv*ATP-PFK1 and yeast ATP-PFK in *S. cerevisiae*. (A) Domain structure of the expressed proteins. N, N-terminal extension; C, catalytic domain; R, regulatory domain; GFP, green fluorescent protein tag. (B) Immunofluorescence microscopy. *Tv*ATP-PFK1 was expressed in yeasts with C-terminal GFP (green). Mitochondria were detected using MitoTracker dye (Invitrogen) (red). *Tv*ATP-PFK1, complete short *T. vaginalis* PFK; *Sc*PFK, complete long yeast PFK; 1/2*Sc*PFK, N-terminal extension (205 aa) and catalytic domain (359 aa) of *Sc*PFK; ΔN1/2*Sc*PFK, catalytic domain with deleted N-terminal extension. (C) Immunoblotting of subcellular fractions and protein protection assay. GFP-tagged proteins were detected using an anti-GFP antibody. Cytochrome oxidase subunit IV (CoxIV) was used as a mitochondrial marker, which was detected using a rabbit anti-CoxIV antibody. Ly, total cell lysate; Ct, cytosol; M, mitochondria; Tr, hydrogenosomes treated with trypsin; Tx, hydrogenosomal fraction treated with trypsin and Triton X-100.

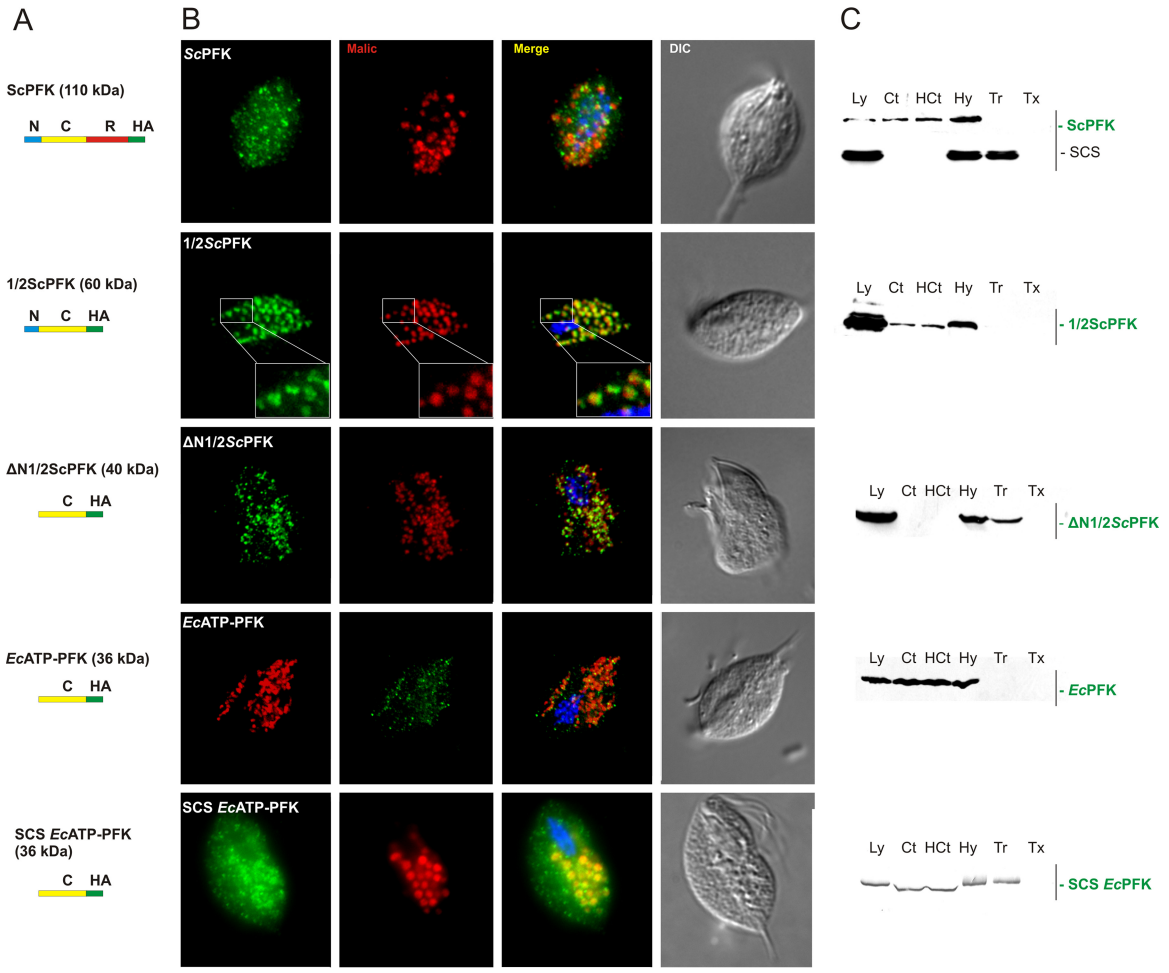


FIG 7 Cellular localization of *S. cerevisiae Sc*ATP-PFK and *Ec*ATP-PFK in *T. vaginalis*. (A) Domain structure of the expressed constructs. N, N-terminal extension; C, catalytic domain; R, regulatory domain; HA, hemagglutinin tag. (B) Recombinant HA-tagged proteins were expressed in *T. vaginalis* cells under the control of the *Tv*ATP-PFK1 promoter. SCS *Ec*ATP-PFK was expressed under the control of the SCSα promoter. HA-tagged proteins were visualized with mouse monoclonal anti-HA antibody (green). The hydrogenosomal marker protein (malic enzyme) was detected using a polyclonal rabbit antibody (red). (C) Immunoblotting of subcellular fractions and protein protection assay. Recombinant HA-tagged proteins were detected using monoclonal anti-HA antibody. The hydrogenosomal marker protein SCSα was detected using a rabbit polyclonal antibody. Ly, total cell lysate; Ct, cytosol; HCt, high-speed cytosol; Hy, hydrogenosomes; Tr, hydrogenosomes treated with trypsin; Tx, hydrogenosomal fraction treated with trypsin and Triton X-100.

highly specific and ATP-dependent process, even though the protein is not predicted to possess a cleavable NTS, which is typical of hydrogenosomal matrix proteins (44, 45).

The replacement of ATP with PP_i as a phosphate donor in the phosphorylation of fructose-6-phosphate allows an increased glycolytic ATP yield (3), conceivably a significant feature for a fermenting organism. Examples of organisms that express both PP_i-PFK and ATP-PFK are rare. The actinomycete *Amycolatopsis methanolica* possesses both genes, but their expression depends strictly on the carbon source (46). *Entamoeba histolytica* possesses two genes for PP_i-PFK orthologues; however, one of the gene products has been shown to utilize ATP instead of PP_i, and it has been suggested that the two enzymes might be expressed during different life stages (37). In plants, PP_i-PFK and ATP-PFK are both cytosolic enzymes with reciprocal expression responding to environmental perturbations (47). Whereas the expression of PP_i-PFK is upregulated by anoxia or orthophosphate deficiency, ATP-PFK is downregulated under such conditions. The spatial separation in *T. vaginalis* of PP_i-PFK and ATP-PFK to the cytosol and hydrogenosomes, respectively, could be an alternative solution to avoid interference between the two enzymes.

Specific targeting of *Tv*ATP-PFK to the organelle was demonstrated *in vivo* by episomal expression of tagged *Tv*ATP-PFK1 under SCSα and its native promoters, as well as the *in vitro* import of radiolabeled protein into isolated hydrogenosomes. Through the HA-tagged *Tv*ATP-PFK1, products of four paralogous *Tv*ATP-PFK genes were immunoprecipitated from isolated hydrogenosomes and identified by mass spectrometry. Earlier proteomic studies suggested association of the glycolytic pathway, including *Tv*ATP-PFK, with the hydrogenosome (13, 15), which raises the question of whether glycolytic enzymes form functional protein complexes on the hydrogenosomal outer membrane, as has been shown for mitochondria. For example, in *Arabidopsis thaliana*, 5 to 10% of each glycolytic enzyme is associated with the outer mitochondrial surface. Mammalian and fish heart mitochondria bind hexokinase and ATP-PFK (48), which has been discussed in the context of an increased glycolytic rate under hypoxic conditions (49). However, in *T. vaginalis*, expression of seven glycolytic

enzymes, including PP_i-PFK, showed exclusively cytosolic localization of these proteins (15, 26). Moreover, available cell fractionation studies of glyceraldehyde-3-phosphate dehydrogenase (50) and PP_i-PFK (this study) indicated that the corresponding activities are not associated with the organelle. These data do not support the formation of functional glycolytic complexes at the hydrogenosomal membrane and make the interpretation of previous proteomic analysis problematic, although systematic studies of glycolytic enzyme activities in cellular fractions of *T. vaginalis* are currently lacking. The localization of *Tv*ATP-PFK in the hydrogenosomal matrix, as shown in this study, is new for trichomonads.

Organellar forms of ATP-PFK have been found in glycosomes (51) and chloroplasts (52) thus far, where ATP-PFK operates within a known biochemical context. Kinetoplastids catalyze the "upper" six glycolytic steps in glycosomes, exporting 3-phosphoglycerate to the cytosol. Microalgae, such as *Chlamydomonas reinhardtii*, possess four glycolytic enzymes that convert glucose to glyceraldehyde-3-phosphate in chloroplasts, whereas the rest of glycolysis is localized in the cytosol (7). The most complicated glycolytic network has been found in diatoms, such as *P. tricornutum*, in which the complete set of glycolytic enzymes is present in the cytosol; nine glycolytic enzymes, including ATP-PFK, catalyze the conversion of glucose-1-phosphate to pyruvate in the chloroplast, and five glycolytic enzymes convert glyceraldehyde-3-phospate to pyruvate in the mitochondrion (8). In these organisms, the specific targeting of various glycolytic enzymes into the organelles is mediated by NTS (mitochondria), peroxisomal targeting signals (glycosomes), and plastid targeting signal (chloroplasts). The organellar *Tv*ATP-PFK found in *T. vaginalis* is unique with respect to three features: (i) it is a single glycolytic enzyme that is compartmentalized without apparent distal and proximal partners in the pathway, (ii) it is the only PFK that was observed to be imported into mitochondrion-related organelles, and (iii) the import into hydrogenosomes is mediated by ITS. The overall low hydrogenosomal ATP-PFK activity (approximately 2% of the PP_i-dependent activity), together with the lack of organellar glycolytic partners, raises questions regarding the metabolic role of *Tv*ATP-PFK and whether another function, unrelated to glycolysis, might be a possible alternative. Various moonlighting functions have been suggested for ATP-PFK in eukaryotes and bacteria, such as participation in the microautophagy of peroxisomes (53), RNA processing and degradation (54), and surface binding of plasminogen (55) and mannan (56). In our view, however, none of these functions currently appear likely for *Tv*ATP-PFK.

Heterologous expression of *Tv*ATP-PFK1 in *S. cerevisiae* revealed that the trichomonad enzyme is imported into yeast mitochondria, in addition to hydrogenosomes. This result indicates that *Tv*ATP-PFK1 possesses a targeting signal that is recognized by the hydrogenosomal, as well as the mitochondrial, import machinery. From an evolutionary perspective, these data suggest that the "short" ancient ATP-PFK might be predisposed to being recognized and imported into mitochondria, which might be a relict from the early phases of mitochondrial evolution. If so, the evolving eukaryotic cell had not only to develop a mechanism for retargeting nuclear-encoded proteins to mitochondria, but also to prevent the organellar translocation of some proteins, such as ATP-PFK, that are components of cytosolic pathways. Interestingly, unlike short bacterial ATP-PFK, eukaryotes frequently possess structurally modified long ATP-PFK that consists of catalytic and regulatory domains. In addition, the ATP-PFK of yeast and other fungi is equipped with a negatively charged N-terminal extension that may interfere with organellar import. Indeed, when we expressed the catalytic domain of *Sc*ATP-PFK with the N-terminal extension (1/2*Sc*PFK) in *T. vaginalis*, the protein was not delivered to the hydrogenosomal matrix, indicating that the extension prevents translocation. However, the hydrogenosomal import machinery was able to recognize and partially import truncated yeast *Sc*ATP-PFK, consisting of only the catalytic domain (ΔN1/2ScPFK), and the short proteobacterial *Ec*ATP-PFK, which are both homologous to *Tv*ATP-PFK. These results are consistent with the idea that ancient ATP-PFKs were predisposed to target the organelle. They also support previous analysis of proteins encoded by *E. coli* that predicted the presence of mitochondrial targeting information in about 5% of bacterial proteins (57).

The cell localization studies performed need to be interpreted with caution. Import of *Ec*ATP-PFK was observed when the gene was expressed under a strong SCSα promoter, while expression under the *Tv*ATP-PFK1 promoter resulted in partial association of *Tv*ATP-PFK1 with the outer hydrogenosomal membrane. Similarly, we observed promoter-dependent variation in the cell localization of Fdx, which possesses both NTS and ITS. Although we cannot exclude the possibility that hydrogenosomal localization of proteins expressed under strong promoters reflects protein mislocalization, it has been shown previously that six glycolytic enzymes expressed under the SCSα promoter remained exclusively in the cytosol, as expected, which argues against protein mislocalization (26). Therefore, it is more likely that, in addition to ITS, a suitable level of protein is required for protein translocation into the hydrogenosomes, while proteins without ITS are not targeted to the organelle regardless of the protein level. Importantly, expression of ΔN1/2ScPFK under *Tv*ATP-PFK1 was sufficient for its partial translocation into hydrogenosomes.

In conclusion, we identified ATP-PFK in *T. vaginalis* that is efficiently delivered into mitochondria and hydrogenosomes via NTS-independent mechanisms. Although NTS-independent targeting of membrane proteins is well documented, little is known about NTS-independent targeting of soluble proteins and the characters of multiple inner signals that are embedded within the protein structure (23, 58). The import of ATP-PFK into *T. vaginalis* hydrogenosomes can be used to investigate the molecular mechanisms that facilitate NTS-independent targeting and underpins the importance of internal targeting motifs that, in the case of PFK, are recognized in species spanning different eukaryotic supergroups. Intriguingly, the function of *Tv*ATP-PFK in *T. vaginalis* hydrogenosomes remains mysterious.

ACKNOWLEDGMENTS

This work was supported by the Czech Grant Foundation (13-09208J); the Biomedicine Centre of the Academy of Sciences and Charles University, Prague, Czech Republic (CZ.1.05/1.1.00/02.0109), from the European Regional Development Fund, project no. CZ.1.07/2.3.00/30.0061 (OPVK); and DFG grants to S.B.G. (GO1825/3–1). S.B.G. was additionally supported by a DFG grant to William F. Martin (MA1426/19-1).

REFERENCES

1. **Lane N, Martin W.** 2010. The energetics of genome complexity. Nature **467:**929–934. http://dx.doi.org/10.1038/nature09486.
2. **Hannaert V, Brinkmann H, Nowitzki U, Lee JA, Albert MA, Sensen CW, Gaasterland T, Muller M, Michels P, Martin W.** 2000. Enolase from *Trypanosoma brucei*, from the amitochondriate protist *Mastiga-*

moeba balamuthi, and from the chloroplast and cytosol of *Euglena gracilis*: pieces in the evolutionary puzzle of the eukaryotic glycolytic pathway. Mol Biol Evol **17:**989–1000. http://dx.doi.org/10.1093/oxfordjournals.molbev.a026395.

3. **Müller M, Mentel M, van Hellemond JJ, Henze K, Woehle C, Gould SB, Yu RY, van der Giezen M, Tielens AG, Martin WF.** 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. Microbiol Mol Biol Rev **76:**444–495. http://dx.doi.org/10.1128/MMBR.05024-11.
4. **Embley TM, Martin W.** 2006. Eukaryotic evolution, changes and challenges. Nature **440:**623–630. http://dx.doi.org/10.1038/nature04546.
5. **Gualdron-Lopez M, Brennand A, Hannaert V, Quinones W, Caceres AJ, Bringaud F, Concepcion JL, Michels PA.** 2012. When, how and why glycolysis became compartmentalised in the Kinetoplastea. A new look at an ancient organelle. Int J Parasitol **42:**1–20. http://dx.doi.org/10.1016/j.ijpara.2011.10.007.
6. **Schnarrenberger C, Pelzer-Reith B, Yatsuki H, Freund S, Jacobshagen S, Hori K.** 1994. Expression and sequence of the only detectable aldolase in *Chlamydomonas reinhardtii*. Arch Biochem Biophys **313:**173–178. http://dx.doi.org/10.1006/abbi.1994.1374.
7. **Johnson X, Alric J.** 2013. Central carbon metabolism and electron transport in *Chlamydomonas reinhardtii*: metabolic constraints for carbon partitioning between oil and starch. Eukaryot Cell **12:**776–793. http://dx.doi.org/10.1128/EC.00318-12.
8. **Kroth PG, Chiovitti A, Gruber A, Martin-Jezequel V, Mock T, Parker MS, Stanley MS, Kaplan A, Caron L, Weber T, Maheswari U, Armbrust EV, Bowler C.** 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. PLoS One **3:**e1426. http://dx.doi.org/10.1371/journal.pone.0001426.
9. **Nakayama T, Ishida K, Archibald JM.** 2012. Broad distribution of TPI-GAPDH fusion proteins among eukaryotes: evidence for glycolytic reactions in the mitochondrion? PLoS One **7:**e52340. http://dx.doi.org/10.1371/journal.pone.0052340.
10. **Mertens E, Ladror US, Lee JA, Miretsky A, Morris A, Rozario C, Kemp RG, Müller M.** 1998. The pyrophosphate-dependent phosphofructokinase of the protist, Trichomonas vaginalis, and the evolutionary relationships of protist phosphofructokinases. J Mol Evol **47:**739–750. http://dx.doi.org/10.1007/PL00006433.
11. **Winkler C, Delvos B, Martin W, Henze K.** 2007. Purification, microsequencing and cloning of spinach ATP-dependent phosphofructokinase link sequence and function for the plant enzyme. FEBS J **274:**429–438. http://dx.doi.org/10.1111/j.1742-4658.2006.05590.x.
12. **Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UC, Besteiro S, Sicheritz-Ponten T, Noel CJ, Dacks JB, Foster PG, Simillion C, Van de PY, Miranda-Saavedra D, Barton GJ, Westrop GD, Muller S, Dessi D, Fiori PL, Ren Q, Paulsen I, Zhang H, Bastida-Corcuera FD, Simoes-Barbosa A, Brown MT, Hayes RD, Mukherjee M, Okumura CY, Schneider R, Smith AJ, Vanacova S, Villalvazo M, Haas BJ, Pertea M, Feldblyum TV, Utterback TR, Shu CL, Osoegawa K, de Jong PJ, Hrdy I, Horvathova L, Zubacova Z, Dolezal P, Malik SB, Logsdon JM, Jr, Henze K, Gupta A, Wang CC, Dunne RL, Upcroft JA, Upcroft P, White O, Salzberg SL, Tang P, Chiu CH, Lee YS, Embley TM, Coombs GH, Mottram JC, Tachezy J, Fraser-Liggett CM, Johnson PJ.** 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science **315:**207–212. http://dx.doi.org/10.1126/science.1132894.
13. **Schneider RE, Brown MT, Shiflett AM, Dyall SD, Hayes RD, Xie Y, Loo JA, Johnson PJ.** 2011. The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. Int J Parasitol **41:**1421–1434. http://dx.doi.org/10.1016/j.ijpara.2011.10.001.
14. **Henze K.** 2007. The proteome of *T. vaginalis* hydrogenosomes, p 163–178. *In* Tachezy J (ed), Hydrogenosomes and mitosomes: mitochondria of anaerobic eukaryotes. Springer, Berlin, Germany.
15. **Rada P, Dolezal P, Jedelsky PL, Bursac D, Perry AJ, Sedinova M, Smiskova K, Novotny M, Beltran NC, Hrdy I, Lithgow T, Tachezy J.** 2011. The core components of organelle biogenesis and membrane transport in the hydrogenosomes of *Trichomonas vaginalis*. PLoS One **6:**e24428. http://dx.doi.org/10.1371/journal.pone.0024428.
16. **Liapounova NA, Hampl V, Gordon PM, Sensen CW, Gedamu L, Dacks JB.** 2006. Reconstructing the mosaic glycolytic pathway of the anaerobic eukaryote Monocercomonoides. Eukaryot Cell **5:**2138–2146. http://dx.doi.org/10.1128/EC.00258-06.
17. **Bapteste E, Moreira D, Philippe H.** 2003. Rampant horizontal gene transfer and phospho-donor change in the evolution of the phosphofructokinase. Gene **318:**185–191. http://dx.doi.org/10.1016/S0378-1119(03)00797-2.
18. **Shirakihara Y, Evans PR.** 1988. Crystal structure of the complex of phosphofructokinase from *Escherichia coli* with its reaction products. J Mol Biol **204:**973–994. http://dx.doi.org/10.1016/0022-2836(88)90056-3.
19. **Poorman RA, Randolph A, Kemp RG, Heinrikson RL.** 1984. Evolution of phosphofructokinase: gene duplication and creation of new effector sites. Nature **309:**467–469. http://dx.doi.org/10.1038/309467a0.
20. **Mony BM, Mehta M, Jarori GK, Sharma S.** 2009. Plant-like phosphofructokinase from *Plasmodium falciparum* belongs to a novel class of ATP-dependent enzymes. Int J Parasitol **39:**1441–1453. http://dx.doi.org/10.1016/j.ijpara.2009.05.011.
21. **Vogtle FN, Wortelkamp S, Zahedi RP, Becker D, Leidhold C, Gevaert K, Kellermann J, Voos W, Sickmann A, Pfanner N, Meisinger C.** 2009. Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. Cell **139:**428–439. http://dx.doi.org/10.1016/j.cell.2009.07.045.
22. **Gakh O, Cavadini P, Isaya G.** 2002. Mitochondrial processing peptidases. Biochim Biophys Acta **1592:**63–77. http://dx.doi.org/10.1016/S0167-4889(02)00265-3.
23. **Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N.** 2009. Importing mitochondrial proteins: machineries and mechanisms. Cell **138:**628–644. http://dx.doi.org/10.1016/j.cell.2009.08.005.
24. **Bradley PJ, Lahti CJ, Plumper E, Johnson PJ.** 1997. Targeting and translocation of proteins into the hydrogenosome of the protist Trichomonas: similarities with mitochondrial protein import. EMBO J **16:**3484–3493. http://dx.doi.org/10.1093/emboj/16.12.3484.
25. **Dolezal P, Smid O, Rada P, Zubacova Z, Bursac D, Sutak R, Nebesarova J, Lithgow T, Tachezy J.** 2005. Giardia mitosomes and trichomonad hydrogenosomes share a common mode of protein targeting. Proc Natl Acad Sci U S A **102:**10924–10929. http://dx.doi.org/10.1073/pnas.0500349102.
26. **Zimorski V, Major P, Hoffmann K, Bras XP, Martin WF, Gould SB.** 2013. The N-terminal sequences of four major hydrogenosomal proteins are not essential for import into hydrogenosomes of *Trichomonas vaginalis*. J Eukaryot Microbiol **60:**89–97. http://dx.doi.org/10.1111/jeu.12012.
27. **Mentel M, Zimorski V, Haferkamp P, Martin W, Henze K.** 2008. Protein import into hydrogenosomes of *Trichomonas vaginalis* involves both N-terminal and internal targeting signals: a case study of thioredoxin reductases. Eukaryot Cell **7:**1750–1757. http://dx.doi.org/10.1128/EC.00206-08.
28. **Katoh K, Kuma K, Toh H, Miyata T.** 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res **33:**511–518. http://dx.doi.org/10.1093/nar/gki198.
29. **Hall TA.** 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser **41:**95–98.
30. **Stamatakis A.** 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22:**2688–2690. http://dx.doi.org/10.1093/bioinformatics/btl446.
31. **Stamatakis A, Hoover P, Rougemont J.** 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol **57:**758–771. http://dx.doi.org/10.1080/10635150802429642.
32. **Lartillot N, Rodrigue N, Stubbs D, Richer J.** 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. Syst Biol **62:**611–615. http://dx.doi.org/10.1093/sysbio/syt022.
33. **Hrdy I, Hirt RP, Dolezal P, Bardonova L, Foster PG, Tachezy J, Embley TM.** 2004. Trichomonas hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. Nature **432:**618–622. http://dx.doi.org/10.1038/nature03149.
34. **Niedenthal RK, Riles L, Johnston M, Hegemann JH.** 1996. Green fluorescent protein as a marker for gene expression and subcellular localization in budding yeast. Yeast **12:**773–786.
35. **Sutak R, Dolezal P, Fiumera HL, Hrdy I, Dancis A, Delgadillo-Correa MG, Johnson PJ, Müller M, Tachezy J.** 2004. Mitochondrial-type assembly of FeS centers in the hydrogenosomes of the amitochondriate eukaryote *Trichomonas vaginalis*. Proc Natl Acad Sci U S A **101:**10368–10373. http://dx.doi.org/10.1073/pnas.0401319101.
36. **Drmota T, Proost P, Van Ranst M, Weyda F, Kulda J, Tachezy J.** 1996. Iron-ascorbate cleavable malic enzyme from hydrogenosomes of

Trichomonas vaginalis: purification and characterization. Mol Biochem Parasitol **83:**221–234. http://dx.doi.org/10.1016/S0166-6851(96)02777-6.

37. **Chi AS, Deng Z, Albach RA, Kemp RG.** 2001. The two phosphofructokinase gene products of *Entamoeba histolytica*. J Biol Chem **276:**19974–19981. http://dx.doi.org/10.1074/jbc.M011584200.
38. **Mertens E, Van Schaftingen E, Müller M.** 1989. Presence of a fructose-2,6-bisphosphate-insensitive pyrophosphate: fructose-6-phosphate phosphotransferase in the anaerobic protozoa *Tritrichomonas foetus*, *Trichomonas vaginalis* and *Isotricha prostoma*. Mol Biochem Parasitol **37:**183–190. http://dx.doi.org/10.1016/0166-6851(89)90150-3.
39. **Gregg C, Kyryakov P, Titorenko VI.** 2009. Purification of mitochondria from yeast cells. J Vis Exp **30:**1417. http://dx.doi.org/10.3791/1417.
40. **Michels PA, Chevalier N, Opperdoes FR, Rider MH, Rigden DJ.** 1997. The glycosomal ATP-dependent phosphofructokinase of *Trypanosoma brucei* must have evolved from an ancestral pyrophosphate-dependent enzyme. Eur J Biochem **250:**698–704. http://dx.doi.org/10.1111/j.1432-1033.1997.00698.x.
41. **Mustroph A, Sonnewald U, Biemelt S.** 2007. Characterisation of the ATP-dependent phosphofructokinase gene family from *Arabidopsis thaliana*. FEBS Lett **581:**2401–2410. http://dx.doi.org/10.1016/j.febslet.2007.04.060.
42. **Delgadillo MG, Liston DR, Niazi K, Johnson PJ.** 1997. Transient and selectable transformation of the parasitic protist *Trichomonas vaginalis*. Proc Natl Acad Sci U S A **94:**4716–4720. http://dx.doi.org/10.1073/pnas.94.9.4716.
43. **Gerber J, Neumann K, Prohl C, Muhlenhoff U, Lill R.** 2004. The yeast scaffold proteins Isu1p and Isu2p are required inside mitochondria for maturation of cytosolic Fe/S proteins. Mol Cell Biol **24:**4848–4857. http://dx.doi.org/10.1128/MCB.24.11.4848-4857.2004.
44. **Smid O, Matuskova A, Harris SR, Kucera T, Novotny M, Horvathova L, Hrdy I, Kutejova E, Hirt RP, Embley TM, Janata J, Tachezy J.** 2008. Reductive evolution of the mitochondrial processing peptidases of the unicellular parasites *Trichomonas vaginalis* and *Giardia intestinalis*. PLoS Pathog **4:**e1000243. http://dx.doi.org/10.1371/journal.ppat.1000243.
45. **Burstein D, Gould SB, Zimorski V, Kloesges T, Kiosse F, Major P, Martin WF, Pupko T, Dagan T.** 2012. A machine learning approach to identify hydrogenosomal proteins in *Trichomonas vaginalis*. Eukaryot Cell **11:**217–228. http://dx.doi.org/10.1128/EC.05225-11.
46. **Alves AM, Euverink GJ, Santos H, Dijkhuizen L.** 2001. Different physiological roles of ATP- and PP(i)-dependent phosphofructokinase isoenzymes in the methylotrophic actinomycete *Amycolatopsis methanolica*. J Bacteriol **183:**7231–7240. http://dx.doi.org/10.1128/JB.183.24.7231-7240.2001.
47. **Plaxton WC, Tran HT.** 2011. Metabolic adaptations of phosphate-starved plants. Plant Physiol **156:**1006–1015. http://dx.doi.org/10.1104/pp.111.175281.
48. **Treberg JR, MacCormack TJ, Lewis JM, Almeida-Val VM, Val AL, Driedzic WR.** 2007. Intracellular glucose and binding of hexokinase and phosphofructokinase to particulate fractions increase under hypoxia in heart of the Amazonian armored catfish (*Liposarcus pardalis*). Physiol Biochem Zool **80:**542–550. http://dx.doi.org/10.1086/520129.
49. **Graham JW, Williams TC, Morgan M, Fernie AR, Ratcliffe RG, Sweetlove LJ.** 2007. Glycolytic enzymes associate dynamically with mitochondria in response to respiratory demand and support substrate channeling. Plant Cell **19:**3723–3738. http://dx.doi.org/10.1105/tpc.107.053371.
50. **Markos A, Miretsky A, Müller M.** 1993. A glyceraldehyde-3-phosphate dehydrogenase with eubacterial features in the amitochondriate eukaryote, *Trichomonas vaginalis*. J Mol Evol **37:**631–643. http://dx.doi.org/10.1007/BF00182749.
51. **Opperdoes FR, Borst P.** 1977. Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: the glycosome. FEBS Lett **80:**360–364. http://dx.doi.org/10.1016/0014-5793(77)80476-6.
52. **Plaxton WC.** 1996. The organization and regulation of plant glycolysis. Annu Rev Plant Physiol Plant Mol Biol **47:**185–214. http://dx.doi.org/10.1146/annurev.arplant.47.1.185.
53. **Gancedo C, Flores CL, Gancedo JM.** 2014. Evolution of moonlighting proteins: insight from yeasts. Biochem Soc Trans **42:**1715–1719. http://dx.doi.org/10.1042/BST20140199.
54. **Commichau FM, Rothe FM, Herzberg C, Wagner E, Hellwig D, Lehnik-Habrink M, Hammer E, Volker U, Stulke J.** 2009. Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. Mol Cell Proteomics **8:**1350–1360. http://dx.doi.org/10.1074/mcp.M800546-MCP200.
55. **Kinnby B, Booth NA, Svensater G.** 2008. Plasminogen binding by oral streptococci from dental plaque and inflammatory lesions. Microbiology **154:**924–931. http://dx.doi.org/10.1099/mic.0.2007/013235-0.
56. **Katakura Y, Sano R, Hashimoto T, Ninomiya K, Shioya S.** 2010. Lactic acid bacteria display on the cell surface cytosolic proteins that recognize yeast mannan. Appl Microbiol Biotechnol **86:**319–326. http://dx.doi.org/10.1007/s00253-009-2295-y.
57. **Lucattini R, Likic VA, Lithgow T.** 2004. Bacterial proteins predisposed for targeting to mitochondria. Mol Biol Evol **21:**652–658. http://dx.doi.org/10.1093/molbev/msh058.
58. **Garg S, Stolting J, Zimorski V, Rada P, Tachezy J, Martin WF, Gould SB.** 2015. Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. Genome Biol Evol **7:**2716–2726. http://dx.doi.org/10.1093/gbe/evv175.

Molecular Microbiology (2018) 0(0), 1–16 doi:10.1111/mmi.14175

Targeting of tail-anchored proteins to *Trichomonas vaginalis* hydrogenosomes

Petr Rada, Abhijith Makki, Vojtěch Žárský and Jan Tachezy *

Department of Parasitology, Faculty of Science, Charles University, BIOCEV, Průmyslová 595, Vestec 25242, Czech Republic.

Summary

Tail-anchored (TA) proteins are membrane proteins that are found in all domains of life. They consist of an N-terminal domain that performs various functions and a single transmembrane domain (TMD) near the C-terminus. In eukaryotes, TA proteins are targeted to the membranes of mitochondria, the endoplasmic reticulum (ER), peroxisomes and in plants, chloroplasts. The targeting of these proteins to their specific destinations correlates with the properties of the C-terminal domain, mainly the TMD hydrophobicity and the net charge of the flanking regions. *Trichomonas vaginalis* is a human parasite that has adapted to oxygen-poor environment. This adaptation is reflected by the presence of highly modified mitochondria (hydrogenosomes) and the absence of peroxisomes. The proteome of hydrogenosomes is considerably reduced; however, our bioinformatic analysis predicted 120 putative hydrogenosomal TA proteins. Seven proteins were selected to prove their localization. The elimination of the net positive charge in the C-tail of the hydrogenosomal TA4 protein resulted in its dual localization to hydrogenosomes and the ER, causing changes in ER morphology. Domain mutation and swap experiments with hydrogenosomal (TA4) and ER (TAPDI) proteins indicated that the general principles for specific targeting are conserved across eukaryotic lineages, including *T. vaginalis*; however, there are also significant lineage-specific differences.

Accepted 25 November, 2018. *For correspondence. E-mail tachezy@natur.cuni.cz; Tel. +420 325 874 144; Fax +420 224 919 704.

Introduction

The proteome of mitochondria consists of over a thousand proteins that are encoded in the nucleus, synthesized in the cytosol and targeted to the organelles via N-terminal or internal targeting signals (Wiedemann and Pfanner, 2017). At the outer mitochondrial membrane (OMM), these proteins are recognized by an elaborate complex called the translocase of the outer mitochondrial membrane (TOM) that is coupled to the translocase of the inner mitochondrial membrane (TIM) to mediate the import of proteins to their final destinations (Chacinska *et al.*, 2009; Wiedemann and Pfanner, 2017). Only a few proteins are encoded in the mitochondrial genome, which represents a remnant genome of a premitochondrial ancestor of α-proteobacterial origin (Embley and Martin, 2006; Roger *et al.*, 2017). In certain forms of mitochondria, such as hydrogenosomes and mitosomes, the genome and most mitochondrial functions were entirely lost during the course of reductive evolution (Clemens and Johnson, 2000; Tovar *et al.*, 2003; Hrdý *et al.*, 2008). The biogenesis and functions of these reduced mitochondria are completely dependent on protein import.

The α-helical tail-anchored (TA) proteins represent a specific set of mitochondrial proteins that are delivered to the OMM. The TA proteins have a single transmembrane domain (TMD) near their C-terminus and a long N-terminal domain facing the cytosol. During biosynthesis, the TMD of TA proteins emerges from the ribosome only after the termination of translation. Consequently, all TA proteins are imported in membranes posttranslationally (Chio *et al.*, 2017; Costello *et al.*, 2017a; 2017b). In addition to mitochondria, TA proteins are components of other membrane-bound organelles, such as the endoplasmic reticulum (ER), peroxisomes, and in plants, the outer membrane of plastids (Kriechbaumer *et al.*, 2009; Chio *et al.*, 2017; Costello *et al.*, 2017a; 2017b). Altogether, TA proteins represent 3–5% of the eukaryotic membrane proteome (Hegde and Keenan, 2011). It is well known that the targeting signal of TA proteins is embedded in their C-terminal domain; however, in spite of a progress in understanding of the mechanisms involved in the sorting of TA proteins into ER (Mateja *et al.*, 2015; Cho and Shan, 2018), the targeting into the OMM, is still poorly understood.

The TMDs and short C-tail segments (CTSs) of mitochondrial TA proteins do not display similarity in their primary structures. The ability to be recognized and incorporated into the OMM is instead conferred by the physicochemical properties (Borgese and Fasana, 2011). Based on the studies in mammalian and yeast cells, these properties include a moderate hydrophobicity of the TMD, a short TMD sequence of less than 20 amino acid (AA) residues, and the presence of basic AA residues that provide positive charges to one or both TMD flanking regions. Subtle changes in the properties of the C-tail anchor can mistarget mitochondrial TA proteins to the cytosol or ER (Kuroda *et al.*, 1998; Borgese *et al.*, 2001; Hwang *et al.*, 2004; Henderson *et al.*, 2007). Particularly, interplay between the tail charge and TMD hydrophobicity appeared to be critical to control the correct targeting of TA proteins into the cellular organelles including ER, mitochondria and peroxisome (Costello *et al.*, 2017a; 2017b).

Several modes of targeting and insertion have been proposed for the incorporation of TA proteins into the OMM. Unassisted insertion, which involves the translocation of the C-terminal domain of TA proteins into the OMM independent of any cytosolic or membrane proteins, has been reported for the mitochondrial isoform of cytochrome b_5 (Colombo *et al.*, 2009) and mitochondrial fission 1 protein (Kemper *et al.*, 2008). Insertion with the assistance of membrane receptors was found for three small TA subunits of the TOM complex (Tom5, Tom6 and Tom7) that are recognized by Tom40 (Allen *et al.*, 2002; Horie *et al.*, 2003). The apoptosis regulators Bcl-2 and Bax were proposed to interact with the Tom20 and Tom22 receptors, respectively, which assist in their import into the OMM, bypassing the Tom40 import pore (Motz *et al.*, 2002; Bellot *et al.*, 2007). It has been hypothesized that the involvement of the cytosolic proteins in the targeting of TA proteins to the OMM may compete with other pathways for the substrate or may maintain the substrate in an insertion-competent form or an unproductive complex (Colombo *et al.*, 2009; Borgese and Fasana, 2011; Marty *et al.*, 2014). However, the specific cytosolic factors that recognize mitochondrial TA proteins remain elusive. There is also evidence that the inherent lipid composition, particularly the level of ergosterol, plays an important role in the specific targeting of TA proteins to the OMM (Krumpe *et al.*, 2012).

Considerably more information is available on the biogenesis of TA proteins in the ER and their subsequent transport to other compartments of the secretory pathway (Rabu *et al.*, 2009; Borgese and Fasana, 2011; Borgese, 2016). There are multiple pathways that assist in the posttranslational targeting of TA proteins to the ER. They are based on the interaction of TA proteins with (i) the signal recognition particle (SRP-assisted insertion), (ii) the Hsc70/Hsp40 system of chaperones, (iii) targeting and insertion via the Guided Entry of TA protein (GET) pathway (Abell *et al.*, 2004; 2007; Schuldiner *et al.*, 2008; Rabu *et al.*, 2008; Brkljacic *et al.*, 2009; Colombo and Fasana, 2011; Chio *et al.*, 2017; Costello *et al.*, 2017a; 2017b) and (iv) the recently discovered SRP-independent pathway (SND) that can partially substitute for the SRP and GET pathways (Aviram *et al.*, 2016; Casson *et al.*, 2017). The characteristic feature of the yeast ER proteins is a higher hydrophobicity of the TMD in comparison to the mitochondrial TA proteins, although this feature is not clearly different in human cells (Costello *et al.*, 2017a; 2017b). The sorting of TA proteins into peroxisomes is based on their recognition by the import receptor Pex19. The feature that defines peroxisomal proteins is the high net positive charge of the CTS (Yagita *et al.*, 2013; Costello *et al.*, 2017a; 2017b).

Considering the diversity of eukaryotes, studies of mitochondrial TA proteins are limited mainly to a few model organisms from the eukaryotic supergroup Opisthokonta (*S. cerevisiae* and mammalian cells) and the Plantae (*Arabidopsis thaliana*) (Abell and Mullen, 2011; Borgese and Fasana, 2011). More recently, the targeting of three mitochondrial TA proteins was investigated in *Toxoplasma gondii* of the Stramenopila/Alveolata/Rhizaria supergroup (Padgett *et al.*, 2017). Here, we decided to test the conservation of mitochondrial TA protein targeting in *Trichomonas vaginalis*, a member of Excavata supergroup that possesses hydrogenosomes (Hrdý *et al.*, 2008; Hampl *et al.*, 2009). These organelles produce hydrogen and ATP by substrate-level phosphorylation, but they lack respiratory chain complexes, the F_oF_1 ATP synthase, the citric acid cycle and other mitochondrial functions. Hydrogenosomes are bounded by a double membrane, as known for mitochondria; however, the inner membrane does not form cristae (Benchimol, 2009). A previous proteomic study of hydrogenosomal membranes revealed 70 putative membrane proteins, including core components of highly simplified TOM and TIM complexes (Rada *et al.*, 2011). TA proteins with known functions in mitochondria were entirely absent; however, 12 putative TA proteins were identified that seem to be unique to *T. vaginalis* (Rada *et al.*, 2011). In this study, we performed *in silico* searches for TA proteins in *T. vaginalis* genome, investigated the topology of seven selected TA proteins and tested the properties of the C-terminal domain that are decisive for the targeting of TA proteins to either hydrogenosomes or the ER. Our results demonstrated that the general mode of TA protein insertion is conserved across eukaryotic supergroups, including the excavate *T. vaginalis*, whereas specific features of TA proteins likely evolved after the split of the main eukaryotic lineages.

Table 1. Characteristics of C-terminal domains of hydrogenosomal TA proteins.

TrichoDB	Annotation	CTS length	TMD length	TMD sequence	Charge –10aa	Charge +10aa	Hydrophobicity
TVAG_090120	TA1, TPR	6	22	LYLILGIGAAIGVGAFCFYKIW	–2	4	1,73
TVAG_190830	TA2, Homp38, HSP20,TPR	6	19	IAIGGGIAAFVAGFSYAIY	1	4	1,67
TVAG_458060	TA3, HSP20,TPR	4	19	AVMAGLAIAGFAALAFSYM	1	4	1,86
TVAG_272350	TA4, HSP20,TPR	11	19	IGTMVAIGVGAGLATHWLI	0	5	1,56
TVAG_240680	TA5	16	22	PYLIIAIVVVIGVAVGLGFYF	4	1	2,48
TVAG_137270	TA6, HSP20,TPR	6	19	ILTGIAGAAVLVGAIIIIA	0	4	2,64
TVAG_277930	TA7, HSP20,TPR	7	22	ISKPLIVGGAVIAAGFLLYKGI	1	3	1,50
TVAG_283120	TA8	13	18	ITAVAVSVSICAAAYFLF	0	4	2,16
TVAG_174010	TA9	6	22	FSTIIGLGITVGVIVGVVVALY	2	3	2,29
TVAG_369980	TA10	6	19	YNKFWGIFSVVAFFGVIIF	-1	1	1,62
TVAG_211970	TA12	3	22	STYVIAGTAVLAASAAAFLFFS	3	3	1,64

Results

Physicochemical characteristics of the C-terminal domains of hydrogenosomal TA proteins

To analyze the targeting signal of hydrogenosomal TA proteins, we first compared the C-terminal domains of TA proteins identified in the proteome of the hydrogenosomal membranes (Rada *et al*., 2011). The predicted TMD using the TMHMM server confirmed the presence of α-helixes of 18–22 AA residues close to the C-terminus with a mean hydrophobicity of 1.95 (Table 1). The length of the CTS was short, ranging from 3 to 16 AA residues. The mean positive net charge calculated for 10 AA residues flanking the TMD at the N-terminus was slightly positive (0.82) with a wide range from –2 to 4 (Table 1). The C-terminal flanking regions were considerably more enriched in the net positive-charged residues, with a higher mean charge (3.27, range 1–5). Inspection of the latter flanking regions showed that they all contained the dibasic motif R/K-R/K or longer stretches of up to four basic AA (i.e., KRRK, RKKK) (Table 1). Next, we searched for proteins with a single C-terminal α-helix and CTS up to 30 AA in the *T. vaginalis* protein database, which identified 1452 proteins (Table S2). Then, we applied criteria based on the characteristics of hydrogenosomal TA proteins listed in Table 1 that includes: hydrophobicity ≤2.64, a net positive charge within 10 AA on the CTS side ≥1, CTS is longer than 2 AA residues, and the protein possessing a dibasic motif K/R/H-K/R/H within 10 AA of the CTS. In addition, all proteins with the predicted secretory signal and the mitochondrial presequence were excluded. This approach led to a set of 120 predicted hydrogenosomal TA proteins (Table S3). Most of the predicted hydrogenosomal TA candidates (75%) are conserved hypothetical proteins. The protein domain predictions revealed that five TA proteins previously identified in the hydrogenosomal membrane proteome (TA2, 3, 4, 6 and 7) (Rada *et al*., 2011) and three additional proteins that were predicted *in silico* (TA13, 14, 15) (Table S3) contain an N-terminal Hsp20-like domain followed by tetratricopeptide repeats (TPRs).

Localization and topology of the putative hydrogenosomal TA proteins

To investigate the cellular localization and the topology of the putative hydrogenosomal TA proteins, the genes for TA4, TA5, TA7, TA8, TA10 and TA11, which were previously identified in the membrane proteome (Rada *et al*., 2011), and the predicted protein TA16, were episomally expressed in the *T. vaginalis* T1 strain with an HA tag at the N-terminus. Immunofluorescence microscopy confirmed that six of the seven TA proteins localized

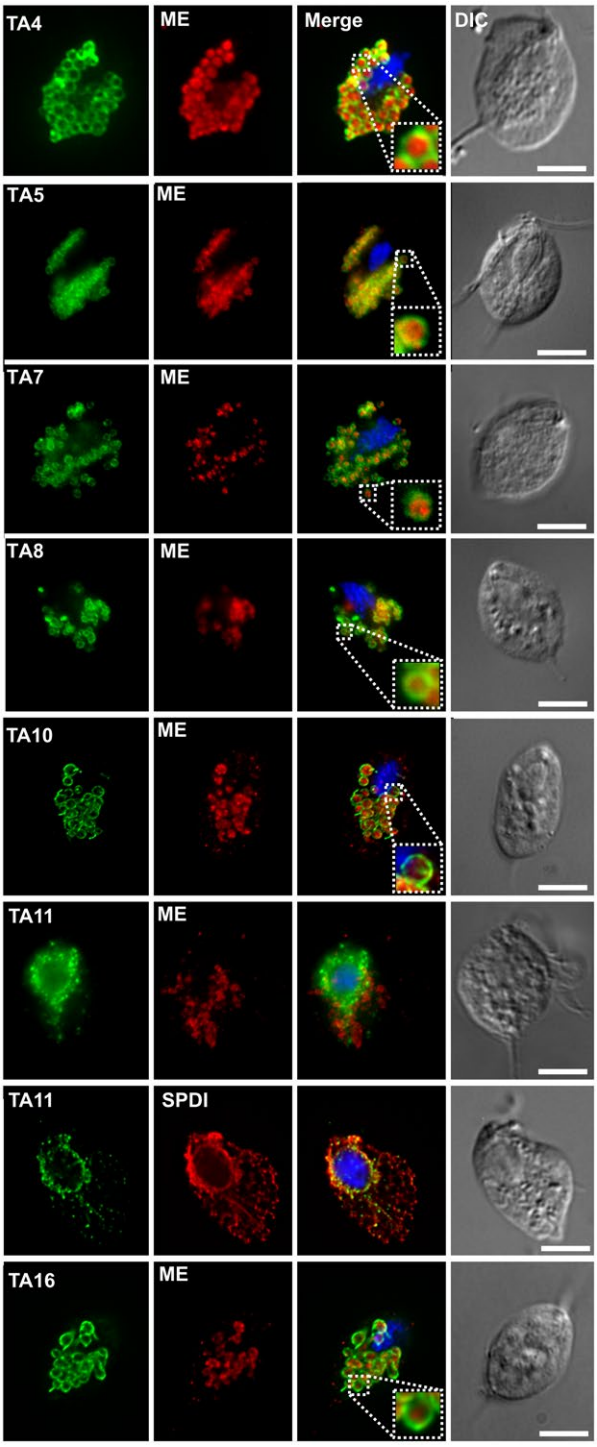
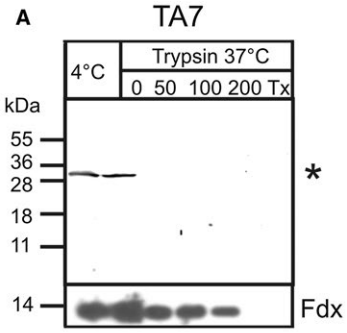


Fig. 1. Localization of TA proteins with N-terminal HA tag in *T. vaginalis*. TA proteins were expressed with an N-terminal HA tag in trichomonads and visualized with a mouse monoclonal anti-HA antibody (green). The hydrogenosomal marker malic enzyme (ME) was detected with a polyclonal rabbit anti-malic enzyme antibody (red). ER marker SPDI was expressed in trichomonads with a C-terminal V5 tag and detected with a rabbit monoclonal anti-V5 antibody (red). The nucleus was stained with DAPI (blue). DIC, differential interference contrast. The scale bar represents 5 μm.

to hydrogenosomes and appeared as rings decorating the hydrogenosomal membranes. Malic enzyme was used as a marker of the hydrogenosomal matrix. Only TA11 was not observed in the hydrogenosomes and was localized mostly to structures surrounding the nucleus that colocalized with soluble protein disulfide isomerase (SPDI), an ER marker (Fig. 1). Weak TA11 signal was also observed in vesicular structures scattered in the cytosol (Fig. 1). Therefore, T11 was most likely a contaminant from the ER in the previous study of the hydrogenosomal proteome (Rada *et al*., 2011). Next, we investigated the topology of the hydrogenosomal TA proteins. Initially, we used a polyclonal antibody raised against TA7 that recognized the complete TA7 in intact isolated hydrogenosomes (Fig. 2A). Treatment of the hydrogenosomes with trypsin completely erased the signal indicating that the N-terminus of TA7 is facing the cytosol. However, the antibody did not allow us to visualize the short membrane-protected part of TA7. Therefore, we decided to express TA7 and other six hydrogenosomal TA proteins with a C-terminal hemagglutinin (HA) tag. Although, the C-terminal tag has been shown to interfere with the correct targeting of some TA proteins to mitochondria (Horie *et al*., 2002), we found that all tested hydrogenosomal TA proteins remained associated with the hydrogenosomal membrane except TA5 that appeared mostly in the cytosol (Fig. S1). Therefore, five TA proteins including TA7 with the hydrogenosomal localization using both N- and C-terminal HA tag were used for protein protection assay. The addition of trypsin to the hydrogenosomes isolated from each transfected strain resulted in a shift of protein mobility from the size corresponding to the complete non-cleaved recombinant protein to a smaller size of the membrane-protected C-terminal domain that includes the TMD, CTS and HA tags (Fig. 2B). The experimental sizes of the protected domains visualized on immunoblots were slightly higher (~4 kDa) than the theoretical sizes calculated from the closest lysine or argine to the TMD from the cytosolic side. TA8 and TA10 were highly sensitive to proteolysis and were partially cleaved at any conditions. The C-terminal domains were degraded only when Triton X-100 was added to solubilize the lipid bilayers of the hydrogenosomes. PFO, a hydrogenosomal matrix protein was used as a control to assess the membrane intactness during the trypsin treatment. The protein protection assays clearly confirmed that all the tested proteins localized to the outer hydrogenosomal membrane (OHM) with the N-terminal domain facing the cytosol, the property that defines TA proteins. To obtain more accurate information about the protein distribution in the OHM, we randomly selected TA10 that had appeared as a ring under standard confocal microscopy for visualization by STED microscopy (Fig. 3). This



C

TA protein	Full size kDa	Protected peptides cal.+ kDa	exp. kDa
TA4	50	7	11
TA7	37	7	11
TA8	18	7	10
TA10	21	6	10
TA16	37	6	10

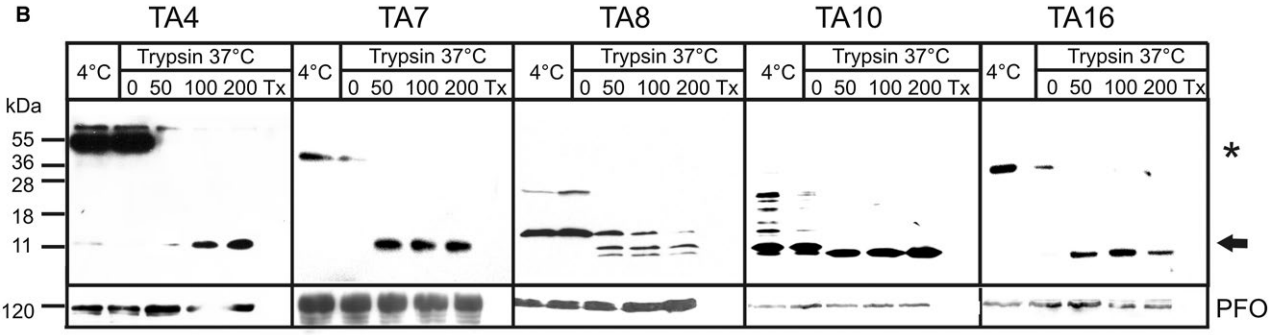


Fig. 2. Topology of hydrogenosomal TA proteins tested by the protein protection assay.
A. Hydrogenosomes were isolated from *T. vaginalis* cells and incubated for 30 min at 4°C and 37°C with 0, 50, 100 or 200 µg/ml trypsin or with trypsin and 0.5% Triton X-100 (Tx). TA7 were analyzed by SDS-PAGE and immunoblotting using a rat polyclonal anti-TA7 antibody. The ferredoxin (Fdx, used as a control matrix protein) was visualized by polyclonal rabbit anti-Fdx1 antibody.
B. Hydrogenosomes were isolated from *T. vaginalis* cells expressing TA proteins with a C-terminal HA tag. The hydrogenosomes were incubated for 30 min at 4°C and 37°C with 0, 50, 100 or 200 µg/ml trypsin or with trypsin and 0.5% Triton X-100 (Tx). TA proteins (asterisk) and their protected domains (arrow) were analyzed by SDS-PAGE and immunoblotting using a mouse monoclonal anti-HA antibody. Pyruvate ferredoxin oxidoreductase (PFO, used as a control matrix protein) was visualized by a mouse monoclonal anti-PFO antibody.
C. Molecular weight of complete TA proteins including HA tag and membrane protected peptides. Cal., calculated molecular weight of protected peptides. The weight was calculated from the closest trypsine clevage site (arginine or lysisine) to N-terminus of TMD up to end of CTS.

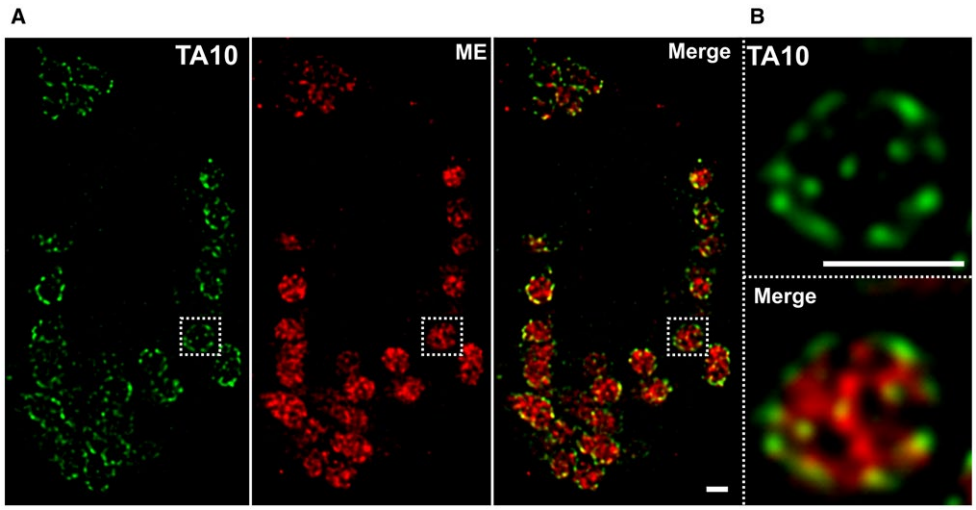


Fig. 3. Distribution of TA10 in the OHM.
A. TA10 fused with an N-terminal HA tag was expressed in *T. vaginalis* and visualized by STED super-resolution microscopy using a mouse monoclonal anti-HA antibody (green). The hydrogenosomal marker malic enzyme (ME) was detected with a polyclonal rabbit anti-malic enzyme antibody (red).
B. The hydrogenosome in detail. TA10 was detected in distinct spots in the outer hydrogenosomal membrane. The scale bar represents 0.5 µm.

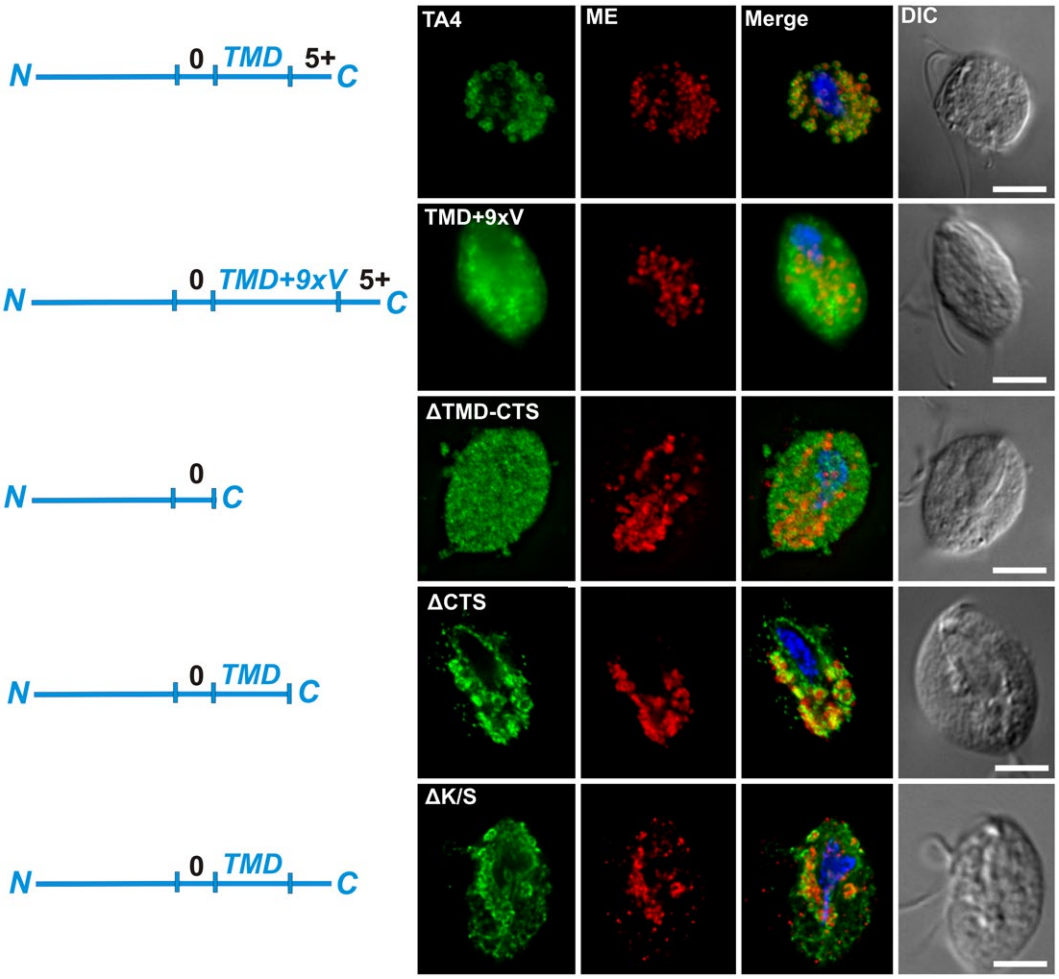


Fig. 4. Mutations in the tail-anchored domains cause the redistribution of TA4 to the cytosol and ER. Various versions of TA4 were expressed with an N-terminal HA tag in trichomonads and visualized with a mouse monoclonal anti-HA antibody (green). The hydrogenosomal marker malic enzyme (ME) was detected with a polyclonal rabbit anti-malic enzyme antibody (red). The nucleus was stained with DAPI (blue). DIC, differential interference contrast. TMD, transmembrane domain. The numbers 0 and 5+ represent the overall charges within the left and right TMD-flanking regions. 9xV, insertion of 9 residues of valine. CTS, C-terminal sequence. Δ, deletion. ΔK/S, lysine amino acid residues were exchanged with serine. The scale bar represents 5 μm.

approach revealed that the protein is not distributed evenly in the OHM, but is present in distinct spots.

TMD length and C-terminal net positive charge are critical for the hydrogenosomal localization of TA4

To test the role of C-tail anchor domains in protein targeting to hydrogenosomes, we developed a series of mutations in TA4 (Fig. 4). First, we tested whether the length of the TMD is critical for protein accommodation within the OHM. The TMD was extended by introducing 9 valines, which are small non-polar, uncharged amino acids. We introduced 9 valines as a previous study had showed that introduction of 2–7 valines gradually decreased efficiency of TA protein targeting to OMM, however, the targeting was not completely abolished (Horie *et al.*, 2002). TA4 with an extended TMD by 9 valines localized to the cytosol without any labeling of the hydrogenosomes (Fig. 4). Next, we deleted the TMD together with the positively charged CTS. As expected, this deletion abolished the hydrogenosomal targeting of the protein, and the protein localized to the cytosol. However, when we deleted only the CTS (ΔCTS, 11 AA), this mutation resulted in the dual localization of TA4-ΔCTS to both hydrogenosomes and structures corresponding to ER morphology (Figs 4 and 5). A similar effect was found when we replaced five positively charged lysine residues with serine residues. Moreover, both electron and structured illumination microscopy revealed that the expression of TA4-ΔCTS was associated with an unusual ER morphology (Fig. 5). While the ER formed typical flattened sacs around the nucleus in

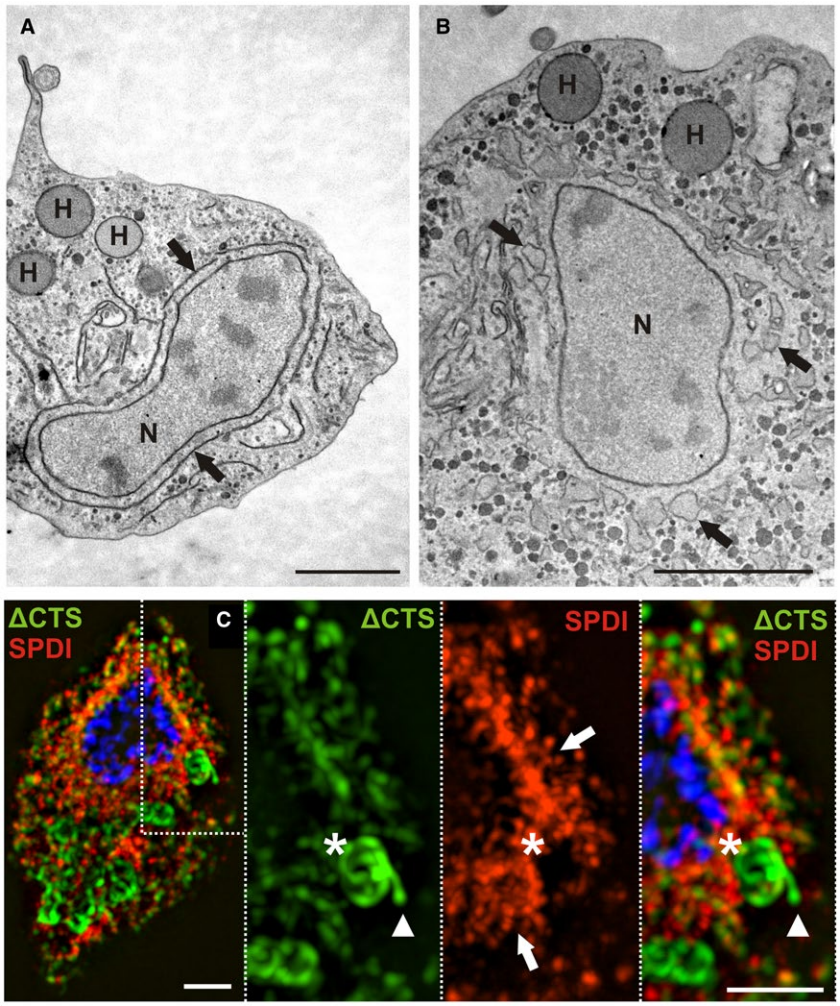


Fig. 5. Dissipation of the ER upon expression of TA4 without its CTS.
(A–B) Transmission electron microscopy of *T. vaginalis* expressing TA4 and TA4 without its CTS (ΔCTS), both with an N-terminal diHA tag. H, hydrogenosome; N, nucleus. The black arrow indicates membranes of the endoplasmic reticulum. The scale bar represents 1 μm.
C. SIM of *T. vaginalis* expressing N-terminally HA-tagged TA4-ΔCTS (green) and C-terminally V5-tagged ER marker SPDI (red). The expressed proteins were detected with a mouse monoclonal anti-HA antibody (green) and a rabbit monoclonal anti-V5 antibody (red). The nucleus was stained with DAPI (blue). The white arrows indicate the ER. Asterisks mark areas of proximity between hydrogenosomes and the ER. Arrowheads mark an extension of the hydrogenosomal membrane. The scale bar represents 1 μm.

Trichomonas cells (Benchimol, 2008) and in the cells expressing TA4, the ER in the cells expressing TA4-ΔCTS was dissipated into multiple vesicles surrounding the nucleus (Fig. 5). In cells expressing TA4-ΔCTS, we also observed an extension of the membranes in some hydrogenosomes (Fig. 5).

ER proteins with a single C-terminal TMD domain

The ER localization of the mutated TA4 protein prompted us to investigate the targeting signals of both hydrogenosomal and ER proteins containing a single C-terminal TMD that mediate the organelle-specific localization of TA proteins. First, we filtered all proteins with a predicted ER localization in our dataset of proteins with a C-terminal TMD (Table S2). Altogether, we identified 54 proteins with a TMD of 18–23 AA residues and CTS up to 30 AA residues (Table S4). According to TargetP predictions, 20 proteins possessed signal peptides (SP) that likely target these proteins to ER as type I transmembrane proteins (Goder and Spiess, 2001), and 34 proteins were devoid of a predictable SP. The latter proteins might be considered as putative TA proteins in ER. However, absence of predictable SP needs to be considered with caution. Our previous study revealed that the prediction of SP has limited reliability for *T. vaginalis* protein sequences as it produces false negative results frequently (Štáfková *et al.*, 2018). The physicochemical characteristics of the C-terminal domains of putative TA proteins and

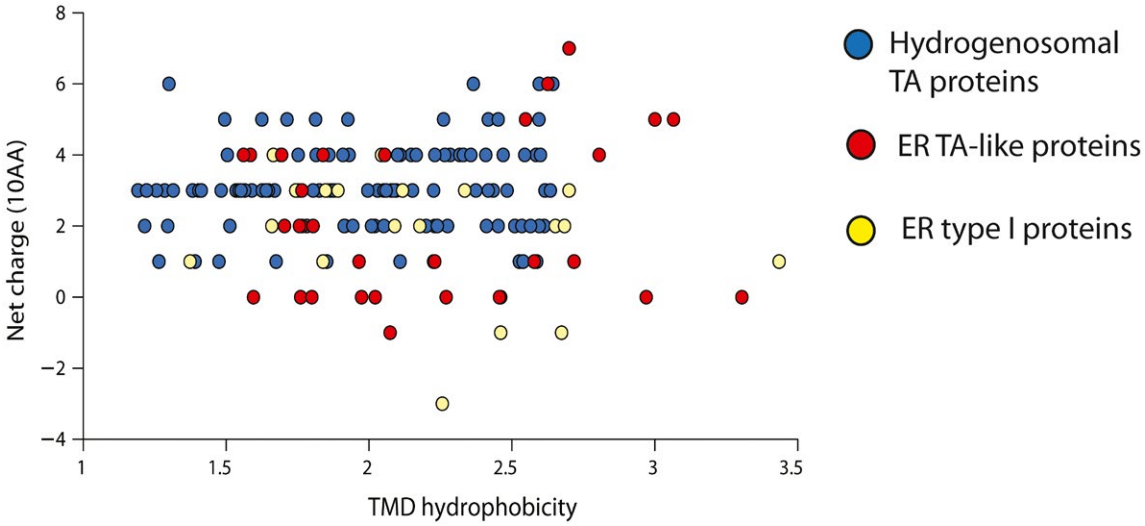


Fig. 6. Comparison of physicochemical parameters of predicted hydrogenosomal TA proteins, and ER proteins with a single C-terminal TMD domain. Hydrogenosomal proteins are in blue circels, ER proteins without predictable SP (ER TA protein) are in red circle, and type I transmembrane proteins with SP are in yellow circle.

type I transmembrane proteins appeared rather similar (Fig. 6). Both groups of ER proteins revealed a similar average hydrophobicity (~2.18) with a wide range of values (Fig. 6 and Table S4). The average positive net charges calculated for the flanking 10 AA residues in the N-terminal region of TA proteins and type I transmembrane proteins were 0.82 and 0.40, respectively, and in the C-terminal region were 1.97 and 1.85 respectively (Table S4).

As we were not confident with the prediction of *bona fide* TA proteins in ER, for topology studies, we selected protein disulfide isomerase named TAPDI (TVAG_255840) that possesses predicted an SP at the N-terminus and TA-like structure at the C-terminus. We expected that if the predicted SP is deleted, truncated TAPDI may act as a TA protein. The expression of TAPDI in *T. vaginalis* confirmed its localization in the ER (Fig. 7). SPDI was co-expressed as an ER marker that possesses an N-terminal SP and a C-terminal ER retention signal KQEL (Pagny *et al.*, 2000). The protein protection assay revealed that the N-terminal domain of TAPDI is protected against trypsin treatment by the ER membrane, which is consistent with the type I protein topology (Fig. 7B). The deletion of TMD-CTS had no effect on TAPDI targeting to the ER (Fig. 7). When, 25 N-terminal AA (ΔN-25AA) of the SP were deleted, TAPDI was still associated with the ER and the protein remained protected against trypsin treatment (Fig. 7). To investigate the topology of ΔN-25AA-TAPDI in the ER membrane, we prepared a double transfectant expressing ΔN-25AA-TAPDI and SPDI as a control. Immunoblotting of the cellular fractions confirmed that a significant part of ΔN-25AA-TAPDI is associated with the ER-enriched fraction; however, the signal disappeared after trypsin treatment, whereas trypsin has no effect on the control SPDI (Fig. 7B). These results suggest that TAPDI with deleted SP acts as a TA protein that is targeted to the ER membrane with the N-terminal domain facing the cytosol.

TMD and its flanking regions are critical for organelle-specific targeting

Next, we prepared a series of chimeric proteins by swapping domains between hydrogenosomal TA4 and TAPDI to investigate the role of the charged domains flanking the TMD, the TMD domain and the SP. TA4 and TAPDI have TMD lengths of 20 and 22 AA residues, respectively, with similar TMD hydrophobicity values of 1.62 and 1.74 respectively. Net charge of TMD flanking domain (10 AA) at the N-terminus was 0 and –1, and at the C-terminus +5 and +3 for TA4 and TAPDI respectively. When we fused the N-terminal portion of TA4 with the C-terminal domains of TAPDI, including the TMD and both flanking regions, the chimeric protein was targeted to the ER (Fig. 8A). This chimeric protein was not protected from trypsin treatment (Fig. 8B), as observed in the case of TAPDI without the signal peptide (Fig. 7B). Localization to the ER was also observed for the same chimeric protein that contained a region with a net neutral charge flanking the TMD at the N-terminus (Fig. 8A). However, the TA4 protein that contained only the CTS from TAPDI with a net charge of 3 was targeted to hydrogenosomes. The exchange of only the TMD had no effect on TA4 targeting to hydrogenosomes, which is consistent with similar length and hydrophobicity of TMD in TA4 and TAPDI (Fig. 8A). However, the TMD

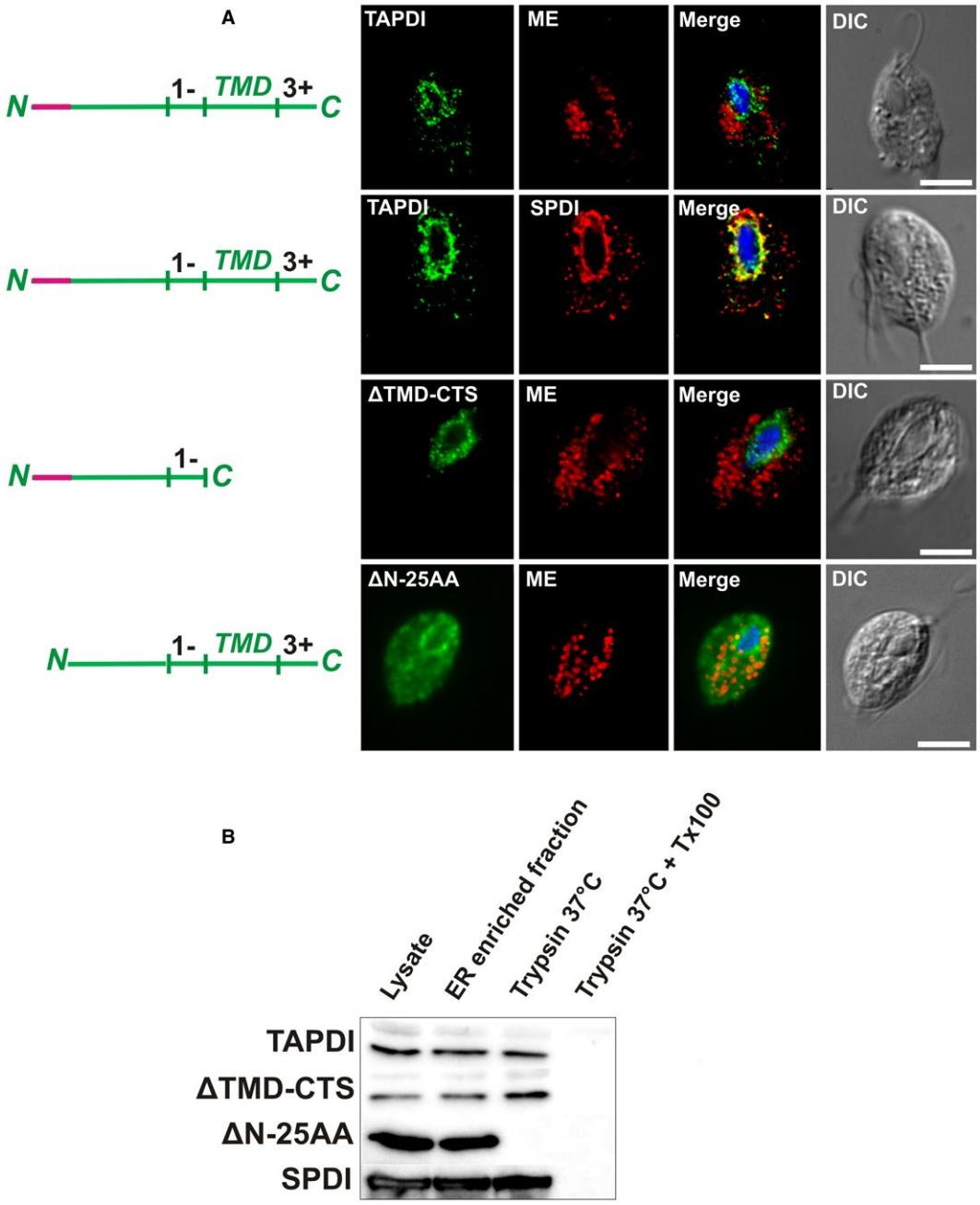


Fig. 7. The role of signal peptide and C-tail anchored domain for topology of TAPDI.
A. Immunofluorescence microscopy. Various versions of TAPDI were expressed with an N-terminal HA tag in trichomonads and visualized with a mouse monoclonal anti-HA antibody (green). The hydrogenosomal marker malic enzyme (ME) was detected with a polyclonal rabbit anti-malic enzyme antibody (red). ER marker SPDI was expressed in trichomonads with a C-terminal V5 tag and detected with a rabbit monoclonal anti-V5 antibody (red). The nucleus was stained with DAPI (blue). DIC, differential interference contrast. TMD, transmembrane domain. The numbers 1– and 3+ represent the overall net charges within the left and right TMD-flanking regions. CTS, C-terminal sequence. ΔTMD-CTS, deletion of the TMD and CTS domains. ΔN-25AA, the first 25 amino acids were deleted from the TAPDI N-terminus.
B. Protein protection assay. ER-enriched fractions were isolated from trichomonads expressing various TAPDI versions and treated with trypsin or with trypsin and Triton X-100 (Tx100). SPDI (ER marker) was expressed in trichomonads with a C-terminal V5 tag. Samples were analyzed by immunoblotting using a monoclonal mouse anti-HA antibody and a monoclonal rabbit anti-V5 antibody. The scale bar represents 5 μm.

in the context of the charged region flanking the TMD at the C-terminus was decisive for the relocation of the protein to the ER.

Conversely, we constructed a chimeric protein that consists of an N-terminal domain of TAPDI and a C-terminal domain of TA4. Surprisingly, although this chimera

contains a SP, it appeared exclusively in hydrogenosomes. However, the extension of the TAPDI domain with the flanking region at the N-terminal end of the TMD (net charge –1) caused the relocation of the protein to the ER. Localization to the ER was also observed for TAPDI with only the CTS derived from TA4. Thus, in this series of chimeric proteins, the charged region flanking the TMD at the N-terminus was decisive for ER or hydrogenosomal localization.

Discussion

The investigation of hydrogenosomal TA proteins in *T. vaginalis* revealed that the general properties required for their specific targeting to the OHM are similar to those defined for mitochondrial TA proteins: (i) the TA proteins possess a single C-terminal TMD of a defined length, (ii) the TMD is flanked at the N-terminus, C-terminus or both termini by basic residues and (iii) the TMD domain is of moderate hydrophobicity (Isenmann *et al.*, 1998; Kuroda *et al.*, 1998; Borgese *et al.*, 2001; Motz *et al.*, 2002; Kaufmann *et al.*, 2003).

Despite these general properties, there are more subtle characteristics of mitochondrial/hydrogenosomal TA proteins that seem to be lineage-specific and that are related to their sorting to different cellular compartments. The TMD of hydrogenosomal TA proteins and single-spanning ER proteins (TA and type I proteins) in *T. vaginalis* appear to be similar, mostly consisting of 23 AA residues. In contrast, the length of the mitochondrial TMD tends to be shorter (<20 AA residues) (Hwang *et al.*, 2004). The net positive charge of the C-terminal region flanking the TMD is considerably higher in hydrogenosomal TA proteins than in proteins targeted to the *T. vaginalis* ER. We have demonstrated that the net charge difference is a critical property for targeting of hydrogenosomal TA proteins to the correct destination. A similar difference in the net positive charges of mitochondrial TA proteins (mean 1.1) and ER proteins (mean 0.2) was observed in mammalian and yeast cells (Costello *et al.*, 2017a; 2017b). However, the net positive charges are considerably higher in *T. vaginalis* than in other cells. This difference might be explained by the different repertoires of target organelles. In addition to mitochondria and the ER, TA proteins in yeast and mammalian cells are also targeted to peroxisomes or to both mitochondria and peroxisomes. The net positive charge of the tail regions of these two types of TA proteins is in the range of 2.5–6 (Costello *et al.*, 2017a; 2017b). As *T. vaginalis* lack peroxisomes, a higher net positive charge for hydrogenosomal TA proteins does not interfere with peroxisomal localization and might be employed for hydrogenosomal targeting.

Interestingly, we identified a type I ER protein, TAPDI, with an N-terminal SP, a C-terminal anchor domain and a low net positive charge in the flanking regions. Classical PDIs are soluble proteins and PDI8 that has been identified in terrestrial plants is the only known PDI with TMD close to C-terminus (Yuen *et al.*, 2016). We showed that TAPDI protein is inserted into the ER membrane with an N-terminal domain facing the ER lumen. It could be expected that the TAPDI SP is recognized by the SRP and delivered to the ER translocation machinery for co-translational insertion into the ER (Higy *et al.*, 2004). However, when the C-terminal domain of TAPDI was replaced with a hydrogenosomal C-terminal domain, the chimeric protein appeared in hydrogenosomes. Therefore, the hydrogenosomal signal embedded in the higher net charge of the flanking regions was dominant over the ER signal sequence. This result suggests that TAPDI is inserted into the ER post-translationally and that the import system that delivers TAPDI to the ER may compete with unidentified chaperones that deliver the protein to hydrogenosomes based on net positive charge recognition. In mammalian cells, a high net positive charge in the tail of peroxisomal TA proteins was shown to promote their interaction with the peroxisomal import receptor Pex19 and their delivery to peroxisomes (Costello *et al.*, 2017a; 2017b). Thus, a similar protein-protein interaction might be expected to support the specific delivery of hydrogenosomal TA proteins. Alternatively, hydrogenosomes might represent the primary site that sequesters TA proteins with positively charged residues without the contribution of any cytosolic factors, whereas targeting to the ER is dependent on the recognition of a targeting signal by ER pathways, such as the GET pathway. This possibility is supported by the observation in yeast that a subset of TA proteins mislocalized to mitochondria upon the deletion of GET pathway components (Schuldiner *et al.*, 2008). Although, the GET system has not been studied in *T. vaginalis*, genes encoding putative GET components are present in the *T. vaginalis* genome (TrichoDB). Noteworthy, when SP of TAPDI was deleted, the protein was partially inserted to ER membrane from the cytosolic side mimicking a TA protein. The tantalizing question is: how did proteins with an N-terminal SP and a C-terminal TA-like structure evolve? We can speculate that TAPDI evolved from a soluble PDI version and TMD was added later during the evolution to gain new specific function in ER lumen. It has been suggested that plant PDI8 with a C-terminal TMD may play a role in protein folding as they translocate across ER membrane (Yuen *et al.*, 2016). However, we cannot exclude an alternative possibility that some ER type I proteins served originally in the cytosol and at cytosolic side of ER membrane as TA proteins and later gained SP to be targeted to ER lumen.

Mutation of the C-terminal flanking region of the hydrogenosomal protein TA4 showed the critical importance

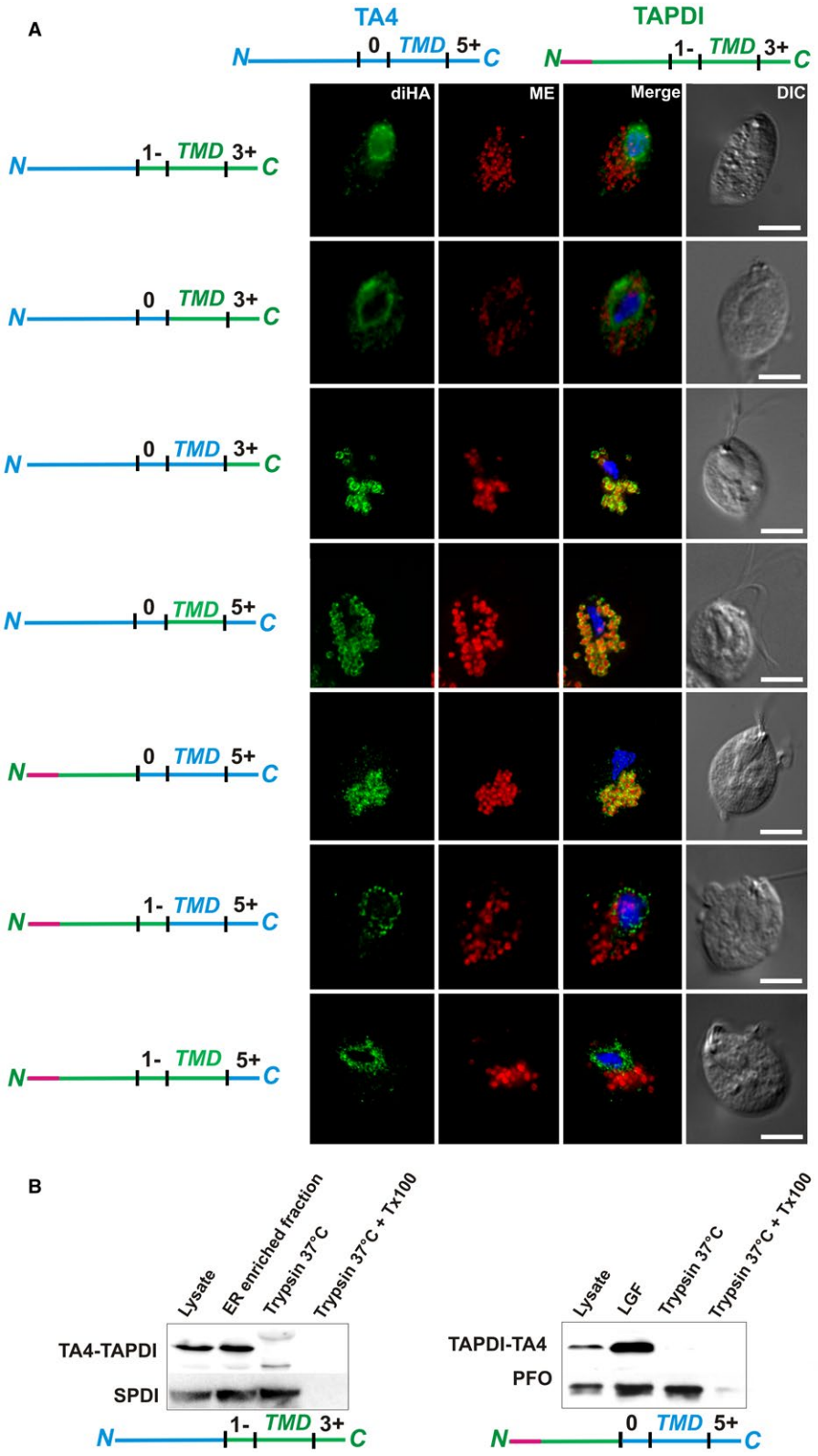

A
TA4
N 0 TMD 5+ C
TAPDI
N 1- TMD 3+ C
diHA
ME
Merge
DIC
N 1- TMD 3+ C
N 0 TMD 3+ C
N 0 TMD 3+ C
N 0 TMD 5+ C
N 0 TMD 5+ C
N 1- TMD 5+ C
N 1- TMD 5+ C
B
Lysate
ER enriched fraction
Trypsin 37°C
Trypsin 37°C + Tx100
TA4-TAPDI
SPDI
N 1- TMD 3+ C
Lysate
LGF
Trypsin 37°C
Trypsin 37°C + Tx100
TAPDI-TA4
PFO
N 0 TMD 5+ C

Fig. 8. Effect of C-tail anchored domain swapping between TA4 and TAPDI.
A. Immunofluorescence microscopy. Various chimeric versions of TA4 and TAPDI were expressed with an N-terminal HA tag in trichomonads and visualized with a mouse monoclonal anti-HA antibody (green). The hydrogenosomal marker malic enzyme (ME) was detected with a polyclonal rabbit anti-malic enzyme antibody (red). The nucleus was stained with DAPI (blue). DIC, differential interference contrast. TMD, transmembrane domain. The numbers 1–, 0, 3+ and 5+ represent the overall charges within the left and right TMD-flanking regions.
B. Protein protection assay. ER-enriched and hydrogenosome-enriched fractions were isolated from transfected trichomonads and treated with trypsin or with trypsin and Triton X-100 (Tx100). SPDI (ER marker) was expressed in trichomonads with a C-terminal V5 tag. Samples were analyzed by immunoblotting using a monoclonal mouse anti-HA antibody (green) and a monoclonal rabbit anti-V5 antibody (red). PFO (pyruvate ferredoxin oxidoreductase), a hydrogenosomal matrix protein, was detected by a mouse monoclonal anti-PFO antibody. LGT, large granule fraction. The scale bar represents 5 μm.

of the net positive charge for its correct localization to hydrogenosomes. The reduction of the net positive charge in the CTS or the deletion of the CTS resulted in the dual localization of TA4 to hydrogenosomes and the ER. A similar dual localization to mitochondria and the ER was observed for cytochrome b_5 with a mutated CTS in mammalian cells (Kuroda *et al.*, 1998; Borgese *et al.*, 2001; Henderson *et al.*, 2007; Costello *et al.*, 2017a; 2017b). Of note, we observed that the mistargeting of TA4 caused swelling and vacuolization of the ER and the formation of membrane extensions in hydrogenosomes. These changes remind the response of the mammalian ER to stress conditions, which is interconnected with the response of the mitochondrial morphology (Urra and Hetz, 2012; Vannuvel *et al.*, 2013) and highlights the fundamental importance of the correct sorting of TA proteins for cell physiology.

The overall proteome of *T. vaginalis* hydrogenosomes is considerably reduced in comparison to that of mitochondria (Rada *et al.*, 2011; Schneider *et al.*, 2011). The absence of typical mitochondrial pathways most likely reflects an adaptation of this parasite to anaerobic environments (Hrdy *et al.*, 2004). However, the number of TA proteins in the OHM seems to be comparable to or even larger than that observed in mitochondria. Our *in silico* predictions identified 120 putative hydrogenosomal TA proteins in *T. vaginalis.* This number is similar to the 161 mitochondrial proteins predicted in *A. thaliana* (Kriechbaumer *et al.*, 2009). Human, yeast and *T. gondii* mitochondria possess ~3–12 mitochondrial TA proteins; however, the total number is not known (Beilharz *et al.*, 2003; Kalbfleisch *et al.*, 2007; Padgett *et al.*, 2017). The function of the hydrogenosomal TA proteins remains elusive. We did not find any homolog of mitochondrial TA proteins that are known to be involved in the stabilization of the TOM complex, electron transport (Kuroda *et al.*, 1998), organelle fission (Kemper *et al.*, 2008) or apoptosis (Kaufmann *et al.*, 2003). However, we found seven proteins with an N-terminal Hsp20-like domain followed by array of TPR domains. A similar domain structure was recently found in the mitochondrial TA protein ATOM69 in *Trypanosoma brucei* (Mani *et al.*, 2015), which belongs to the eukaryotic supergroup Excavata with *T. vaginalis*. ATOM69 serves as a specific receptor for the divergent TOM complex in trypanosomes. Thus, we can speculate that hydrogenosomal TA proteins with Hsp20/TPR domains may play roles similar to ATOM69 in *T. vaginalis.*

Collectively, this is the first study that has focused on the insertion of TA proteins into the outer membrane of hydrogenosomes, an unusual form of mitochondria that are present in some anaerobic protists. Moreover, *T. vaginalis* is the only member of the eukaryotic supergroup Excavata in which the mode of TA protein insertion has been explored to date. Despite the evolutionary distance of excavates and changes in the hydrogenosomal proteome that have occurred due to adaptation to anaerobiosis, hydrogenosomal TA proteins share common characteristics with proteins inserted to mitochondria of other eukaryotes that belong to the supergroup Opisthokonta (fungi and mammalian cells), the Plantae (*Arabidopsis thaliana*) and *Toxoplasma gondii* of the Alveolate lineage. This finding strongly suggests that TA proteins with similar properties of the C-terminal domain were present in the last common eukaryotic ancestor. However, the unique primary structure of the N-terminal domains of hydrogenosomal proteins without detectable homology to proteins with known function in the other eukaryotic supergroups indicates that they evolved independently, and their lineage-specific functions remain to be established.

Experimental procedures

Cell cultivation

T. vaginalis strain T1 (provided by J. H. Tai at the Institute of Biomedical Sciences in Taipei, Taiwan) was grown in Diamond's tryptone-yeast extract-maltose (TYM) medium supplemented with 10% (vol/vol) heat-inactivated horse serum (Diamond, 1957).

Gene cloning and transformation

Genes that encode TA proteins TVAG_272350 (TA4), TVAG_240680 (TA5), TVAG_277930 (TA7), TVAG_283120 (TA8), TVAG_369980 (TA10), TVAG_393390 (TA11), TVAG_290590 (TA14) and TVAG_069740 (TA16), and tail-anchored protein disulfide isomerase (TAPDI, TVAG_255840) were amplified by PCR from *T. vaginalis* genomic DNA. Chimeric constructs based on the TA4 and TAPDI genes were designed and amplified with primers

according to Table S1. The amplified genes were cloned into the plasmids pTagVag (Šuťák *et al*., 2004) and pTagVag-N-HA-Neo, enabling the expression of recombinant proteins with C-terminal and N-terminal dihemagglutinin (HA) tags under the control of α-subunit succinyl-coenzyme A synthetase promotor, respectively. pTagVag-N-HA-Neo was derived from pTagVag. The 300 bp promotor sequence was amplified from pTagVag and fused by PCR with an oligonucleotide sequence encoding the HA tag that was synthesized by Sigma-Aldrich. The amplified cassette with the promotor and the HA tag was cloned into pTagVag via SacII and NdeI restriction sites. The primers are listed in Table S1.

Plasmids with subcloned genes were electroporated into trichomonads and selected with geneticin (200 µg/ml), as previously described (Šuťák *et al*., 2004). In double-labeling experiments, trichomonads were cotransfected with two plasmids: pTagVag-N-HA-Neo expressing TAPDI with a deleted 25 N terminal AA or chimeric constructs and pTagVag-V5-Pur (Štáfková *et al*., 2018) expressing SPDI (TVAG_267400) with a V5 tag and the KQEL sequence at the C-terminus as an ER control. Both plasmids were electroporated into trichomonads, and the transformants were selected in the presence of geneticin (200 µg/ml) and puromycin (40 µg/ml) (Štáfková *et al*., 2018). The sequences of the PCR primers are listed in Table S1.

Preparation of cellular fractions

The cells were harvested and homogenized by sonication, and the cellular fractions were separated by differential centrifugation (hydrogenosome-enriched fraction) and Percoll gradient centrifugation (hydrogenosomes), as described previously (Šuťák *et al*., 2004). The cytosolic fraction was isolated by centrifugation of the total cell lysates at 20,000 × g and subsequent centrifugation at 190,000 × g.

The ER-enriched fraction was isolated from trichomonads as follows. The cells were disrupted by sonication at an amplitude of 40, with a pulse of 1 s and a duration of 3 × 1 min on ice. Unbroken cells and cell debris were removed by centrifugation at 500 × g and 900 × g respectively. The supernatant was then centrifuged at 3,000 × g, and the collected pellet (ER-enriched fraction) was washed three times in isolation medium (225 mM sucrose, 20 mM KCl, 10 mM KH_2PO_4, 5 mM $MgCl_2$, 1 mM EDTA and 20 mM Tris-HCl, pH 7.2).

Protease protection assay

Aliquots of Percoll-purified intact hydrogenosomes (3 mg) were resuspended in 1 ml of ST buffer (250 mM sucrose, 10 mM Tris, pH 7.4, 0.5 mM KCl, 50 µg/ml TLCK and 10 µg/ml leupeptin). The ER-enriched fraction (3 mg) was resuspended in isolation medium (225 mM sucrose, 20 mM KCl, 10 mM KH_2PO_4, 5 mM $MgCl_2$, 1 mM EDTA and 20 mM Tris, pH 7.2) with 50 µg/ml TLCK and 10 µg/ml leupeptin. Trypsin was added to a final concentration of 50–200 µg/ml, and the samples were incubated on ice or in a water bath at 37°C for 30 min. After incubation, soybean trypsin inhibitor was added (5 mg/ml) and the samples were analyzed by immunoblotting using a rat polyclonal anti-TA7antibody (BIOCEV, Czech Republic), rabbit polyclonal anti-ferredoxin antibody (a kind gift from Patricia Johnson at the University of California Los Angeles, USA), mouse monoclonal anti-HA antibody (Exbio, Czech Republic), a mouse monoclonal anti-pyruvate:ferredoxin oxidoreductase antibody (a kind gift from Guy Brugerolle at the University of Clermont Ferrand in France) and a rabbit monoclonal anti-V5 antibody (Abcam, Cambridge, UK).

Electron microscopy

Trichomonads were centrifuged at 3,000 × g for 10 min and fixed in 2.5% glutaraldehyde and 5 mM $CaCl_2$ diluted in 0.1 M cacodylate buffer, pH 7.2, overnight at 4°C. The pellets were then washed in cold PBS, pH 7.2 and post-fixed in 0.1 M cacodylate buffer containing 1.6% ferricyanide, 10 mM $CaCl_2$ and 2% OsO_4 at 4°C for 15 min. The pellets were then washed in PBS, pH 7.2 and dehydrated in a graded acetone series. Finally, the dehydrated pellets were embedded in medium hard Epoxy resin (Electron Microscopy Sciences) and incubated at 60°C for 2 days. Ultrathin sections were then stained with uranyl acetate and observed using a JEOL JEM-1011 microscope (JEOL USA, Inc.).

Immunofluorescence confocal microscopy

The cells were prepared for immunofluorescence microscopy as described elsewhere (Rada *et al*., 2015). The recombinant proteins were detected in trichomonads using a mouse monoclonal anti-HA antibody or a rabbit monoclonal anti-V5 antibody (Abcam). The hydrogenosomal marker malic enzyme was detected using a rabbit polyclonal antibody (Drmota *et al*., 1997). Secondary Alexa Fluor 488 donkey anti-mouse and Alexa Fluor 594 donkey anti-rabbit antibodies were used for the visualization of the target proteins. The cells were examined using a Leica TCS SP8 inverted confocal microscope system (Leica Microsystems, Wetzlar, Germany).

Structured illumination microscopy (SIM)

SIM was performed using a 3D N-SIM microscope (Nikon Eclipse Ti-E, Nikon, Japan) equipped with a Nikon CFI SR Apo TIRF objective (100x oil, NA 1.49), as described previously (Štáfková *et al*., 2018).

Stimulated emission depletion (STED) microscopy

The cells were fixed on high-precision cover glasses (Carl Zeiss, Germany) for super-resolution microscopy and prepared for immunofluorescence microscopy as described elsewhere (Rada *et al*., 2015). The recombinant proteins were detected in trichomonads using a mouse monoclonal anti-HA antibody (Exbio, Czech Republic). The hydrogenosomal marker malic enzyme was detected using a rabbit polyclonal antibody (Drmota *et al*., 1997). The secondary

Abberior STAR580 and STAR 635P antibodies were used to visualize the target proteins. STED microscopy was performed using an Abberior STED 775 QUAD Scanning microscope (Abberior Instruments GmbH, Göttingen, Germany) equipped with a Nikon CFI Plan Apo Lambda objective (60x Oil, NA 1.40). Abberior STAR580- and STAR 635P-labeled proteins were illuminated by pulsed 561 and 640 nm lasers and depleted by a pulsed 775 nm STED depletion laser with a 2D donut beam. The fluorescence signal was filtered (emission bandpasses: 605–625 nm and 650–720 nm; pinhole 40 µm) and detected on single-photon counting modules with time gates of 0.8–8.8 ns. The images were scanned with a pixel size of 20 nm × 20 nm, a 10 µs dwell time, and the in-line interleaved acquisition mode using the Imspector software (Abberior Instruments). All images were deconvolved with Huygens Professional version software 17.04 (Scientific Volume Imaging, The Netherlands, http://svi.nl).

Bioinformatics

Transmembrane helixes and topology were predicted for protein sequences downloaded from the TrichoDB server (http://trichdb.org/trichdb/) using the TMHMM server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/). Proteins with a single C-terminal TMD and a short CTS of up to 30 AA in length (1456 proteins, Table S2) were selected and annotated according to TrichoDB and Pfam 31.0 (http://pfam.xfam.org/). The cell localization for each protein was prediction using TargetP (http://www.cbs.dtu.dk/services/TargetP/) and Cellular Component Ontology (http://geneontology.org/), the net charge was calculated as the sum of the 10 AA flanking the TMD on each side, and the hydrophobicity of the TMD was calculated as the sum of the hydropathy values of all AA residues according to the Kyte–Doolittle scale (Kyte and Doolittle, 1982) divided by the TMD length. To predict hydrogenosomal TA proteins, we first removed all proteins annotated as ER proteins (59) and we then removed all proteins with the following features: (i) hydrophobicity >2.64, (ii) net positive charge within 10 AA on the CTS side <1, (iii) CTS shorter than 3 AA residues, (iv) predicted secretory signal sequences and mitochondrial presequences and (v) proteins without the dibasic motif K/R/H-K/R/H within 10 AA of the CTS.

Authorship

J.T. designed the study, P.R., A.M. and V.Ž. performed the research, P.R. and J.T. analyzed the data, and P.R. and J.T. wrote the manuscript.

Acknowledgements

This work was supported by the programs KONTAKT II (LH15254) and NPU II (LQ1604) provided by the Ministry of Education, Youth and Sport (MEYS) of the Czech Republic, project BIOCEV and project CePaViP (CZ.02.1.01/0.0/0.0/16_019/0000759) supported by ERD Funds. We acknowledge the Imaging Methods Core Facility at BIOCEV, which is supported by MEYS (the Czech-BioImaging RI project LM2015062). The authors declare no conflict of interest.

References

Abell, B.M. and Mullen, R.T. (2011) Tail-anchored membrane proteins: exploring the complex diversity of tail-anchored-protein targeting in plant cells. *Plant Cell Reports*, **30**, 137–151.

Abell, B.M., Pool, M.R., Schlenker, O., Sinning, I. and High, S. (2004) Signal recognition particle mediates post-translational targeting in eukaryotes. *EMBO Journal*, **23**, 2755–2764.

Abell, B.M., Rabu, C., Leznicki, P., Young, J.C. and High, S. (2007) Post-translational integration of tail-anchored proteins is facilitated by defined molecular chaperones. *Journal of Cell Science*, **120**, 1743–1751.

Allen, R., Egan, B., Gabriel, K., Beilharz, T. and Lithgow, T. (2002) A conserved proline residue is present in the transmembrane-spanning domain of Tom7 and other tail-anchored protein subunits of the TOM translocase. *FEBS Letters*, **514**, 347–350.

Aviram, N., Ast, T., Costa, E.A., Arakel, E.C., Chuartzman, S.G., Jan, C.H., *et al.* (2016) The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature*, **540**, 134–138.

Beilharz, T., Egan, B., Silver, P.A., Hofmann, K. and Lithgow, T. (2003) Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, **278**, 8219–8223.

Bellot, G., Cartron, P.-F., Er, E., Oliver, L., Juin, P., Armstrong, L.C., *et al.* (2007) TOM22, a core component of the mitochondria outer membrane protein translocation pore, is a mitochondrial receptor for the proapoptotic protein Bax. *Cell Death and Differentiation*, **14**, 785–794.

Benchimol, M. (2008) The hydrogenosome peripheral vesicle: similarities with the endoplasmic reticulum. *Tissue and Cell*, **40**, 61–74.

Benchimol, M. (2009) Hydrogenosomes under microscopy. *Tissue and Cell*, **41**, 151–168.

Borgese, N. (2016) Getting membrane proteins on and off the shuttle bus between the endoplasmic reticulum and the Golgi complex. *Journal of Cell Science*, **129**, 1537–1545.

Borgese, N. and Fasana, E. (2011) Targeting pathways of C-tail-anchored proteins. *BBA – Biomembr*, **1808**, 937–946.

Borgese, N., Gazzoni, I., Barberi, M., Colombo, S. and Pedrazzini, E. (2001) Targeting of a tail-anchored protein to endoplasmic reticulum and mitochondrial outer membrane by independent but competing pathways. *Molecular Biology of the Cell*, **12**, 2482–2496.

Brkljacic, J., Zhao, Q. and Meier, I. (2009) WPP-domain proteins mimic the activity of the HSC70-1 chaperone in preventing mistargeting of RanGAP1-anchoring protein WIT1. *Plant Physiology*, **151**, 142–154.

Casson, J., McKenna, M., Haßdenteufel, S., Aviram, N., Zimmerman, R. and High, S. (2017) Multiple pathways

facilitate the biogenesis of mammalian tail-anchored proteins. *Journal of Cell Science*, **130**, 3851–3861.
Chacinska, A., Koehler, C.M., Milenkovic, D., Lithgow, T. and Pfanner, N. (2009) Importing mitochondrial proteins: machineries and mechanisms. *Cell*, **138**, 628–644.
Chio, U.S., Cho, H. and Shan, S. (2017) Mechanisms of tail-anchored membrane protein targeting and insertion. *Annual Review of Cell and Developmental Biology*, **33**, 417–438.
Cho, H. and Shan, S. (2018) Substrate relay in an Hsp70-cochaperone cascade safeguards tail-anchored membrane protein targeting. *EMBO Journal*, **37**, e99264.
Clemens, D.L. and Johnson, P.J. (2000) Failure to detect DNA in hydrogenosomes of *Trichomonas vaginalis* by nick translation and immunomicroscopy. *Molecular and Biochemical Parasitology*, **106**, 307–313.
Colombo, S.F. and Fasana, E. (2011) Mechanisms of insertion of tail-anchored proteins into the membrane of the endoplasmic reticulum. *Current Protein and Peptide Science*, **12**, 736–742.
Colombo, S.F., Longhi, R. and Borgese, N. (2009) The role of cytosolic proteins in the insertion of tail-anchored proteins into phospholipid bilayers. *Journal of Cell Science*, **122**, 2383–2392.
Costello, J.L., Castro, I.G., Camões, F., Schrader, T.A., Mcneall, D., Yang, J., *et al.* (2017a) Predicting the targeting of tail-anchored proteins to subcellular compartments in mammalian cells. *Journal of Cell Science*, 1675–1687.
Costello, J.L., Castro, I.G., Schrader, T.A., Islinger, M. and Schrader, M. (2017b) Peroxisomal ACBD4 interacts with VAPB and promotes ER-peroxisome associations. *Cell Cycle*, **16**, 1039–1045.
Diamond, L.S. (1957) The establishment of various trichomonads of animals and man in axenic cultures. *Journal of Parasitology*, **43**, 488–490.
Drmota, T., Tachezy, J. and Kulda, J. (1997) Isolation and characterization of cytosolic malate dehydrogenase from *Trichomonas vaginalis*. *Folia Parasitologica*, **44**, 103–108.
Embley, T.M. and Martin, W. (2006) Eukaryotic evolution, changes and challenges. *Nature*, **440**, 623–630.
Goder, V. and Spiess, M. (2001) Topogenesis of membrane proteins: determinants and dynamics. *FEBS Letters*, **504**, 87–93.
Hampl, V., Hug, L., Leigh, J.W., Dacks, J.B., Lang, B.F., Simpson, A.G.B. *et al.* (2009) Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". *Proceedings of the National Academy of Sciences*, **106**, 3859–3864.
Hegde, R.S. and Keenan, R.J. (2011) Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nature Reviews Molecular Cell Biology*, **12**, 787–798.
Henderson, M.P.A., Hwang, Y.T., Dyer, J.M., Mullen, R.T. and Andrews, D.W. (2007) The C-terminus of cytochrome b5 confers endoplasmic reticulum specificity by preventing spontaneous insertion into membranes. *The Biochemical Journal*, **709**, 701–709.
Higy, M., Junne, T. and Spiess, M. (2004) Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry*, **43**, 12716–12722.
Horie, C., Suzuki, H., Sakaguchi, M. and Mihara, K. (2002) Characterization of signal that directs C-tail-anchored proteins to mammalian mitochondrial outer membrane. *Molecular Biology of the Cell*, **13**, 1615–1625.
Horie, C., Suzuki, H., Sakaguchi, M. and Mihara, K. (2003) Targeting and assembly of mitochondrial tail-anchored protein Tom5 to the TOM complex depend on a signal distinct from that of tail-anchored proteins dispersed in the membrane. *Journal of Biological Chemistry*, **278**, 41462–41471.
Hrdý, I., Hirt, R.P., Doležal, P., Bardoňová, L., Foster, P.G., Tachezy, J. *et al.* (2004) Trichomonas hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. *Nature*, **432**, 618–622.
Hrdý, I., Tachezy, J. and Müller, M. (2008) Metabolism of trichomonad hydrogenosomes. In: Tachezy, J. (Ed.), *Hydrogenosomes and Mitosomes: Mitochondria of Anaerobic Eukaryotes*. Berlin Heidelberg: Springer-Verlag, pp. 114–145.
Hwang, Y.T., Pelitire, S.M., Henderson, M.P.A., Andrews, D.W., Dyer, J.M. and Mullen, R.T. (2004) Novel targeting signals mediate the sorting of different isoforms of the tail-anchored membrane protein cytochrome b5 to either endoplasmic reticulum or mitochondria. *The Plant Cell*, **16**, 3002–3019.
Isenmann, S., Khew-Goodall, Y., Gamble, J., Vadas, M. and Wattenberg, B.W. (1998) A splice-isoform of vesicle-associated membrane protein-1 (VAMP-1) contains a mitochondrial targeting signal. *Molecular Biology of the Cell*, **9**, 1649–1660.
Kalbfleisch, T., Cambon, A. and Wattenberg, B.W. (2007) A bioinformatics approach to identifying tail-anchored proteins in the human genome. *Traffic*, **8**, 1687–1694.
Kaufmann, T., Schlipf, S., Sanz, J., Neubert, K., Stein, R. and Borner, C. (2003) Characterization of the signal that directs Bcl-xL, but not Bcl-2, to the mitochondrial outer membrane. *Journal of Cell Biology*, **160**, 53–64.
Kemper, C., Habib, S.J., Engl, G., Heckmeyer, P., Dimmer, K.S. and Rapaport, D. (2008) Integration of tail-anchored proteins into the mitochondrial outer membrane does not require any known import components. *Journal of Cell Science*, **121**, 1990–1998.
Kriechbaumer, V., Shaw, R., Mukherjee, J., Bowsher, C.G., Harrison, A.M. and Abell, B.M. (2009) Subcellular distribution of tail-anchored proteins in arabidopsis. *Traffic*, **10**, 1753–1764.
Krumpe, K., Frumkin, I., Herzig, Y., Rimon, N., Özbalci, C., Brügger, B., *et al.* (2012) Ergosterol content specifies targeting of tail-anchored proteins to mitochondrial outer membranes. *Molecular Biology of the Cell*, **23**, 3927–3935.
Kuroda, R., Ikenoue, T., Honsho, M., Tsujimoto, S., Mitoma, J.Y. and Ito, A. (1998) Charged amine acids at the carboxyl-terminal portions determine the intracellular locations of two isoforms of cytochrome b5. *Journal of Biological Chemistry*, **273**, 31097–31102.
Kyte, J. and Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, **157**, 105–132.
Mani, J., Desy, S., Niemann, M., Chanfon, A., Oeljeklaus, S., Pusnik, M., *et al.* (2015) Mitochondrial protein

import receptors in Kinetoplastids reveal convergent evolution over large phylogenetic distances. *Nature Communications*, **6**, 6646.
Marty, N.J., Teresinski, H.J., Hwang, Y.T., Clendening, E.A., Gidda, S.K., Sliwinska, E., *et al.* (2014) New insights into the targeting of a subset of tail-anchored proteins to the outer mitochondrial membrane. *Frontiers in Plant Science*, **5**, 1–20.
Mateja, A., Paduch, M., Chang, H.-Y., Szydlowska, A., Kossiakoff, A.A., Hegde, R.S. *et al.* (2015) Protein targeting. Structure of the Get3 targeting factor in complex with its membrane protein cargo. *Science*, **347**, 1152–1155.
Motz, C., Martin, H., Krimmer, T. and Rassow, J. (2002) Bcl-2 and porin follow different pathways of TOM-dependent insertion into the mitochondrial outer membrane. *Journal of Molecular Biology*, **323**, 729–738.
Padgett, L.R., Arrizabalaga, G. and Sullivan, W.J. (2017) Targeting of tail-anchored membrane proteins to subcellular organelles in *Toxoplasma gondii*. *Traffic*, **18**, 149–158.
Pagny, S., Cabanes-Macheteau, M., Gillikin, J.W., Leborgne-Castel, N., Lerouge, P., Boston, R.S., *et al.* (2000) Protein recycling from the Golgi apparatus to the endoplasmic reticulum in plants and its minor contribution to calreticulin retention. *The Plant Cell*, **12**, 739–756.
Rabu, C., Wipf, P., Brodsky, J.L. and High, S. (2008) A precursor-specific role for Hsp40/Hsc70 during tail-anchored protein integration at the endoplasmic reticulum. *Journal of Biological Chemistry*, **283**, 27504–27513.
Rabu, C., Schmid, V., Schwappach, B. and High, S. (2009) Biogenesis of tail-anchored proteins: the beginning for the end? *Journal of Cell Science*, **122**, 3605–3612.
Rada, P., Doležal, P., Jedelský, P.L., Bursac, D., Perry, A.J., Šedinová, M., *et al.* (2011) The core components of organelle biogenesis and membrane transport in the hydrogenosomes of *Trichomonas vaginalis*. *PLoS ONE*, **6**, e24428.
Rada, P., Makki, A.R., Zimorski, V., Garg, S., Hampl, V., Hrdý, I., *et al.* (2015) N-Terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*. *Eukaryotic Cell*, **14**, 1264–1275.
Roger, A.J., Muñoz-Gómez, S.A. and Kamikawa, R. (2017) The origin and diversification of mitochondria. *Current Biology*, **27**, R1177–R1192.
Schneider, R.E., Brown, M.T., Shiflett, A.M., Dyall, S.D., Hayes, R.D., Xie, Y., *et al.* (2011) The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. *International Journal for Parasitology*, **41**, 1421–1434.
Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., *et al.* (2008) The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell*, **134**, 634–645.
Štáfková, J., Rada, P., Meloni, D., Žárský, V., Smutná, T., Zimmann, N., *et al.* (2018) Dynamic secretome of *Trichomonas vaginalis*: Case study of β-amylases. *Molecular & Cellular Proteomics*, **17**, 304–320.
Šuťák, R., Doležal, P., Fiumera, H.L., Hrdy, I., Dancis, A., Delgadillo-Correa, M., *et al.* (2004) Mitochondrial-type assembly of FeS centers in the hydrogenosomes of the amitochondriate eukaryote *Trichomonas vaginalis*. *Proceedings of the National Academy of Sciences*, **101**, 10368–10373.
Tovar, J., León-Avila, G., Sánchez, L.B., Sutak, R., Tachezy, J., van der Giezen, M., *et al.* (2003) Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature*, **426**, 172–176.
Urra, H. and Hetz, C. (2012) The ER in 4D: a novel stress pathway controlling endoplasmic reticulum membrane remodeling. *Cell Death and Differentiation*, **19**, 1893–1895.
Vannuvel, K., Renard, P., Raes, M. and Arnould, T. (2013) Functional and morphological impact of ER stress on mitochondria. *Journal of Cellular Physiology*, **228**, 1802–1818.
Wiedemann, N. and Pfanner, N. (2017) Mitochondrial machineries for protein import and assembly. *Annual Review of Biochemistry*, **86**, 685–714
Yagita, Y., Hiromasa, T. and Fujiki, Y. (2013) Tail-anchored PEX26 targets peroxisomes via a PEX19-dependent and TRC40-independent class I pathway. *Journal of Cell Biology*, **200**, 651–666.
Yuen, C.Y.L., Shek, R., Kang, B.-H., Matsumoto, K., Cho, E.J. and Christopher, D.A. (2016) Arabidopsis protein disulfide isomerase-8 is a type I endoplasmic reticulum transmembrane protein with thiol-disulfide oxidase activity. *BMC Plant Biology*, **16**, 181.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

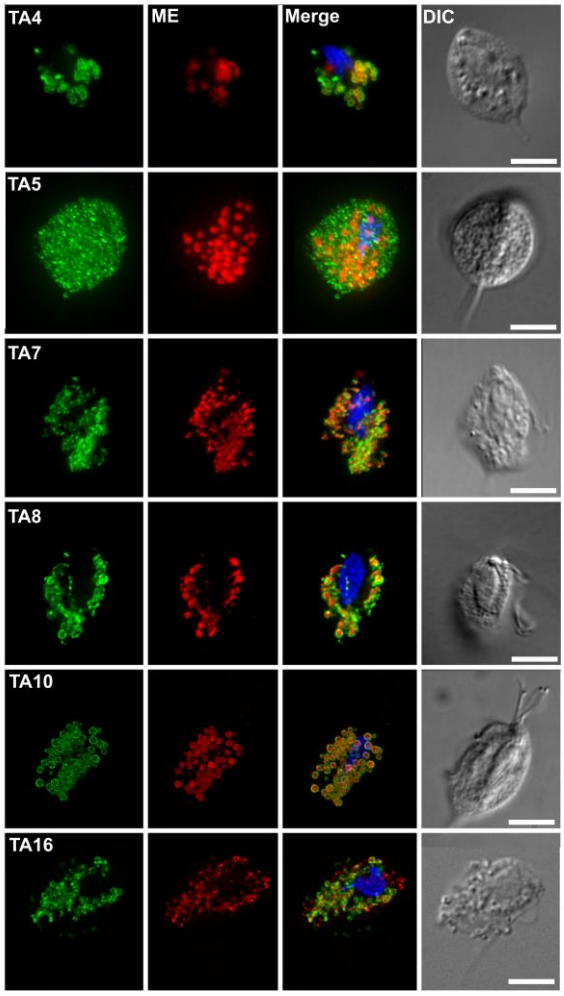


Figure S1. Effect of C-terminal tagging on cellular localization of hydrogenosomal TA proteins.

TA proteins were expressed with a C-terminal HA tag in trichomonads and visualized with a mouse monoclonal anti-HA antibody (green). The hydrogenosomal marker malic enzyme (ME) was detected with a polyclonal rabbit anti-malic enzyme antibody (red). The nucleus was stained with DAPI (blue). DIC, differential interference contrast. The scale bar represents 5 μm.

Metadata of the chapter that will be visualized online

Chapter Title	Protein Import into Hydrogenosomes and Mitosomes	
Copyright Year	2019	
Copyright Holder	Springer Nature Switzerland AG	
Corresponding Author	Family Name	**Dolezal**
	Particle	
	Given Name	**Pavel**
	Suffix	
	Division	Faculty of Science, Department of Parasitology
	Organization	Charles University
	Address	Vestec, Czech Republic
	Email	pavel.dolezal@natur.cuni.cz
Author	Family Name	**Makki**
	Particle	
	Given Name	**Abhijith**
	Suffix	
	Division	Faculty of Science, Department of Parasitology
	Organization	Charles University
	Address	Vestec, Czech Republic
	Email	abhijith.makki@natur.cuni.cz
Author	Family Name	**Dyall**
	Particle	
	Given Name	**Sabrina D.**
	Suffix	
	Division	Department of Biosciences and Ocean Studies
	Organization	University of Mauritius
	Address	Réduit, Mauritius
	Email	s.dyall@uom.ac.mu

Abstract

In the past decade, studies on protein targeting to hydrogenosomes and mitosomes have revealed several characteristics in common with mitochondrial protein targeting. Proteins from one system can readily be imported into another, strongly suggesting that targeting signals on hydrogenosomal, mitosomal and mitochondrial preproteins are conserved. By extension, these observations, together with the proposed common origin of hydrogenosomes, mitosomes and mitochondria, led to the proposition that components of the respective protein import machineries for these organelles are conserved. With the advent of complete genome sequence databases for diverse eukaryotes, we are now in a better position to examine this proposition. In this review, we report and integrate the latest experimental

and bioinformatics data on the state of protein import in hydrogenosomes, mitosomes and mitochondria.

Protein Import into Hydrogenosomes and Mitosomes

Pavel Dolezal, Abhijith Makki, and Sabrina D. Dyall

Contents

P. Dolezal (✉) · A. Makki
Faculty of Science, Department of Parasitology, Charles University, Vestec, Czech Republic
e-mail: pavel.dolezal@natur.cuni.cz; abhijith.makki@natur.cuni.cz

S. D. Dyall
Department of Biosciences and Ocean Studies, University of Mauritius, Réduit, Mauritius
e-mail: s.dyall@uom.ac.mu

J. Tachezy (ed.), *Hydrogenosomes and Mitosomes: Mitochondria of Anaerobic Eukaryotes*, Microbiology Monographs 9,
https://doi.org/10.1007/978-3-030-17941-0_3

Abstract In the past decade, studies on protein targeting to hydrogenosomes and mitosomes have revealed several characteristics in common with mitochondrial protein targeting. Proteins from one system can readily be imported into another, strongly suggesting that targeting signals on hydrogenosomal, mitosomal and mitochondrial preproteins are conserved. By extension, these observations, together with the proposed common origin of hydrogenosomes, mitosomes and mitochondria, led to the proposition that components of the respective protein import machineries for these organelles are conserved. With the advent of complete genome sequence databases for diverse eukaryotes, we are now in a better position to examine this proposition. In this review, we report and integrate the latest experimental and bioinformatics data on the state of protein import in hydrogenosomes, mitosomes and mitochondria.

1 Introduction

Eukaryotic cells have internal membranes defining subcellular compartments, each of which has discrete metabolic and/or biosynthetic functions. To fulfil such functions, specific sets of proteins must be precisely targeted to, and quantitatively imported and localized within, the compartment in a timely fashion. Protein trafficking has been extensively studied in fungi and mammals, and a number of elaborate machines have been described that specifically import certain proteins into mitochondria. Recent studies have demonstrated that there are several mechanistic and structural features in common between import into hydrogenosomes, mitosomes and mitochondria (Makiuchi and Nozaki 2014; Pyrihová et al. 2018; Makki et al. 2019). Despite their different morphologies and non-mitochondrial functional pathways, hydrogenosomes and mitosomes are thought to be related descendants of the endosymbiont that gave rise to mitochondria (Roger et al. 2017). Due to the limited number of the experimental models, we focus the current review to *Trichomonas vaginalis, Neocallimastix* sp. and *Nyctotherus ovalis* hydrogenosomes and to *Entamoeba histolytica*, *Giardia intestinalis*, *Cryptosporidium* sp. and microsporidian mitosomes. We also include additional information inferred from the recent genomic and transcriptomic data. We shall examine what is known about hydrogenosomal and mitosomal biogenesis in those species and shall discuss the evolution of protein import in relation to bacterial, mitochondrial and other organellar systems.

2 Protein Trafficking in Eukaryotes

Of over 6500 proteins encoded in the nuclear genome of *Saccharomyces cerevisiae*, about 1000 proteins are targeted to the endoplasmic reticulum (ER) and from there are subsequently localized throughout the endomembrane system, and a similar number of proteins are targeted to mitochondria. Nearly 50% of the proteins are folded and localized in the cytoplasm, while around 25% of them are retargeted to the nucleus, and a smaller portion is distributed in other compartments such as peroxisomes (Kumar et al. 2002; Picotti et al. 2013). In general, a sophisticated system of membrane translocases with associated propelling machines recognizes the address on individual protein molecules. In most cases, protein trafficking is fuelled by the hydrolysis of either ATP or GTP. Additionally, proteins travelling through or into mitochondrial inner membranes require an electrochemical membrane potential generated by the mitochondrial electron transport chain (Wickner and Schekman 2005).

Despite these general similarities, the respective mechanisms responsible for protein import into the nucleus, the ER and the mitochondrion differ fundamentally, with each employing distinct molecular machine. While the molecular machines operating in the ER were recruited during the evolution of the eukaryotic cell from an ancestral prokaryote, and adapted towards current needs, nuclear and mitochondrial protein import systems seem to be almost entirely created de novo by the eukaryotic cell (Dolezal et al. 2006; Lithgow and Schneider 2010; Cautain et al. 2015; Mani et al. 2016; Fukasawa et al. 2017).

2.1 The Nucleus

The nuclear envelope is perforated with huge macromolecular assemblies of ~30 different proteins that form nuclear pore complexes with a central channel of 25–30 nm in diameter. This channel fuses the inner and the outer nuclear membranes and allows proteins smaller than 40 kDa to passively traverse. Larger proteins are actively transported across the nuclear envelope and contain nuclear localization signal (NLS) sequence motifs. These signals consist of one or two clusters of four or five basic residues localized usually within the polypeptide chain. The import of proteins with NLS through the channel is facilitated by the carrier heterodimer of importin-α/β (Pemberton and Paschal 2005; Lange et al. 2007). Upon passing through the nuclear pore, the interaction of the complex with RanGTP initiates the release of cargo protein from the importins. The whole process of translocation is regulated by the nucleotide state of Ran, which accordingly cycles between the nucleus and the cytoplasm (Stewart 2007).

2.2 The Endoplasmic Reticulum

In contrast to cytosolic, nuclear and most mitochondrial proteins that are synthesized on free ribosomes, mRNA transcripts encoding ER-destined proteins are translated on ribosomes tightly bound to the ER membrane. Nascent luminal proteins are equipped with an N-terminal signal sequence that consists of a basic amino-terminus followed by a stretch of 8–14 non-polar residues and a cleavage motif for the signal peptidase (Blobel and Dobberstein 1975a, b; von Heijne 1990) (Blobel and Dobberstein 1975a, b). Membrane proteins usually contain internal topological signals instead of the N-terminal signal peptide. Translation and translocation are coordinated by the signal recognition particle (SRP), a complex of 7S RNA and six protein subunits (Nyathi et al. 2013). Initially, SRP binds the signal peptide emerging from the ribosome. Translation slows down until SRP is recognized by its ER-bound receptor, whereupon translocation can resume following the binding of GTP to both SRP and its receptor. The passage through the membrane is formed by the Sec61 translocon consisting of a Sec61α channel and two accessory subunits β and γ (Park and Rapoport 2012). The co-translational transport of substrate protein through the channel is driven by the elongation of the polypeptide by the ribosome (Connolly and Gilmore 1986; Ménétret et al. 2007).

In case of post-translational transport, the chaperones protecting the translated polypeptides are released upon the contact with Sec61 translocon, which is accompanied by the additional Sec62/Sec63 complex of so far unclear role. The translocation is then driven by the action of lumenal Hsp70 (Osborne et al. 2005). Analogously, in bacteria, the SecYEG translocon is used for secretion of proteins across the plasma membrane, and the signal peptides of secreted proteins share similar characteristics with ER proteins (Park and Rapoport 2012). While protein import into ER requires nucleotide triphosphates, bacteria need additional membrane electrochemical gradient to export proteins across the plasma membrane. After translocation into the ER, the signal peptide is cleaved from the precursor polypeptide by the signal peptidase. This step is necessary for releasing the protein from the membrane lipid bilayer to which it is bound via the hydrophobic signal peptide. The ER signal peptidase is a membrane protein that shares ancestry with both the bacterial signal peptidase and the inner membrane protease complex in mitochondria (Dalbey et al. 1997).

2.3 The Mitochondrion

About 99% of the mitochondrial proteins are nuclear-encoded and are synthesized in the cytosol from where they are imported into mitochondria (Wiedemann and Pfanner 2017). Protein import into mitochondria is mostly post-translational, although recent data suggest that mRNA subpopulations are organized in the proximity of the mitochondrial outer membrane (García-Rodríguez et al. 2007).

Besides having to be targeted to the organelle, these proteins have to be internally sorted to either of four distinct sub-compartments: the outer membrane, the intermembrane space (IMS), the inner membrane or the mitochondrial matrix. Proteins destined to mitochondria possess N-terminal and/or internal targeting signals which ensure their correct delivery to the organelle. A large majority of mitochondrial proteins are synthesized with N-terminal cleavable presequences of variable length, ranging from 10 to 80 amino acids (Wiedemann and Pfanner 2017). The presequence is not generally conserved at the primary sequence level among the different preproteins, but it is rather the α-helical structure engaged by the presequence upon interaction with the outer membrane receptor Tom20 that appears to be the common factor (Schatz and Dobberstein 1996; Abe et al. 2000). This α-helix is amphipathic, containing patches of positively charged, and hydrophobic amino acids, respectively, on opposite surfaces of the theoretical cylinder. The presequence is usually processed by the mitochondrial processing peptidase (MPP), and the mature protein is sorted to either the matrix or to the inner membrane if it bears a hydrophobic stop-transfer sequence. Some mitochondrial proteins, mostly destined to the membranes, do not have cleavable N-terminal presequences but have internal targeting signals that are not well characterized (Wiedemann and Pfanner 2017).

Translocation across mitochondrial membranes is carried out by several molecular machines. These machines consist of core transmembrane translocases complemented by additional components that provide specificity. Such a modular structure enables the independent evolution and function of each molecular machine (Dolezal et al. 2006). The translocase of the outer membrane (TOM) complex constitutes the central recognition point and gate for all nuclear-encoded mitochondrial proteins (Fig. 1). Tom70 and Tom20 are receptor subunits that recognize the precursor proteins and release them subsequently into the translocation channel. This transfer is assisted by Tom22, which, together with Tom40 and Tom5, represents the core and essential part of the TOM complex (Meisinger et al. 2001). Two other small proteins, Tom6 and Tom7, participate in the maintenance of the complex. The translocation pore is formed by several subunits of Tom40, which is most likely a β-barrel protein (Hill et al. 1998; Bausewein et al. 2017). After passing through the TOM complex, preproteins may interact with either of three distinct molecular machines (Fig. 1), depending on their final destination.

The insertion and assembly of β-barrel outer membrane proteins, including Tom40, are assisted by the sorting and assembly machinery (SAM) complex (Fig. 1). The SAM complex consists of four subunits, the core translocase Sam50, which is itself a β-barrel protein (Kozjak et al. 2003; Paschen et al. 2003; Gentle et al. 2004), and the additional proteins Sam35, Sam37 and Mdm10 (Bohnert et al. 2007). In addition to SAM complex Mdm10 takes part in the ER-mitochondria tethering complex known as ERMES (Kornmann et al. 2009). The structure of β-barrel precursors does not allow their lateral insertion into the lipid bilayer directly from the TOM channel, and the precursors must first be released into the IMS. The passage of β-barrel precursors from the TOM complex to the SAM complex is assisted by the so-called small translocase of the inner membrane (TIM) chaperone

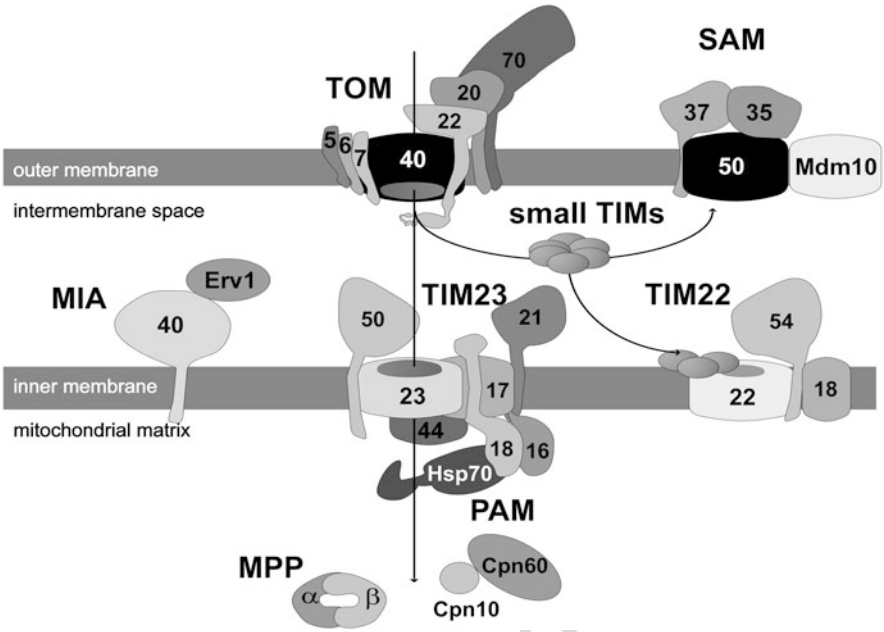


Fig. 1 The mitochondrial protein import machinery as defined in *S. cerevisiae*. *TOM* translocase of the outer mitochondrial membrane, *SAM* sorting and assembly machinery, *TIM* translocase of the inner mitochondrial membrane, *MIA* mitochondrial IMS assembly machinery, *PAM* presequence-associated motor, *IMP* inner membrane protease, *MPP* mitochondrial processing peptidase. The numbers on the individual TOM, SAM, TIM or PAM components represent their approximate molecular masses in kDa. See text for mechanistic details

complexes (small Tims). These soluble complexes are trimeric assemblies of either Tim9/Tim10 heterodimers or of Tim8/Tim13 heterodimers and protect the exposed hydrophobic epitopes of some membrane proteins from the aqueous environment of the IMS (Hoppins and Nargang 2004; Wiedemann et al. 2004; Koehler 2004). Specific class of outer membrane proteins with a-helical transmembrane segment are inserted into the membrane via distinct mitochondrial import (MIM) complex (Becker et al. 2008; Hulett et al. 2008).

Some of the precursors destined for the IMS have a sorting signal at the N-terminus. The inner membrane protease (IMP) complex is responsible for the maturation of these proteins. IMP comprises the two proteases Imp1 and Imp2 and a regulatory subunit Som1. Some of the precursors contain bipartite presequences consisting of a matrix-targeting signal followed by an IMS-sorting signal (Gakh et al. 2002). However, majority of the IMS proteins lack any N-terminal presequences and instead contain internal CX3C or CX9C motives (Stojanovski et al. 2012), which are recognized by the mitochondrial IMS import and assembly (MIA) machinery (Mesecke et al. 2005). The machinery consists of IMS components Mia40, Erv1 and Hot13p (Fig. 1). According to the proposed model, Mia40 binds the precursors in the IMS via disulphide bridges, thereby trapping them after

their entrance through the Tom40 pore. Further isomerization of disulphide bridges releases the precursors from Mia40, which is subsequently oxidized by Erv1 (Mesecke et al. 2005). In this process, Hot13p might perform a reducing action on the precursors (Curran et al. 2004).

Some precursors that are to be integrated in the inner membrane, such as the mitochondrial carrier family (MCF) proteins, are inserted by the TIM22 machine (Fig. 1), which is built around a Tim22 subunit and contains another two membrane-integral subunits Tim54 and Tim18 (Rehling et al. 2003). The precursors are shuttled from the TOM to the TIM22 complex by the Tim9/Tim10 chaperone through the IMS. Tim12, peripherally associated with TIM22, serves as a docking site for the Tim9/Tim10 complex that detaches from the precursor upon contact with the TIM22 complex (Koehler 2004).

Soluble matrix-destined preproteins, usually synthesized with a cleavable N-terminal presequence, are passed from the TOM complex to the distinct TIM23 machine through interaction with the IMS domain of Tim50 (Fig. 1). In yeast, TIM23 consists of Tim23, Tim17, Tim50 and Tim21 (Bohnert et al. 2007). Tim23 forms a protein-conducting channel that is regulated by the action of Tim50 and Tim17 (Meinecke et al. 2006; Martinez-Caballero et al. 2007). The TIM23 complex is also capable of inserting preproteins into the inner mitochondrial membrane (Koehler 2004). These inner membrane preproteins are synthesized with a short hydrophobic sorting sequence downstream of the N-terminal presequence (Glick et al. 1992; Stuart 2002). To discriminate between and to coordinate the dual translocation and insertion activities, TIM23 either interacts with Tim21 or associates with the presequence translocase-associated motor (PAM) complex (Fig. 1). This cycle is regulated by Tim17, which recruits the PAM complex to the TIM23 complex (Chacinska et al. 2005). The PAM complex, together with the membrane potential, drives translocation of the precursor through the TIM23 complex (Rehling et al. 2003). First, the negatively charged matrix face of the inner membrane generates an electrophoretic force on the predominantly positively charged presequence and mediates preprotein insertion into the Tim23 channel. Consequently, the PAM complex completes the translocation of the bulk polypeptide in an ATP-dependent manner (Voos and Röttgers 2002).

The central component of the PAM complex is the molecular chaperone Hsp70 (Fig. 1). The activity of Hsp70 is regulated by two inner membranes, J-domain-containing proteins Pam18 and Pam16, and the soluble nucleotide exchange factor Mge1. The PAM complex is tethered to the translocation channel probably by the peripheral membrane protein Tim44. Upon translocation into the matrix, the N-terminal presequence of precursors is processed by MPP, and the mature protein is thereafter folded into its native conformation. Some preproteins with an octapeptide-containing presequence require sequential processing of the targeting presequence by MPP followed by the mitochondrial intermediate peptidase (MIP), which removes the octapeptide sequence (Gakh et al. 2002).

After being driven in by the PAM complex and having their presequence processed, some newly imported matrix proteins require further assistance from molecular chaperones. First, soluble Hsp70 and its co-chaperone Mdj1 accost the

substrate proteins to partially fold them. Two populations of Hsp70 thus exist in mitochondria: (1) a Tim44-bound membrane-associated form serving as the protein import driver and (2) a matrix-soluble chaperone that assists protein folding (Horst et al. 1997). After release from Hsp70, the protein is passed along an Hsp60/Hsp10 system, which is the major chaperone system for protein folding in the matrix (Manning-Krieg et al. 1991). A homo-oligomer of Hsp60 provides a protected cavity for protein folding, while Hsp10 regulates the ATPase cycle of Hsp60 and the behaviour of individual subunits (Martin et al. 1991).

3 The Evolution of the Mitochondrial Protein Import Machinery

Mitochondria are of endosymbiotic origin and have descended from bacteria. Extensive sequence analyses have shown that mitochondria form a monophyletic group and have demonstrated strong affinities between mitochondrial genomes and present-day alphaproteobacteria-related organisms (Andersson et al. 1998; Gray et al. 1999; Martijn et al. 2018). Consequently, the endosymbiotic theory for the origin of mitochondria purports that the mitochondrion originates from a single endosymbiont, which formed a symbiotic relationship with a pre-eukaryotic or a primitive eukaryotic cell around two billion years ago. Over time, the endosymbiont lost its capacity to function and reproduce as an independent organism, and its fate was sealed within the host as it transferred the bulk of its genome to the host nucleus or simply discarded some of it. The possible reasons or driving forces behind the symbiosis, and the subsequent loss of the endosymbiont genome and why and how that happened, are beyond the scope of this chapter and are comprehensively covered in Chap. 2 and in reviews, e.g. (Martin et al. 2015; Roger et al. 2017). The outcome of, or perhaps the support for, the endosymbiont transferring its genes to the nucleus was the evolution of new machinery in the eukaryotic cell to send the nuclear-encoded proteins back to the degenerate endosymbiont to allow the latter to function. Moreover, it is of note that the large majority of extant mitochondrial proteins are not of endosymbiotic or α-proteobacterial origin. These proteins have either been recruited from other bacterial sources or have been invented de novo by the evolving eukaryote (Andersson et al. 2003; Gabaldón and Huynen 2003; Gray 2015). All these proteins would have had to develop targeting signals, while the eukaryote was inventing a new machine, either from scratch, or by tinkering existing protein targeting components, to intake the nuclear-encoded precursors into the proto-mitochondrion.

Of the six protein import machines characterized to date in yeast mitochondria, namely, TOM, SAM, MIM, TIM22, TIM23 and MIA, only the SAM complex bears a component, Sam50, which is clearly related to a bacterial translocase (Gentle et al. 2004). Several import components such as Pam18 and Tim44 contain domains found in bacteria (Clements et al. 2009). Thus, the majority of translocase

components are a product of eukaryotic invention. Much can be inferred about the evolution of mitochondrial biogenesis by examining these translocases and their features, as discussed later in this chapter. It is of particular interest whether hydrogenosomes and mitosomes use phylogenetically similar translocases as mitochondria.

The majority of mitochondrial proteins have an N-terminal presequence that is both necessary and sufficient to target a passenger protein to mitochondria (von Heijne et al. 1989). How and when these presequences were initially acquired and how they have been distributed to genes on different loci are intriguing questions. It has been demonstrated that synthetic mitochondrial presequences can penetrate either artificial or bacterial lipid bilayers (Roise et al. 1986; Maduke and Roise 1993; Neupert 1997). Of the presequence properties, it is mainly their amphiphilic character and the net charge which have a decisive role in targeting proteins to mitochondria and other organelles. It is therefore plausible that presequences were developed prior to the existence of specific receptors or pores on the outer surface of the proto-mitochondrion and that the latter were later evolved to enhance the efficiency of translocation. Conversely, it was hypothesized that in order to avoid crosstalk between targeting to host cell translocases at the plasma membrane and the newly developing organelle, the early mitochondrial targeting might not have relied on the presequence at all (Garg and Gould 2016). It was proposed that a primitive PAM might have driven the presequence and the mature polypeptide into the proto-mitochondrion in the absence of any TOM or TIM component (Herrmann 2003). Sequences with presequence-like features are commonly found in genomes, and the odds of a transferred gene landing in such a locus may have been quite good (Baker and Schatz 1987; Lucattini et al. 2004). Since some presequences are distributed over several exons in some nuclear-encoded mitochondrial genes, exon shuffling and alternative splicing have been proposed as mechanisms for presequence generation (McFadden 1999). Much insight on these aspects has been gained by studying mitochondrial gene transfer processes in flowering plants (Adams et al. 2000; Adams and Palmer 2003). In those species, mitochondrial gene transfer is an ongoing process, such that functional copies of some genes for mitochondrial proteins can be found in (1) both the nucleus and the mitochondrion of one species or (2) in the nucleus of one species, but only in the mitochondrion of a sister species. Thus, the changes required to target the newly transferred genes to the mitochondrion can be examined. Productive gene transfer not only involves the evolution of a targeting signal but is primarily dependent on the acquisition of gene expression signals. It was shown that mitochondrial copies of freshly transferred genes remain active for some indeterminate period and do not immediately get shut off (Choi et al. 2006). Transferred genes can acquire presequences by simply integrating into the locus for a duplicated nuclear-encoded mitochondrial gene (Sandoval et al. 2004; Murcha et al. 2005b; Choi et al. 2006). Curiously, some mitochondrial presequences have independently been acquired from the same donor gene, e.g. from *mt-hsp70* (Adams et al. 2000; Choi et al. 2006). Genes for non-mitochondrial proteins have also acted as donors for gene control regions and for fortuitous presequence-like stretches (Murcha et al. 2005b). Or, sometimes, the transferred gene does not gain a

presequence but uses internal signals from the mature polypeptide (Murcha et al. 2005b; Choi et al. 2006). Additionally, changes in local hydrophobicity in the protein sequence of certain transferred genes have been implied in enhancing protein import (Daley et al. 2002). Therefore, a plethora of tricks exists to append presequences to, or to create internal signals within, mitochondrial proteins, and though the precise mechanisms are shady, it appears that these tricks occur repeatedly and independently.

It has been hypothesized that the process of inventing a protein import machine for mitochondria would have been so intricate and critical that it is unlikely to have occurred more than once (Cavalier-Smith 1987). By extension, similarities found between mitosomal and hydrogenosomal and mitochondrial protein import have been presented as a strong support that all these organelles use common components for import and are therefore one and the same. These observations have prompted a number of studies to shed light on the constitution and the evolution of the protein import machineries of hydrogenosomes and mitosomes.

4 Studying Hydrogenosomal and Mitosomal Protein Import

4.1 Laboratory Techniques and Tools

Unfortunately, a limited set of tools is available to study the hydrogenosomal or mitosomal species. In yeast, much has been deduced about mitochondrial biogenesis through extensive genetic manipulation, and a variety of mutants can readily be obtained that can be used to assess the function of individual components in the biogenetic pathway (Bonnefoy et al. 2007). These studies are complemented with a wealth of highly honed biochemical techniques such as in organello import, membrane separation, creation of mitoplasts and generation of protein import intermediates, to list a few (Stojanovski et al. 2007). Nonetheless, these techniques provide a basis for developing new methods to study protein import in hydrogenosomal or mitosomal species.

Among the species under study, *T. vaginalis* is currently one of the most experimentally tractable. It can be genetically transformed to express endogenous or exogenous proteins (Delgadillo et al. 1997) or to delete genes by homologous recombination (Land et al. 2003; Brás et al. 2013). More recently, a CRISPR/Cas9-mediated gene modification and gene knockout methods were developed for *T. vaginalis* that are based on techniques originally developed for *Cryptosporidium parvum* (Vinayak et al. 2015; Janssen et al. 2018). Establishment of a gene knockout technique can significantly advance the field and can help the biologists to prove the gene function. Of all the mitochondria-related organelles, only *T. vaginalis* hydrogenosomes (~0.8 μm in diameter) have been isolated to high purity, on a Percoll gradient, and have been demonstrated to be protein-import competent

(Bradley et al. 1997). The development of an assay for importing precursor proteins into isolated *T. vaginalis* hydrogenosomes has revealed several requirements that appear to be in common with mitochondrial protein import (Bradley et al. 1997). For this assay, recombinant precursor proteins are either metabolically radiolabelled in *Escherichia coli* for detection by autoradiography or purified with a C-terminal hexahistidine (His6) tag for detection by western analysis. The isolated organelles and the precursor are incubated in an isotonic import buffer supplemented with ATP and crude *T. vaginalis* cytosol. In mitochondrial import systems, the precursor protein is synthesized and radiolabelled in rabbit reticulocyte lysate or in wheat germ extract; these extracts contain cytosolic chaperones that support mitochondrial protein import. However, hydrogenosomal protein import is absolutely dependent on *T. vaginalis* crude cytosol that cannot be substituted for by either rabbit reticulocyte lysate or wheat germ extract (Bradley et al. 1997), emphasizing the specificity of the cytosolic factor(s). Successful import is measured as follows: resistance of the imported protein to externally added protease and presequence cleavage, as detected by faster electrophoretic migration on SDS-PAGE. The import of a matrix precursor protein, ferredoxin, has been shown to be linear and saturable, and dependent on a protease-sensitive component(s) on the outer hydrogenosomal surface, indicating the presence of a specific receptor (Plümper et al. 2000). Precursor ferredoxin import has been shown to be dependent on the presence of a specific presequence and, on ATP, weak electrochemical potential and temperature (Bradley et al. 1997), which are all requirements for import into mitochondria (Schleyer et al. 1982). Import studies have also been carried out to explore the functional conservation of import pathways between hydrogenosomes and mitochondria (Dyall et al. 2000, 2003) or between hydrogenosomes and mitosomes (Dolezal et al. 2005).

The second most studied hydrogenosomal species is *Neocallimastix* sp., although not many techniques are available for thorough studies. An enriched hydrogenosomal fraction can be prepared by differential centrifugation of disrupted cells of this species and can be used for rudimentary sub-organellar fractionation studies (Marvin-Sikkema et al. 1993). However, most of the studies on *Neocallimastix* sp. hydrogenosomal proteins have been done in heterologous fungal systems (van der Giezen et al. 1998, 2002, 2003).

Giardia and *Entamoeba* represent the best cellular models for studying the mitosomal biology. Genetic transformation techniques are used for both organisms and have been applied to localize mitosomal proteins and to investigate putative targeting signals, e.g. (Mai et al. 1999; Tovar et al. 2003; Regoes et al. 2005; Dolezal et al. 2005; Mi-ichi et al. 2009). Transcriptional silencing techniques exist for *Entamoeba*, e.g. (Linford et al. 2009), including elusive G3 strain-specific silencing (Bracha et al. 2003), although none of them represent a robust and reliable genetic tool. Similarly, ribozyme- (Dan et al. 2000) and morpholino-based (Carpenter and Cande 2009). RNA silencing has been developed for *G. intestinalis*, but the techniques have not been widely applied in the field. Recently, CRISPRi has been introduced to *Giardia*, opening new possibilities in manipulating expression of even multiple genes at once (McInally et al. 2018). Due to polyploid nature of both organisms, gene knockouts are not very feasible method, although complete

knockout of *Giardia* cyst wall protein 1 was recently accomplished by employing Cre-Lox recombination to recycle selection marker for consecutive elimination of the four alleles (Ebneter et al. 2016).

A high-speed differential centrifugal fraction for *Giardia* mitosomes has been generated that has been successfully used to reconstitute Fe-S cluster formation (Tovar et al. 2003). Further enrichment of these fractions was obtained on sucrose gradients that yielded organelles of ~150 nm in diameter (Regoes et al. 2005; Dolezal et al. 2005). These organelles have been used for localization studies, but so far only limited protein import assays have been performed yet (Dagley et al. 2009). The proteome of mitosomes is far from complete, and only a few proteins have been physically localized in the mitosomes by specific polyclonal antibodies or by detection of their tagged recombinant versions. In vivo enzymatic tagging and immunoprecipitation techniques increased the number of newly identified mitosomal proteins, most of which remain of unknown function (Martincová et al. 2015; Rout et al. 2016). Gradient centrifugation of *E. histolytica* cellular lysates led to the purification of the mitosomes and the identification of unique sulphate-activation pathway within the organelles (Mi-ichi et al. 2009). While genetic manipulation of *Microsporidia* remains to be established, CRISPR/Cas9-based strategy was successfully introduced to *Cryptosporidium* parasites (Vinayak et al. 2015), although the experimental system still depends on a cumbersome propagation of the parasites. The likelihood of mitosomes being ever purified from *Microsporidia* or *Cryptosporidium* is not very high as the organelles are extremely small, ranging from 70 nm for the microsporidian *Trachipleistophora hominis* (Williams et al. 2002) to between 150 and 300 nm for the *C. parvum* mitosome (Riordan et al. 2003; Putignani et al. 2004). Moreover, the single *C. parvum* mitosome is entangled by the rough ER (Riordan et al. 2003; Putignani et al. 2004), which will render any disruption technique quite tricky. Despite these limitations, the studies of protein import into mitosomes of *G. intestinalis* (Regoes et al. 2005; Dolezal et al. 2005; Pyrihová et al. 2018), *E. histolytica* (Mai et al. 1999; Tovar et al. 1999) and of the two microsporidian species *Encephalitozoon cuniculi* and *Antonospora* (*Nosema*) *locustae* (Burri et al. 2006) brought exciting insight into the degree of functional similarity between mitosomes, hydrogenosomes and mitochondria, and also into the degree of adaptation of mitosomes within the microsporidia (Burri and Keeling 2007). Although both mitosomes and hydrogenosomes have arisen independently and repeatedly, the molecular basis of the reduced protein import machinery may offer clues as to the composition of the original sets of translocases installed in the membranes of proto-mitochondria.

4.2 Mining Genome Sequence Data

Thanks to the completed genome sequencing projects for hydrogenosomal and mitosomal species, e.g.(McArthur et al. 2000; Katinka et al. 2001; Abrahamsen et al. 2004; Xu et al. 2004; Loftus et al. 2005; Carlton et al. 2007), we have a

tremendous amount of information about the biology and evolution of these organisms at hand. Having all these data available, we are presented with the difficult issue of efficient data mining. Our attempts to identify possible homologous sequences in the genomes of evolutionary diverse species are very often faced with the danger of false-negative results and therefore of incorrect conclusions. The widely used approach is BLAST based on pairwise sequence analyses (Altschul et al. 1990, 1997). BLAST searches are sometimes inefficient simply because a particular query may be too divergent to pick the target sequence from genome databases. In the field of mitochondrial protein import, queries originate primarily from *S. cerevisiae* or other fungal sequences. While pairwise sequence analyses were sufficient to identify equivalent components in animals, they work less well on plants and often fail to identify homologous sequences in other phylogenetic groups, especially protists (Hoogenraad et al. 2002). If the whole family of proteins instead of a single sequence is available, a search based on the hidden Markov model (HMM) offers a significantly more sensitive mining method when compared with BLAST (Eddy 1998; Finn et al. 2011; Alva et al. 2016). In practice, analyses based on HMMs represent a reversed search of the protein family (PFAM) database (Bateman et al. 2004). Instead of comparing one query sequence with all the available HMMs in PFAM, a single HMM is used to search the genome database. Although HMMs were first designed for speech recognition, they can be applied to a variety of problems, where hidden parameters need to be determined from obvious parameters, such as sequence alignment of homologous proteins. The parameters that are extracted from the sequence alignment, for instance, the probability of occurrence of certain amino acids in a particular position, can be then used for mining data from the conceptual translation of a genome sequence. Depending on the selection of sequences for the alignment, HMMs can even pick structural information otherwise hidden in the primary sequence (Dolezal et al. 2006; Likic et al. 2010). For example, the alignment of some homologous β-barrel proteins might provide enough information to find any β-barrel protein in the examined genome. Importantly, newly identified homologous sequences can be included into the alignment used for the building of new HMM, thus providing a refined and more sensitive tool for the next round of searches. Usually several cycles of refinement are used to craft a reliable HMM that is powerful enough to pick very divergent homologous sequences, but that is insensitive to unrelated sequences. The freely available HMMER software (https://www.ebi.ac.uk/Tools/hmmer/) enable the building of tailored HMMs based on the user' s protein sequence alignment. A more recent addition to the list of search tools is HHpred (https://toolkit.tuebingen.mpg.de/#/tools/hhpred). HHpred is a sensitive server for detecting remote homologues based on protein function and protein structure (Söding et al. 2005). Thus, the efficient searches using single or several HMMs in a large database can now be performed by the typical bench-work biologist. Collectively, in silico studies and laboratory studies have begun to reveal much about hydrogenosomal and mitosomal targeting signals, translocases, chaperones and processing peptidases.

5 Organellar Targeting Signals

Targeting signals contain the minimal information necessary for a protein precursor to be recognized by targeting machinery and to be directed to the correct compartment in a cell. These sequences are both "necessary and sufficient" to target a passenger protein to a given organelle. Other sequence stretches within the precursor may be necessary to target the protein to the correct sub-organellar location (Neupert 1997). Several categories of targeting signals have been defined for mitochondrial precursors and, to a lesser extent, for hydrogenosomal and mitosomal precursors.

5.1 Mitochondrial Targeting Signals

The majority of mitochondrial precursors are synthesized with N-terminal cleavable presequences. Typical presequences contain 10–80-amino acid residues, many of which are positively charged, hydrophobic and hydroxylated (Pfanner and Geissler 2001). Negatively charged amino acid residues are notoriously absent in most presequences (von Heijne et al. 1989). Generally, the amino acid residues on presequences are disposed in an amphipathic α-helix, which has one hydrophobic surface opposed by a positively charged surface (Roise et al. 1986; Abe et al. 2000). These contrasting surfaces make contact sequentially with Tom and Tim translocases as the presequence-containing precursor traverses both in mitochondrial membranes (Pfanner and Geissler 2001). For instance, structural studies have shown that the hydrophobic surface of the α-helical presequence makes contact with the Tom20 binding groove (Abe et al. 2000). Subsequently, the precursor is passed over to Tom22 through interaction of the positive surface of the α-helix with negative charges on Tom22 (Brix et al. 1997) and continues through a "binding chain" by contacting the various translocases (Pfanner and Geissler 2001). Eventually, as the precursor reaches TIM23, the positive charges on the presequence are acted upon by the membrane potential that draws the precursor into the channel (Martin et al. 1991). Upon translocation through TIM23, the presequence is generally cleaved by MPP from the majority of proteins (Gakh et al. 2002), but there are exceptions where presequences remain an integral part of the mature protein (Rospert et al. 1993). Curiously, one case of a cleavable presequence at the C-terminus of a matrix protein has been reported, where the precursor is translocated in a C- to N-terminal rather than the common N- to C-terminal orientation (Lee et al. 1999). Following cleavage, if any, the precursors are then either imported for further folding into the matrix or released by TIM23 into the inner membrane if they additionally possess a hydrophobic stop-transfer signal (Glick et al. 1992; Beasley et al. 1993; Bömer et al. 1997). Some inner membrane and IMS preproteins contain a bipartite presequence that comprises an N-terminal positively charged matrix-targeting sequence and a downstream sorting signal that is similar to sorting signals found on bacterial and ER secretory proteins. These preproteins are first processed by MPP and then undergo a

second cleavage by IMP (Schneider et al. 1991; Nunnari et al. 1993). Other types of inner membrane proteins have internal targeting signals as a combination of transmembrane hydrophobic segments together with positively charged loops (Folsch et al. 1996; Davis et al. 1998). Multiple internal targeting signals that act cooperatively have been characterized for the inner membrane ADP/ATP carrier (AAC). No consensus sequence has been computed for these signals, but each segment of about 10-amino acid residues can be recognized individually by the Tom70 receptor (Brix et al. 2000; Wiedemann et al. 2001). In general, internal targeting signals for hydrophobic proteins are poorly characterized. Some outer membrane, monotopic proteins, like Tom70, have a non-cleavable presequence that directs the precursor to mitochondria and drives the insertion of a downstream hydrophobic stretch that acts as a membrane anchor (Hahne et al. 1994). Recent investigations have shown that the targeting signal for outer membrane β-barrel proteins resides in the β-hairpin motif positioned between any two β-strands of the protein. The signal seems to be recognized by Tom20 and in part by Tom70 on the mitochondrial surface.

5.2 *Signals on Precursors of Soluble Hydrogenosomal and Mitosomal Proteins*

Characterization of targeting signals within the hydrogenosomal and mitosomal proteins has mainly focused on putative N-terminal cleavable presequences. These are relatively straightforward to detect, by experimentally determining the N-terminal sequences on isolated endogenous organellar proteins and comparing those with the conceptual translation of the corresponding genes. Most of the putative N-terminal presequences on hydrogenosomal and mitosomal proteins have been either predicted by programmes that have been devised to search for mitochondrial presequences or by sequence comparison with eubacterial homologues. In many cases, the ability of these putative presequences to function as genuine targeting signals has been tested by assessing their efficiency in conducting passenger proteins to mitochondria or to the relevant organelle.

5.2.1 *Trichomonas* Hydrogenosomes

A hydrogenosomal N-terminal cleavable presequence was first noted in *T. vaginalis* ferredoxin, a matrix protein, when purified endogenous ferredoxin was found to lack 8-amino acid residues at the N-terminus, relative to the conceptual gene translation (Johnson et al. 1990). This presequence (Table 1) has an overall positive charge and is significantly shorter than typical mitochondrial targeting sequences which range from 10 to 80-amino acid residues (Pfanner and Geissler 2001). Using an in organello import assay, Bradley et al. (1997) demonstrated that deletion of this eight-aa sequence abolishes binding to, and thus translocation of the protein into,

t1.1 **Table 1** Presequences of precursors to hydrogenosomal and mitosomal proteins[a]

t1.2		N-terminal presequence	References
t1.3	*Trichomonas vaginalis*, hydrogenosomal matrix		
t1.4	Ferredoxin[b]	**MLSQVCRF**^GTITAVKGGVKK	Johnson et al. (1990)
t1.5	Succinyl coA synthetase, α-subunit 1	**MLAGDFSRN**^LHKPLLFIDKD	Lahti et al. (1994)
t1.6	Succinyl coA synthetase, α-subunit 2	MLSSSFERN^LHQPLLFIDKD	Lahti et al. (1994)
t1.7	Succinyl coA synthetase, α-subunit 3	MLSSSFERN^LHQPLLFIDKD	Lahti et al. (1994)
t1.8	Succinyl coA synthetase, β-subunit	**MLSSSFARN**^FNILEWQSKEI	Lahti et al. (1992)
t1.9	Adenylate kinase	**MLSTLAKRF**^ASGKKDRMVVF	Länge et al. (1994)
t1.10	Chaperonin60	MSLIEAAKHFTRAF^AKARDL	Bui et al. (1996)
t1.11	Pyruvate:ferredoxin oxidoreductase subunit A, PFORA	**MLRSF**^GKRIPGDGNTAATSV	Hrdy and Müller (1995a)
t1.12	PFORB	MLRNF^SKRVPGDGNTAATSV	Hrdy and Müller (1995a)
t1.13	Malic enzyme A	MLTSSVSVPVRN^ICRAKVPT	Hrdy and Müller (1995b)
t1.14	Malic enzyme B	**MLTSSVNFPARE**^LSRNVRPT	Hrdy and Müller (1995b)
t1.15	Malic enzyme C	**MLTSVSYPVRN**^ICRSKLPLA	Hrdy and Müller (1995b)
t1.16	Malic enzyme D	**MLTSVSLPVRN**^ICRSKLPVA	Hrdy and Müller (1995b)
t1.17	Isd11[c]	MLSSFLSRTF>ANESVMANLRES	Richards and van der Giezen (2006)
t1.18	[Fe]-hydrogenase maturase HydG[d]	MLASLSRS[]YGKLRADVSKTL	Pütz et al. (2006)
t1.19	[Fe]-hydrogenase maturase HydF	MLTSIGRY[]FAKKGNDLPRTH	Pütz et al. (2006); our unpublished data
t1.20	[Fe]-hydrogenase maturase HydE	MSHDHIVRL[]LNPRTKDEIDA	Pütz et al. (2006); our unpublished data
t1.21	Ferredoxin[e]	MLSQCSPLRF()GSVTVTKGGA	Dolezal et al. (2005)
t1.22	Iron-sulfur cluster assembly protein IscS-1	MLTNLYNKA()FHGHYLDAQATSI	Tachezy et al. (2001), Sutak et al. (2004)
t1.23	IscS-2	MLGSVSRS()YFKGHYLDTQATSV	Tachezy et al. (2001), Sutak et al. (2004)
t1.24	Iron-sulfur cluster assembly protein IscU	MLAAVSRS()SALNMMKPLGIM	Dolezal et al. (2005)
t1.25	Complex I protein Ndh51/NuoF	MLAAYGHRF()QTKFLDPKDRI	Dyall et al. (2004), Hrdy et al. (2004)

t1.26	Complex I protein Ndh24/NuoE	MLASVNTSRF()FARLNKKS>VL	Dyall et al. (2004), Hrdy et al. (2004)
t1.27	Glycine cleavage H protein H1[f]	MITSCFTRA*AKQYSKDHLWF	Mukherjee et al. (2006a)
t1.28	Glycine cleavage H protein H2	MISTLCNCSRNF*TKLYAKT>H	Mukherjee et al. (2006a)
t1.29	Serine hydroxymethyl transferase SHMT	MLKNVFHRF*SSSWILSEKVL	Mukherjee et al. (2006b)
t1.30	Malic enzyme G	MLTSVSLPVRN*ICRSKLPVA	Dyall et al. (2004)
t1.31	Malic enzyme H	MLTSSVSLPARE*LSRKVLPT	Dyall et al. (2004)
t1.32	[Fe]-hydrogenase A	MLASSATAMKGFANSLRM*KD	Bui and Johnson (1996)
t1.33	[Fe]-hydrogenase B	MLASSSRAAANIRW()VDTSHN	Bui and Johnson (1996); our unpublished data
t1.34	DnaK/Hsp70	MLSSVARSTSSLFSRG()FAAG	Bui et al. (1996), Germot et al. (1996), Dyall et al. (2003)
t1.35	Chaperonin 10	MLATFARNF*AAKKVTIKPLG	Bui et al. (1996); our unpublished data
t1.36	Pam18	MSIVNKF*VEKALSLPTYAKA	Dolezal et al. (2005)
t1.37	Mitochondrial processing peptidase β-MPP	MSIISRY*AVPQISKLSNGVRV()	Dolezal et al. (2005)
t1.38	*Trichomonas vaginalis*, hydrogenosomal membrane		
t1.39	Mitochondrial carrier family protein Hmp31	**MAQPAEQILIAT**^SPKPSLSP	Dyall et al. (2000)
t1.40	Hmp31-a[g]	MATEADKVLIAT*SPNGALPT	Our unpublished data; Acc EAX95270
t1.41	Hmp31-b	MKPADKILIAT*SPSDAKLKP	Our unpublished data; Acc EAY15971
t1.42	Hmp31-d	MKIKFSFGQKQKKDNL*SPVQ	Our unpublished data; Acc EAY00062
t1.43	*Neocallimastix frontalis*		
t1.44	Malic enzyme	**MLAPIQTIARPVSSILPATGALAAKRT**^FFA	van der Giezen et al. (1997)
t1.45	Succinyl coA synthetase, β-subunit	MLANVTRSTSKAAPALASIAQTAQKRF*LSV	Brondijk et al. (1996)
t1.46	Neocallimastix sp L2		
t1.47	[Fe] hydrogenase	MSMLSSVLNKAVVNPKLTRSLATAAEKMVNISINGRKF()QV	Voncken et al. (2002)

(continued)

t1.48 **Table 1** (continued)

t1.49		N-terminal presequence	References
t1.50	*Neocallimastix patriciarum*		
t1.51	Chaperonin 60	MLSARSLICKSMIKSGFRRAVAPSVAMAASSMTLTARRNY*SSKY	van der Giezen et al. (2003)
t1.52	DnaK/Hsp70	MFLSTLAKKSTTFGVSNVVKNALSSKVMRTTPRMFQRF*ESSK	van der Giezen et al. (2003)
t1.53	*Giardia intestinalis*		
t1.54	IscU	**MTSLQLSSTSLLQSVARF^**LTKKTSSDEVYSE	Tovar et al. (2003), Dolezal et al. (2005)
t1.55	Ferredoxin	MSLLSSIRRFITFRV()VQQ>	Nixon et al. (2002), Dolezal et al. (2005)
t1.56	Pam18	MLRVLSENRF*PLSLVAGVVA	Dolezal et al. (2005)
t1.57	β-MPP	MLIRD()IVPGALPS>ATVVFSG	Dolezal et al. (2005)
t1.58	*Entamoeba histolytica*		
t1.59	Chaperonin 60	**MLSSSSHYNGKLLSLNIDCRE^**NVL	Clark and Roger (1995), Mai et al. (1999), Tovar et al. (1999)
t1.60	*Antonospora locustae* putative *mitosome*		
t1.61	mtG3PDH	MINKRTYTYAFAAIGTGVLGYVGHRYYRHRK*	Burri et al. (2006)
t1.62	*Cryptosporidium parvum*		
t1.63	Chaperonin 60	MLLRSGINLYKSVEGSIGLRSAAIRFGMRYISSGKE>()LSF	Riordan et al. (2003)
t1.64	DnaK/Hsp70	MSMIINSSFNGVVNSSGIAARILKRSLPLVFSRY[]MSSK	Slapeta and Keithly (2004)
t1.65	IscS	MIVHRYCRQWAPSVVRGISKLAFFSSMSSIAKKRPAY>[]FDY	LaGier et al. (2003)
t1.66	IscU	MLQLRQLIDKRILIKKCVPICQRLFYS>DTV	LaGier et al. (2003)

t1.67 [a]This survey only shows proteins that have been shown to localise to hydrogenosomes, mitosomes or heterologous mitochondria, as reported in the literature, or from unpublished proteomics data

[b]Experimentally determined presequences are shown in bold with the cleavage site marked with ^. Arg residues at the -2 or -3 position relative to the cleavage site are underlined

[c]Cleavage sites predicted by MitoProtII (Claros and Vincens 1996) are indicated by >

[d]Cleavage sites predicted by iPSORT (Bannai et al. 2002) are indicated by [ ]

[e]Cleavage sites predicted by PSORTII (Nakai and Horton 1999) are indicated by ()

[f]Cleavage sites suggested by authors or suggested here are indicated by *

[g]Acc = GenBank accession number

the hydrogenosome (Bradley et al. 1997). This result has been confirmed in vivo, when ferredoxin that lacked residues 2–8 was expressed in *T. vaginalis*, and was found to reside exclusively in the crude cytosolic fraction (Dyall et al. 2000). This suggests an important role for the presequence in binding to any receptor and/or pore that promotes entry into the hydrogenosome. In the last 10 years, dozens of *T. vaginalis* hydrogenosomal presequences have been characterized or predicted for proteins involved in various pathways (Table 1). The emerging picture is that, unlike mitochondrial and plastidic presequences, the *T. vaginalis* hydrogenosomal presequences are highly conserved at primary sequence level. This is even more striking upon examination of over a hundred soluble protein sequences identified during proteomic studies, where about 75% of the translated gene sequences have N-terminal sequences that closely resemble those shown in Table 1 (our unpublished data). How these presequences have been appended and are so well conserved remain a mystery, but it is possible that, as is the case for newly transferred angiosperm mitochondrial genes (Adams et al. 2000; Choi et al. 2006), a small subset of hydrogenosomal protein genes has preferentially been used as presequence donor.

The *T. vaginalis* hydrogenosomal presequences are generally short, ranging from 5 to 14-amino acid residues for those that have been proven experimentally and up to 17 residues for the predicted presequences (Table 1). The presequences are enriched in the amino acid residues Ser (20%), Leu (14%), Arg (11%), Ala (8%), Phe (7%), Val (6%), Thr (6%) and Asn (5%). The other amino acids are significantly underrepresented. Incidentally, or accidentally, the three amino acids most commonly found in these presequences, Ser, Leu and Arg, are the ones that are each encoded by six codons. This may have been relevant in the evolution of these presequences. The mitochondrial matrix N-terminal presequences are enriched in Arg (14%), Leu (12%), Ser (11%) and Ala (14%). On the other hand, chloroplast leader peptides have a different amino acid composition with 19% Ser and 9% Thr (von Heijne et al. 1989). Markedly underrepresented in hydrogenosomal presequences are the acidic residues, as in the case of both mitochondrial and plastidic presequences (von Heijne et al. 1989).

Three of the frequently occurring amino acid residues in hydrogenosomal presequences are positionally conserved as well. Of the 13 hydrogenosomal matrix preproteins for which presequence cleavage sites have been experimentally determined, 12 have Leu at position 2 of the presequence, and the exception has a Leu residue at position 3 (Table 1). Thus, not only the presence but also the position of the Leu residue is conserved. This is even more striking when we examine predicted N-terminal presequences for a further 20 proteins that have been localized to hydrogenosomes, where 85% have Leu at position 2. Mutation of the Leu residue at position 2 in the ferredoxin presequence disrupted binding of the protein precursor to hydrogenosomes (Bradley et al. 1997), suggesting that this particular residue plays a critical role in binding. The Arg residue occurs at the -2 or -3 position relative to the cleavage site in all the experimentally determined presequences, with 77% at the -2 position. Phe residues can be frequently found in the vicinity of the Arg residue. Interestingly, many, but not all, mitochondrial N-terminal presequences

likewise contain Arg at -2 or -3 from the cleavage site (Gavel and von Heijne 1990), but the role of the residue in specifying the cleavage site for MPP is unclear (Gakh et al. 2002). The exact role of these conserved residues, i.e. whether they are important for binding, for translocation or for cleavage, is not known. Nonetheless, some of these conserved features were applied to devise consensus sequences that were used to screen the *T. vaginalis* genome sequence database. A genome-wide search using the consensus sequences M-L-(S/T/A)-$x_{(1-15)}$-R-(N/F/E/xF), M-S-L-$x_{(1-15)}$-R-(N/F/xF) or M-L-R-(S/N)-F picked out 138 sequences with 67% showing similarity to known proteins involved in metabolic pathways, electron transport, protein import, protein folding and oxygen scavenging pathways (Carlton et al. 2007). There are undoubtedly variations on these consensuses, as have been found during proteomic studies (our unpublished data).

Apart from a similar amino acid enrichment, a common feature of these hydrogenosomal presequences and of the mitochondrial N-terminal presequences is their ability to form amphipathic α-helices (Johnson et al. 1990; Lahti et al. 1992; Dolezal et al. 2005). The amphipathic α-helical structure within mitochondrial N-terminal presequences has been shown to be critically important for sequential electrostatic or hydrophobic interaction with various translocases (Pfanner and Geissler 2001). The hydrogenosomal presequence may be interacting with hydrogenosomal translocases using a similar "binding chain" mechanism. To date, it has not been demonstrated that the typical *T. vaginalis*-soluble preprotein presequence is sufficient for translocating the protein into hydrogenosomes, although it has been shown to be necessary (Bradley et al. 1997; Dyall et al. 2000). It is possible that there are additional downstream signals in the mature part of the protein that participate in translocation at stages beyond binding. It has been shown, however, that hydrogenosomal presequences can target a passenger protein into *Trypanosoma brucei* and *S. cerevisiae* mitochondria, but at very low efficiency (Hausler et al. 1997). Some hydrogenosomal matrix proteins, α-subunit of succinyl-CoA synthetase (SCS), Fdx1, malic enzyme and IscA without their predicted N-terminal hydrogenosomal targeting sequence (HTS) were shown to be targeted to the hydrogenosomes suggesting that HTSs are non-essential for targeting (Zimorski et al. 2013). Further, the same set of proteins were found to be localized in mitochondria when they were expressed in yeast without their HTS (Garg et al. 2015). An unusual case is that of ATP-dependent PFK1, which is primarily a glycolytic enzyme that was found to be present in the proteome of hydrogenosomes. Further investigations revealed that it localizes to hydrogenosomes, although it has no predictable HTS, and when expressed in yeast, the protein was targeted to mitochondria (Rada et al. 2011, 2015). Such cases point out that the N-terminal presequence-independent pathway exists for the import of some hydrogenosomal matrix proteins in *T. vaginalis*, and this feature seems to be conserved in yeast as well. In hydrogenosomes, the loss of the respiratory chain complexes and the membrane potential led to the loss of positive charge on the HTS, and subsequently, the HTS might have become dispensable or, in certain cases, the HTS itself was lost (Garg et al. 2015; Rada et al. 2015). It has been hypothesized that this could represent a "primitive" or an ancestral route of protein import into mitochondria.

However, the presence of cryptic or internal targeting sequence in these matrix proteins cannot be ruled out.

5.2.2 *Neocallimastix* Hydrogenosomes

Although there have been several reports of putative *Neocallimastix* sp. hydrogenosomal proteins in silico, few have actually been localized to hydrogenosomes or to heterologous mitochondria (Brondijk et al. 1996; van der Giezen et al. 1997; Voncken et al. 2002; van der Giezen et al. 2003). These proteins have quite similar N-terminal extensions (Table 1), but only one of them has been experimentally confirmed (van der Giezen et al. 1997). The extensions range from 27 to 37-amino acid residues, and are within the range for typical mitochondrial N-terminal presequences, with similar amino acid enrichment and characteristics. Indeed, when expressed in yeast, the hydrogenosomal malic enzyme was targeted to mitochondria in a presequence-dependent fashion (van der Giezen et al. 1998). The predicted N-terminal presequences on *N. patriciarum* Cpn60 and Hsp70 were sufficient to target the green fluorescent protein (GFP) to mammalian mitochondria, although some non-specific targeting was observed for the Cpn60 presequence, suggesting that additional signals are present in the mature part of the protein (van der Giezen et al. 2003).

5.2.3 *Cryptosporidium* Mitosomes

The complexity of mitochondria-related organelles differs significantly among different *Cryptosporidium* species with the most reduced mitosomes found in *C. parvum*, *C. hominis* and *C. ubiquitum* (Liu et al. 2016). Only four proteins have been experimentally localized to mitosome of *C. parvum*. Cpn60 has a putative 38-aa N-terminal presequence (Table 1) which does not follow the Arg -2 rule, but the N-terminal 57-amino acid portion of Cpn60 was necessary and sufficient to target GFP into yeast mitochondria (Riordan et al. 2003). Likewise, the predicted N-terminal extensions on mitosomal IscU and IscS (Table 1) were both sufficient to target GFP to yeast mitochondria (LaGier et al. 2003). The predicted 34-amino acid presequence of Hsp70 closely resembles typical mitochondrial presequences with a predicted amphipathic α-helical domain and similar enrichment in amino acids and has an Arg-2 cleavage site motif (Gavel and von Heijne 1990; Slapeta and Keithly 2004). This predicted presequence could specifically deliver GFP into yeast and *Toxoplasma gondii* mitochondria, and it was shown that the specific presequence region critical for targeting included the predicted amphipathic α-helical domain (Slapeta and Keithly 2004).

5.2.4 *Entamoeba* Mitosomes

Not many *E. histolytica* mitosomal proteins have been identified, leaving us with very little information on protein targeting signals. The only experimental data come from the analysis of Cpn60 that has an N-terminal extension of 15 amino acids (Table 1) shown to be important for mitosomal targeting (Mai et al. 1999; Tovar et al. 1999). This presequence, like most of the *T. vaginalis* presequences, has a Leu residue at position 2 and is highly enriched in Ser residues. While deletion of the extension leads to the accumulation of Cpn60 in the cytosol, the swapping of the extension with the N-terminal presequence from mitochondrial Hsp70 of *Trypanosoma cruzi* delivers the protein back into the enriched mitosomal fraction (Tovar et al. 1999). On the other hand, the components of mitosomal sulphate activation pathway (Mi-ichi et al. 2009) as well the orthologue of Cpn10 (Chan et al. 2005) lack any recognizable N-terminal extension, which thus indicates existence of so far unknown internal targeting signals.

No further functional data are currently available on the processing of targeting presequences in *E. histolytica.*

5.2.5 *Giardia* Mitosomes

A number of proteins have been successfully localized to *G. intestinalis* mitosomes (Tovar et al. 2003; Regoes et al. 2005; Dolezal et al. 2005; Rada et al. 2011; Rout et al. 2016; Pyrihová et al. 2018). The import of giardial homologues of IscU and [2Fe-2S] ferredoxin (Table 1) was shown to be dependent on the N-terminal targeting sequence (Regoes et al. 2005; Dolezal et al. 2005) as their truncated versions were mislocalized and/or degraded in the cytosol. These two presequences are enriched in Ser, Thr, Leu and Arg which are very similar to the *T. vaginalis* presequences. The N-terminal sequences of IscU and [2Fe-2S] ferredoxin, extending beyond the respective predicted presequence cleavage sites, were sufficient to target GFP into mitosomes. The increased electrophoretic mobility of the fusion protein in organellar fractions suggested that the N-terminal presequences were removed upon targeting (Regoes et al. 2005). These two targeting sequences consist of 15–18-amino acid residues that can be projected to form amphipathic α-helices (Dolezal et al. 2005). Interestingly, the gene coding for [2Fe-2S] ferredoxin was demonstrated to contain a spliceable intron just between the exons coding for the N-terminal targeting sequence and the mature ferredoxin (Nixon et al. 2002).

Other soluble proteins that have been localized in mitosomes have recognizable N-terminal presequences (Tovar et al. 2003; Regoes et al. 2005; Dolezal et al. 2005; Rada et al. 2011; Rout et al. 2016; Pyrihová et al. 2018). However, some proteins like IscS and Cpn60 seem to rely on internal signals, which is quite unusual since these mitochondrial proteins typically contain cleavable presequences in other studied organisms. This was experimentally demonstrated when *G. intestinalis* IscS was expressed as a 202-aa N-terminal polypeptide and a 232-aa C-terminal

polypeptide, and both truncated proteins could be successfully delivered to organelles (as tested on *T. vaginalis* hydrogenosomes) showing that targeting information is found in multiple loci within the protein (Dolezal et al. 2005). Deletion of the first five amino acids on *G. intestinalis* Cpn60 did not affect the targeting of the protein to mitosomes (Regoes et al. 2005). Thus, *G. intestinalis* mitosomes display both presequence-dependent and presequence-independent targeting for soluble preproteins (Regoes et al. 2005; Dolezal et al. 2005).

The targeting information on mitosomal proteins can be recognized and processed by the heterologous systems of human and yeast mitochondria, as well as *T. vaginalis* hydrogenosomes (Regoes et al. 2005; Dolezal et al. 2005). The [2Fe-2S] ferredoxin N-terminal presequence was sufficient to deliver a passenger protein into human mitochondria (Regoes et al. 2005), and *T. vaginalis* hydrogenosomes can specifically import *G. intestinalis* [2Fe-2S] ferredoxin, IscU, IscS, Pam18 and β-MPP (Dolezal et al. 2005). Furthermore, the presequence on IscU was sufficient to efficiently target a passenger protein into *T. vaginalis* hydrogenosomes. The N-terminal presequence of IscU can be processed by *S. cerevisiae* mitochondrial extract and also by purified recombinant rat MPP (Dolezal et al. 2005). Altogether, these results strongly suggest that targeting information on *G. intestinalis* mitosomal proteins can be cross-recognized by the respective protein import machineries of mitochondria and hydrogenosomes.

5.2.6 Microsporidia Mitosomes

So far, multiple proteins predominantly of ISC pathway have been localized in situ in *T. hominis* and *E. cuniculi* mitosomes (Williams et al. 2002; Goldberg et al. 2008; Freibert et al. 2017). The presence of N-terminal targeting sequences was not thoroughly investigated in these two species, but at least some *E. cuniculi* protein have predicted cleavable N-terminal presequence (Katinka et al. 2001).

Instead, analysis of *A. locustae* genome has provided interesting insight and surprising differences in mitosomal protein import mechanisms (Slamovits et al. 2004). Of the identified mitosomal proteins, only a handful has amphipathic N-terminal presequences, and others do not appear to have any extensions nor have many characteristics in common (Burri et al. 2006). As no genetic transformation technique has been developed as yet for microsporidia, the targeting information on these proteins was investigated by expressing the full-length and truncated versions of these proteins in *S. cerevisiae* as fusions with GFP (Burri et al. 2006). Of the 16 proteins under investigation, only 6, most from *A. locustae*, could direct GFP to mitochondria. Deletion of the N-terminal predicted extensions from four of these fusion proteins disrupted targeting to mitochondria, showing that the extensions are necessary for cross-organellar targeting. The other two proteins, including mitochondrial glycerol-3-phosphate dehydrogenase (mtG3PDH), could still be delivered to mitochondria, suggesting that internal targeting signals are sufficient for targeting. However, the N-terminal sequence of mtG3PDH (Table 1) was also found to be sufficient to deliver GFP to yeast mitochondria. The N-terminal

extensions from the other proteins were not sufficient to target GFP to mitochondria. This finding undermines the exclusive role of the N-terminal sequence in organellar protein targeting. It is apparent that a combination of N-terminal and mainly internal signals seems to fulfil the targeting role in microsporidian mitosomes (Burri et al. 2006).

5.3 Signals on Hydrogenosomal and Mitosomal Membrane Proteins

Targeting signals on most membrane proteins are generally poorly characterized. Not only do membrane protein precursors require targeting, membrane sorting and insertion signals, but they also require a means of protection against premature folding or aggregation in the hydrophilic environments they encounter during transport to the organelle membrane. A variety of membrane proteins are targeted to mitochondria: β-barrels, tail-anchored and α-helical polytopic and monotopic proteins have been characterized. Given this diversity in structure, specific but sometimes overlapping pathways are utilized for their insertion (Rehling et al. 2003; Koehler 2004; Bohnert et al. 2007). Most of the data available on membrane protein insertion has been generated for members of the mitochondrial carrier family (MCF), particularly for AAC, the model precursor.

The *T. vaginalis* hydrogenosomal Hmp31 precursor protein, a member of the MCF, was found to have a cleavable 12-amino acid N-terminal presequence. Although this sequence is predicted to be mostly α-helical, it does not have an amphipathic disposition but has an overall negative charge. This presequence was found not to be necessary for targeting and integration of mature Hmp31 in the membrane, suggesting that Hmp31 utilizes internal targeting signals, like virtually all MCF proteins. However, the presequence was necessary, and sufficient, to target a passenger protein to the soluble hydrogenosomal fraction and as such acted as a targeting signal. Thus, the Hmp31 precursor has internal targeting signals and a functional N-terminal targeting signal (Dyall et al. 2000). Four more Hmp31 orthologues were found during proteomic analyses (our unpublished data), and two of those, Hmp31-a and Hmp31-b, were each found to have a similar N-terminal extension (Table 1). Another orthologue, Hmp31-d, however, had a putative N-terminal extension that resembles the matrix-targeting N-terminal presequence, with an overall positive charge, and the fourth one does not appear to have an N-terminal extension. However, none of these Hmp31 orthologues have had their N-termini experimentally determined as yet. No such presequence has been found on either the *Neocallimastix* hydrogenosomal AAC (van der Giezen et al. 2002) or on the *Entamoeba* mitosomal AAC (Chan et al. 2005). Although most MCF proteins are synthesized without N-terminal extensions, a subset of precursors has cleavable presequences. Plant mitochondrial AACs are synthesized with long N-terminal presequences, but these are both not necessary and not sufficient to target

passenger proteins to mitochondria and are therefore not acting as targeting signals (Glaser et al. 1998; Murcha et al. 2005a). Another MCF protein, the mammalian phosphate carrier, bears a presequence that may act as an enhancer for translocation but is not strictly necessary, though it was marginally sufficient to target a passenger protein to mitochondria (Zara et al. 2007). It has been suggested that the dispensable N-terminal presequences of mammalian and fish citrate carriers may in fact act as chaperones to increase the solubility of the preprotein in the cytosol through electrostatic interaction (Zara et al. 2007).

All MCF members that have been characterized in hydrogenosomes or mitosomes have been successfully imported into yeast mitochondria (Dyall et al. 2000; van der Giezen et al. 2002; Chan et al. 2005). Therefore, all three precursors must have targeting signals that are compatible with the specific mitochondrial pathway used for mitochondrial carriers (Rehling et al. 2003). Indeed, *T. vaginalis* Hmp31 imported into mitochondria was found not only to be dependent on membrane potential but also on the presence of the small TIM chaperone complex (Fig. 1) that is essential for proper mitochondrial AAC translocation (Dyall et al. 2000). Conversely, mitochondrial AAC was efficiently targeted to *T. vaginalis* hydrogenosomes, showing that targeting signals are compatible between the two systems (Dyall et al. 2000). The targeting of β-barrel membrane proteins is conserved between hydrogenosomes and mitochondria, as a unique hydrogenosomal β-barrel protein, Hmp35, could be targeted to mitochondrial membranes where it associated with, or assembled into, a high molecular weight complex (Dyall et al. 2003). It is notable that β-barrel precursors from eubacteria and plastids can be successfully imported and assembled into mitochondria as well (Röhl et al. 1999; Müller et al. 2002). Thus, targeting and insertion pathways for β-barrel proteins appear to be conserved between eubacteria, mitochondria, plastids and hydrogenosomes. *T. vaginalis* hydrogenosomes may be using a conserved SAM-like pathway (Fig. 1) for insertion of β-barrel proteins, as a homologue of Sam50 has been discovered in the *T. vaginalis* genome (Dolezal et al. 2006). Proteomic analysis of *T. vaginalis* hydrogenosomes had shown the presence of 12 tail-anchored (TA) proteins (Rada et al. 2011). Mitochondrial TA proteins carry a single transmembrane domain at their C-terminus, and their targeting signal often reside in the TMD and its flanking region (Horie et al. 2002). In yeast, the overall charge of the flanking regions or the hydrophobicity determines the destination of the protein as TA proteins are present in ER and peroxisomes as well. However, in *T. vaginalis* since peroxisomes are absent, the TA proteins are targeted either to the hydrogenosomes or the ER. A global search for *T. vaginalis* TA proteins and localization experiments have shown that the upstream regions of the TMD of hydrogenosomal TA proteins have low hydrophobicity compared to those of ER TA proteins (Rada et al. 2019). The TMD flanking regions of TA proteins contain lysine and arginine residues and, thus, have an overall positive charge. Replacement of lysines with serines or extension of the upstream region of the TMD by multiple valines mistargeted the hydrogenosomal TA protein to the ER (Rada et al. 2019).

6 Crossing the Organellar Membranes

All hydrogenosomes and mitosomes examined to date appear to have double membranes, which implies the presence of an intermembrane space. Given that hydrogenosomal and mitosomal preproteins bear signals that are recognized by the mitochondrial protein import machinery, it is likely that some components are phylogenetically and/or functionally conserved between these organelles. In hydrogenosomes and mitosomes, only the core subunits that are conserved in most eukaryotes are readily identifiable, while some subunits that are found in animals and fungi are either absent or too diverged to be identified through sequence-based searches. To date, a very limited number of hydrogenosomal and mitosomal protein import components have been functionally characterized, and some putative players have been identified in the genomes of the protists through sequence comparison with mitochondrial translocases from various species. More insight into potential import processes can be gained by examining in greater detail how mitochondrial preproteins interact with translocases to cross organellar membranes.

6.1 The Outer Membrane

Two major protein import machineries have been characterized to date in the mitochondrial outer membrane: the TOM and the SAM complexes (Fig. 1).

6.1.1 Translocase of the Outer Membrane (TOM Complex)

Most mitochondrial proteins enter mitochondria through a general import pore, the TOM complex. In yeast mitochondria, this complex consists of a pore-forming β-barrel Tom40 and six α-helical proteins: two primary receptors Tom70 and Tom20, core receptor Tom22 and three small Toms Tom5, Tom6 and Tom7 (Pfanner and Geissler 2001; van der Laan et al. 2006a; Bohnert et al. 2007).

Tom70 is the preferred receptor for hydrophobic preproteins with or without presequences (Wiedemann et al. 2001; Chan et al. 2006), although Tom20 also participates in binding (Brix et al. 1997). A typical substrate for Tom70 is the precursor to AAC. AAC has multiple internal targeting signals that are recognized by several Tom70 dimers, which probably act to prevent aggregation of these hydrophobic precursors (Brix et al. 2000; Wiedemann et al. 2001). Preproteins with N-terminal presequences initially make contact with Tom20 (Söllner et al. 1989). This interaction occurs through the hydrophobic surface of the amphipathic helix formed by the presequence, as demonstrated by structural studies (Abe et al. 2000). Thereafter, the two surfaces of the presequence are differentially recognized by binding domains of increasing affinity within the downstream translocases (Pfanner and Geissler 2001). Preproteins from both Tom20 and Tom70 are

subsequently transferred to Tom22, which acts both as a docking point for Tom20 and Tom70 and as a central receptor for preproteins within the TOM complex (Hönlinger et al. 1995; Bolliger et al. 1995; Brix et al. 1997; van Wilpe et al. 1999).

The cytosolic domain of Tom22 interacts with the positively charged surface of the amphipathic helix formed by N-terminal presequences (Brix et al. 1997). Next, the small protein Tom5 transfers preproteins from Tom22 to the Tom40 channel for translocation across the outer membrane (Dietmeier et al. 1997; Künkele et al. 1998; Hill et al. 1998). Besides making up the channel, Tom40 also has a binding site for presequences (Hill et al. 1998). After they pass through the channel, presequence-containing precursors bind to the IMS domain of Tom22 through the positive surface of the presequence and are subsequently sorted to the TIM23 complex. Therefore, a typical N-terminal presequence is recognized at least five times by Tom proteins, through either hydrophobic or ionic interactions (Pfanner and Geissler 2001; Bohnert et al. 2007). Following passage through the Tom channel, other types of preproteins are sorted into their respective specialized biogenesis pathways. Precursors to outer membrane β-barrel and inner membrane carrier proteins are guided by the small TIM chaperone complexes to their respective SAM or TIM22 pathway. Precursors to the small Tims and to other IMS proteins are taken up into the MIA pathway for further processing (Bohnert et al. 2007). Small Toms, Tom5, Tom6 and Tom7, are involved in the assembly and disassembly of the complex (Model et al. 2001).

Structural studies for TOM have been conducted so far only in two organisms, *S. cerevisiae* and *Neurospora crassa*. As observed via cryo-electron microscopy, the ~550 kDa holo TOM complex (trimeric) measuring around 140 Å has three protein translocation channels with each pore measuring around 20 Å, while the core TOM complex (dimeric) measuring 120 Å has two channels (Model et al. 2008; Bausewein et al. 2017). It has been speculated that the trimeric structure seems to be generic for the TOM complex in all forms of mitochondria (Fukasawa et al. 2017). However, this hypothesis needs to be tested in representative organisms that bear divergent and reduced forms of mitochondria. The TOM complex is a highly dynamic structure with the trimeric and dimeric forms switching alternately during the biogenesis of a new Tom40 (Shiota et al. 2015).

Given the intricacy and specificity displayed by the yeast mitochondrial protein import machinery, one might expect that the outer membrane translocases, or Tom proteins, would be conserved across species. Moreover, the demonstrated ability to successfully and specifically import mitosomal and hydrogenosomal preproteins into mitochondria led many to infer that similar and phylogenetically related receptors were present in hydrogenosomes and mitosomes as in mitochondria. This inference has in turn been used as supporting evidence for a common origin for mitochondria and related organelles. However, recent sequence surveys of complete genome databases have taught us that to start with, not all Tom proteins are conserved across all mitochondrial species (Maćasev et al. 2004; Likić et al. 2005; Chan et al. 2006; Perry et al. 2006), let alone mitosomal or hydrogenosomal species. Indeed, a comprehensive survey of available completed eukaryotic genomes revealed that only Tom7, Tom22 and Tom40 sequences are conserved among the

majority of eukaryotes, including animals, plants, fungi and some protists (Maćasev et al. 2004). Other components such as Tom20 and Tom70 have only been found in the genomes of animals and fungi so far, although a functional homologue of Tom70 was found in *Blastocystis* sp. and its homologues were found in other Stramenopiles (Likić et al. 2005; Chan et al. 2006; Tsaousis et al. 2011; Fukasawa et al. 2017). Although "Tom20" has been named and functionally characterized in plants (Heins and Schmitz 1996; Werhahn et al. 2001), it bears no primary sequence similarity to the fungal and animal Tom20 sequences and is likely to be of independent origin (Likić et al. 2005; Perry et al. 2006). Strikingly, though, the plant Tom20 has similar but oppositely orientated structural domains to the fungal Tom20, which appear to fulfil similar functions (Abe et al. 2000; Likić et al. 2005; Perry et al. 2006). These observations, taken together, have led to the hypothesis that the mitochondrial ancestor to eukaryotes had invented a core TOM complex consisting of Tom40, Tom22 and Tom7 and that other components subsequently evolved independently in the descendants as they progressively tweaked their respective mitochondrial protein import apparatuses (Maćasev et al. 2004; Dolezal et al. 2006). Of note is the TOM complex in the excavate *Trypanosoma brucei* termed archaic translocase of the outer membrane (ATOM) that is composed of highly diverged Tom40 (ATOM40) and a partially conserved Tom22-like protein (ATOM14) and four other subunits—two receptors, ATOM69 and ATOM46, ATOM11 and ATOM12 (Mani et al. 2015).

A proteomic study of the *Trichomonas* hydrogenosomes reported the presence of seven Tom40-like proteins that belong to the mitochondrial porin superfamily (Rada et al. 2011). A highly divergent TvTom40 (Isoform-2), present in a high molecular weight complex of ~570 kDa in the hydrogenosomal outer membrane, mediates the translocation of proteins across the OM (Makki et al. 2019). TvTom40-2 is associated with other isoforms of TvTom40, four tail-anchored proteins and Sam50, the core protein involved in β-barrel biogenesis (Makki et al. 2019). Two of the TA proteins, namely, Tom36 and Tom46, carrying an N-terminal Hsp20 chaperone-like and three TPR-like domains were shown to bind two hydrogenosomal preproteins, Fdx1 and α-SCS, and, hence, can function as TOM receptors. The other two TA proteins include a Tom22-like protein with a predicted molecular weight of 6.4 kDa which is present in the TvTOM that has short *cis* domain (cytosolic) and a conserved Tom22 TMD but lacks a *trans* domain (intermembrane space) and Homp19, which has no homology (Makki et al. 2019). Visualization of TvTOM via electron microscopy revealed a triplet-pore structure with an unconventional skull-like shape. EM and biochemical data suggest that TvTOM can associate with Sam50 (Makki et al. 2019).

Entamoeba histolytica mitosomes have a conserved Tom40 and a receptor protein named Tom60 that are part of a ~600 kDa TOM complex (Makiuchi et al. 2013). In the case of mitosomes in *Giardia intestinalis*, a Tom40 protein was identified using an HMM-based search that migrated in ~200 kDa complex, and more recently, GiMOMP35, a mitosomal outer membrane protein, was observed to be enriched when Tom40 was pulled down (Dagley et al. 2009; Martincová et al. 2015). However, it is not known whether GiMOMP35 is present in the GiTOM complex or if it plays any role in the mitosomal protein import. Studies over the

years have shown that no two major eukaryotic groups share the same set of TOM receptors indicating that these proteins were gained after the separation of the groups.

Tom40 homologues were identified in the genome sequences of *Microsporidia* and *Cryptosporidium* species (Abrahamsen et al. 2004; Xu et al. 2004; Heinz and Lithgow 2013; Abrahamsen et al. 2004; Xu et al. 2004). Microsporidian *Nosema bombycis* Tom40 was shown to be targeted to mitochondria when expressed in *S. cerevisiae* (Lin et al. 2012). Moreover, *Microsporidia* contain clear homologue of Tom70 receptor (Waller et al. 2009).

6.1.2 Sorting and Assembling β-Barrel Proteins: The SAM Complex

Upon entering mitochondria through the TOM channel, precursors to β-barrel proteins such as Tom40, porin and Mdm10 are directed by the small TIM chaperone complexes to the SAM pathway for correct sorting and insertion into the outer membrane (Bohnert et al. 2007). It was deduced that the β-hairpin present in the β-barrel proteins act as the targeting signal that is recognized by Tom20 and partly by Tom70 (Jores et al. 2016). Recently, an in-depth crosslinking study dissected some of the crucial steps of β-barrel assembly. The β-signal at the carboxy-terminal of the precursor initiates an opening of the Sam50 between the 1st and the 16th strands, and the β-barrel precursor is assembled in the lumen of Sam50, perhaps using Sam50 itself as the template, and further the newly folded β-barrel protein is released laterally into the membrane (Höhr et al. 2018). In yeast, SAM is composed of Sam50, Sam35, Sam37 and under certain conditions, Mdm10. Out of those, only Sam50 (Kozjak et al. 2003; Paschen et al. 2003; Gentle et al. 2004) and Sam35 (Milenkovic et al. 2004; Waizenegger et al. 2004) are essential for cell viability, whereas Mdm10 (Meisinger et al. 2004) and Sam37 (Wiedemann et al. 2003), though involved in β-barrel biogenesis, are not essential components. In yeast, the TOM complex forms a labile supercomplex with SAM for the efficient translocation and assembly of the OM proteins, and Tom22-Sam37 interaction seems to play a crucial role in the formation of the supercomplex (Qiu et al. 2013; Wenz et al. 2015). More players such as Mim1 and Mdm12/Mmm1 have been characterized that act downstream of the core SAM complex (Ishikawa et al. 2004; Waizenegger et al. 2005; Meisinger et al. 2007). Some components appear to be important only for the biogenesis of subcategories of β-barrels, such that further specific pathways may be uncovered in the near future.

The insertion of β-barrel precursors is one of the two translocation processes, besides the sorting of inner membrane and IMS proteins, which are clearly derived from a eubacterial translocation system. β-Barrel proteins are exclusively found in the outer membranes of Gram-negative bacteria and in the endosymbiotic organelles such as mitochondria and plastids (Wimley 2003). The discovery that Sam50, a protein of eubacterial ancestry, played a critical role in the insertion of mitochondrial β-barrel proteins allowed several parallels to be drawn between the eubacterial and mitochondrial β-barrel biogenesis pathways (Paschen et al. 2005; Dolezal et al.

2006). Sam50 is itself a β-barrel protein that is homologous to the β-barrel bacterial protein Omp85, which is found in all bacteria that have an outer membrane. Omp85 is essential for bacterial viability and has been shown to be involved in the insertion of β-barrel protein precursors into the outer membrane of *Neisseria* (Voulhoux et al. 2003). Phylogenetic analyses revealed that the *sam50* gene is widely distributed among eukaryotes and probably derived from an α-proteobacterial-like bacterium, possibly the mitochondrial endosymbiont (Gentle et al. 2004). Another parallel crops up between the small TIM chaperone complexes and the chaperones Skp and SurA that assist β-barrel precursors as they navigate through the bacterial periplasmic space. In effect, the mitochondrial IMS represents the periplasmic space of the mitochondrial endosymbiont. Though the two chaperone systems are phylogenetically unrelated, they presumably function to prevent aggregation of the substrates according to similar principles (Paschen et al. 2005; Dolezal et al. 2006).

In *T. vaginalis* hydrogenosomes, Sam50 was found to form a stable association with the TOM complex (Makki et al. 2019). With the help of HMM analyses, homologous sequences to Sam50 have been found in the genomes of virtually all eukaryotes with complete genome sequences except *Giardia* and related metamonds (Leger et al. 2017). These putative translocases all have features common to mitochondrial Sam50 and possibly share a common ancestor though no phylogenetic analyses have been performed on the more recently discovered sequences (Dolezal et al. 2006). The distribution of other components of the SAM complex has not yet been thoroughly investigated among mitochondrial eukaryotes, but some components are limited to fungi. No convincing homologues to Sam35, Sam37, Mdm10, Mdm12, Mmm1 or Mim1 have been found by BLAST searches of any of the complete genomes of hydrogenosomal or mitosomal species.

6.2 The Intermembrane Space Chaperones

The small TIM chaperones have been shown to convey "complicated" substrates like β-barrel and polytopic hydrophobic membrane protein precursors across the hydrophilic environment of the IMS (Koehler 2004; Bohnert et al. 2007). The small Tims—Tim8, Tim9, Tim10, Tim12 and Tim13—are about 10 kDa in size and are characterized by a C-x_3-C-$x_{11\text{-}16}$-C-x_3-C motif. Of the small Tims, only Tim9, Tim10 and Tim12 are essential, and Tim8 and Tim13 appear to be dispensable for yeast (Koehler 2004). However, both Tim9-Tim10 and Tim8-Tim13 complexes can bind AAC or β-barrel precursors (Gentle et al. 2007). Recent examination of the distribution of these small proteins using HMM analyses revealed that the small Tims have no prokaryotic homologues and may be eukaryotic inventions devised to assist membrane protein import. One or more small Tim proteins are diversely distributed among eukaryotes, but the only occurrence of small Tim-like homologues in the hydrogenosomal or mitosomal species occurs in *T. vaginalis* and *C. parvum* (Rada et al. 2011; Alcock et al. 2012). The occurrence of the *small tim* genes in diverse eukaryotes suggests an early origin for these genes (Gentle et al.

2007). It is conceivable that some other hydrogenosomal and mitosomal species developed or acquired independent chaperones. As we saw earlier, there are other types of proteins such as Skp and SurA that undertake similar chaperoning functions in bacteria. The dependence of the hydrogenosomal inner membrane protein Hmp31 on the small TIM chaperone complex, particularly Tim10, when imported into mitochondria suggests that this precursor is likely to be sensitive to the IMS, just like its mitochondrial counterparts (Dyall et al. 2000). In support of this, two paralogues of Tim9-Tim10 (A and B) were identified in the proteome of *T. vaginalis* hydrogenosomes (Rada et al. 2011).

6.3 The Inner Membrane

The two complexes that import cytosolic proteins through the mitochondrial inner membrane, TIM22 and TIM23, split the import pathways of hydrophobic inner membrane proteins from that of presequence-containing preproteins (Fig. 1).

6.3.1 The TIM22 Complex

The Tim9-Tim10 chaperone complex delivers MCF proteins such as AAC from the TOM to the TIM22 complex. Tim12, which is peripherally associated with TIM22, acts as a docking receptor for the chaperone complex. The TIM22 complex contains twin pores built from Tim22 that form a voltage-activated channel that is sensitive to synthetic peptides bearing AAC internal targeting signals, but insensitive to synthetic N-terminal presequences. The passage of the substrate through the channel is entirely dependent on the membrane potential and not on ATP hydrolysis (Kovermann et al. 2002; Rehling et al. 2003; Koehler 2004). Both Tim18 and Tim54 do not seem to mediate protein import but act as a scaffold for the TIM22 complex (Koehler et al. 2000; Hwang et al. 2007). However, like Tim22 (Sirrenberg et al. 1996) and Tim12 (Jarosch et al. 1997), Tim54 is essential (Kerscher et al. 1997), whereas Tim18 is not (Kerscher et al. 2000; Koehler et al. 2000). Recently, a metazoan-specific Tim29 was reported that is required for the stability of the TIM22 complex and for forming a contact site with TOM complex for the efficient transfer of hydrophobic proteins in the aqueous intermembrane space (Kang et al. 2016).

All components of the yeast TIM22 complex appear to be restricted to fungi, except for Tim22 which is widely distributed among eukaryotes (Rassow et al. 1999; Dolezal et al. 2006). Among the mitosomal and hydrogenosomal species, only a putative Tim22 sequence was reported in the genome of *E. cuniculi* and other *Microsporidia* (Katinka et al. 2001; Žárský and Doležal 2016; Pyrihová et al. 2018).

6.3.2 The TIM23 Complex

The TIM23 complex, which intakes presequence-containing precursors, is better characterized than the TIM22 complex and exhibits the most intricate import mechanisms (Bohnert et al. 2007). Within the complex, Tim23 forms a cation-selective, voltage-gated, protein-conducting, possibly twin-pore, channel that is specifically sensitive to synthetic presequence peptides (Truscott et al. 2001; Martinez-Caballero et al. 2007). Tim17, though homologous in sequence and a secondary structure to Tim23, does not form part of the channel but modulates its activity (Meier et al. 2005; Martinez-Caballero et al. 2007). Tim21 makes direct contact with the TOM complex by interacting with the IMS domain of Tom22, where it promotes precursor release by competing with presequence binding (Chacinska et al. 2005). Tim50 has a dual role, i.e. acting as a receptor for presequences and regulating the closure of the TIM23 channel (Geissler et al. 2002; Yamamoto et al. 2002; Meinecke et al. 2006). Tim21 also regulates the interaction between PAM and TIM23 by associating with TIM23. This complicated interaction serves to generate two types of TIM23 complexes: one that is matrix-import competent and the other that is competent to sort and insert the presequence-carrying inner membrane proteins (Chacinska et al. 2005; van der Laan et al. 2006a, b). All components of TIM23, except for Tim21 (Chacinska et al. 2005), are essential (Dekker et al. 1993; Emtage and Jensen 1993; Ryan et al. 1994; Maarse et al. 1994; Geissler et al. 2002; Yamamoto et al. 2002; Mokranjac et al. 2003a). This is quite surprising, given the central role played by Tim21 at various levels. Again, this attests to the flexibility of the mitochondrial protein import machinery.

Genes homologous to *tim23* and *tim17* have been found in most mitochondrial eukaryotes (Rassow et al. 1999; Dolezal et al. 2006), and *tim21* homologues can be found in animal, plant and fungal genomes (Chacinska et al. 2005) but not in protists (our unpublished observations). Tim50 contains a LIM domain commonly found in proteins of diverse functions, and no profound sequence analyses have yet been performed to assess its distribution among various species. Sequences related to *tim17* and *tim23* have been detected in the respective genomes of *T. vaginalis* and *C. parvum* (Abrahamsen et al. 2004; Henriquez et al. 2005). Four paralogues of Tim17/22/23 family (A-D), a Tim17-like protein and Tim44 were reported to be present in the hydrogenosomes of *T. vaginalis* (Rada et al. 2011). Thus, a core TIM23 complex could exist in the organelles of these two organisms.

Recently, single Tim17 family protein has been identified in the genome of *G. intestinalis* and localized to its mitosomes (Pyrihová et al. 2018). The protein forms disulphide bond-mediated dimers in the inner mitosomal membrane, where it seems to interact with Tim44 and other components of mitosomal protein import machinery. No clear homologues to any TIM23 component could be detected in the complete genome sequences of *E. histolytica*.

7 The Protein Import Motor

The final step of the journey of the mitochondrial matrix-targeted preprotein across the membranes involves the participation of an ATP-driven protein import motor, PAM, which pulls the preprotein from the Tim23 channel into the matrix (Fig. 1). In yeast, the core component of the PAM complex is Ssc1, or mt-Hsp70, which is assisted by Mge1, Pam18, Pam16, Pam17 and Tim44 (van der Laan et al. 2006a; Bohnert et al. 2007). With the exception of Pam17 (van der Laan et al. 2005), all PAM components identified so far are essential for yeast viability (Craig et al. 1987; Maarse et al. 1992; Bolliger et al. 1994; Mokranjac et al. 2003b; D'Silva et al. 2003; Truscott et al. 2003; Frazier et al. 2004; Kozany et al. 2004; Li et al. 2004).

Mt-Hsp70 is a member of the Hsp70 chaperone family that is distributed in all domains of life. The bacterial cytoplasmic homologues are called DnaK, and various types of *hsp70* genes are found in eukaryotes, with the products localizing to the cytosol, the ER, the mitochondrion or a plastid compartment (Gupta and Singh 1994; Bukau and Horwich 1998). Phylogenetic analyses show strong affinity and conserved signature sequences between mt-Hsp70 and α-proteobacterial DnaK, supporting the endosymbiotic origin of mitochondria from an α-proteobacterial-like ancestor (Boorstein et al. 1994; Falah and Gupta 1994; Gupta 2006). The Hsp70 proteins are the central part of protein folding machines that are powered by ATP. Generally, Hsp70 molecules have a highly conserved amino-terminal region with an ATPase domain and a carboxy-terminal region with a peptide-binding domain. The extensively studied chaperone system in *E. coli* revealed much about the mechanism of action of DnaK, which is assisted in its function by the nucleotide exchange factor GrpE, and the J-protein DnaJ that enhances ATPase activity (Bukau and Horwich 1998). A similar system can be found operating with a likewise mechanism at the matrix side of the TIM23 complex. In this situation, however, mt-Hsp70 is not involved in protein folding per se, but its properties are put to use to bind a largely unfolded incoming preprotein and to drive it completely into the mitochondrial matrix in an action regulated by ATP hydrolysis and co-chaperones. A fraction of mt-Hsp70 docks onto the TIM23 complex through the essential peripheral membrane protein Tim44 (Voos and Röttgers 2002; van der Laan et al. 2006a). Genes homologous to *tim44* have been found in all the completed genome sequences of mitochondrial eukaryotes, and also of α-proteobacteria, where the putative functions of the homologues are unknown (Dolezal et al. 2006; Clements et al. 2009). As the freshly imported preprotein enters the mitochondrial matrix, it is bound and/or pulled in by mt-Hsp70, which is assisted by the soluble matrix protein Mge1 (Bolliger et al. 1994; Voos and Röttgers 2002) and the inner membrane protein Pam18 (Mokranjac et al. 2003b; D'Silva et al. 2003; Truscott et al. 2003). These are the respective homologues of bacterial GrpE and DnaJ. Pam18 has a matrix-oriented J-domain with which it can stimulate mt-Hsp70 ATPase activity (Truscott et al. 2003). Pam18 is tightly bound to Pam16, which contains a degenerate J-domain and acts as a regulatory peripheral inner membrane protein within the motor (Frazier et al. 2004; Kozany et al. 2004; Li et al. 2004). The role of the final

and non-essential component of PAM, Pam17, is unclear, but it is necessary for the stable modular association of Pam16 and Pam18 with TIM23 (van der Laan et al. 2005). This particular component appears to be fungi-specific as convincing homologues could not be found in other mitochondrial species or in eubacteria.

Mitochondrial-type Hsp70 is the only PAM component that has homologues in all mitosome- or hydrogenosome-containing species examined to date, namely, *T. vaginalis* (Bui et al. 1996; Germot et al. 1996), *G. intestinalis* (Morrison et al. 2001; Arisue et al. 2002), *E. histolytica* (Arisue et al. 2002), *C. parvum* (Slapeta and Keithly 2004), *E. cuniculi* (Peyretaillade et al. 1998), *A. locustae* (Germot et al. 1997), *T. hominis* (Williams et al. 2002), *N. ovalis* (Boxma et al. 2005) and *N. patriciarum* (van der Giezen et al. 2003). The complete set of PAM components (mt-Hsp70, Tim44, Pam18 and Pam16) has been identified in *T. vaginalis* hydrogenosomes (Rada et al. 2011; Schneider et al. 2011) and *G. intestinalis* mitosomes (Dolezal et al. 2005; Rada et al. 2011; Martincová et al. 2015) suggesting that a PAM system functions in the organelle. Mt-Hsp70, Pam18 and Tim44 have been found in the genome sequence of *C. parvum* (Abrahamsen et al. 2004), but only mt-Hsp70 has been localized to its mitosome so far (Slapeta and Keithly 2004). In *Microsporidia*, only genes for mt-Hsp70 and Pam18 were identified (Katinka et al. 2001; Waller et al. 2009). However, the most derived and reduced motor complex seems to exist in *E. histolytica* where only homologue of mt-Hsp70 was found (Arisue et al. 2002). The putative origins of most of the mitosomal and hydrogenosomal mt-Hsp70 homologues have been thoroughly pursued through phylogenetic analyses where most sequences group with the mitochondrial homologues with fairly strong support, except in the case of *G. intestinalis* mt-Hsp70, which is divergent (Morrison et al. 2001; Arisue et al. 2002). Generally, it is assumed that the mt-Hsp70 homologues originate from the α-proteobacterial-like endosymbiont that gave rise to the mitochondrion.

8 Preprotein Processing Peptidases

Upon translocation into the matrix, the N-terminal presequence of preproteins is processed by MPP, and the mature protein is thereafter folded into its native conformation (Fig. 1). Some preproteins have a bipartite presequence that is processed in two steps, the first part by MPP and the second part, which includes an octapeptide motif, by the mitochondrial intermediate peptidase (MIP). Precursors destined for the IMS have an IMS-sorting signal at the N-terminus. The IMP complex is responsible for the maturation of these proteins. Some of the precursors contain bipartite presequences consisting of a matrix-targeting signal followed by an intermembrane space-sorting signal (Gakh et al. 2002).

8.1 The Mitochondrial Processing Peptidase (MPP)

The mitochondrial processing peptidase is an essential zinc-dependent metallopeptidase (Yaffe et al. 1985; Luciano and Géli 1996; Gakh et al. 2002; Nomura et al. 2006). It cleaves the N-terminal presequence from precursors to matrix-targeted proteins and from precursors destined for the inner membrane or the IMS. Through sequence comparisons, Gavel and von Heijne (1990) defined four cleavage site motifs for MPP and MIP:

(a) The R-2 motif: x-R-x^x-(S/x)
(b) The R-3 motif: x-R-x-(Y/x)^(S/A/x)-x
(c) The R-10 motif: x-R-x^(F/L/I)-x-x-(S/T/G)-xxxx^, where the second cleavage site is for MIP
(d) The R-none motif: x-x^x-(S/x)

Surveys of mitochondrial presequences showed that, though quite common, these above motifs are not found in all of them and that the primary sequence for the cleavage site is quite degenerate.

The role of the Arg at the -2 or -3 position is unclear and may be presequence-dependent as studies on a variety of precursors revealed that mutating the Arg results in cleavage inhibition or modification in some cases, but not in others. It may be that the structure, rather than the primary sequence composition of the presequence and perhaps of the mature protein, determines the MPP cleavage site (Gakh et al. 2002).

Generally, the enzyme consists of two core subunits, α-MPP and β-MPP, each of about 50 kDa in size, which are widely distributed among mitochondrial eukaryotes. α-MPP and β-MPP are homologous to each other with up to about 30% identical residues in some species. The catalytic unit is β-MPP, which contains the conserved and critical zinc-binding motif H-x-x-H-x_{76}-E. This motif is characteristic of the pitrilysin protease family that includes bacterial proteases (Rawlings and Barrett 1995). The α-subunit is not involved in processing but may be involved in substrate recognition and interaction through a highly conserved glycine-rich loop. However, both subunits are required for processing the presequence in mitochondria (Geli et al. 1990).

MPP has long been thought to have evolved from a bacterial protease of the pitrilysin family (Gakh et al. 2002). Recently, a putative peptidase has been characterized from the α-proteobacterial parasitic bacterium *Rickettsia prowazekii* and related species and was found to have domains typical of both subunits of MPP (Kitada et al. 2007). Strikingly, the N-terminal domain of this rickettsial putative peptidase (RPP) resembles the N-terminal region β-MPP with an H-x-x-H-x_{76}-E motif, and the C-terminal domain of RPP resembles the C-terminal region of α-MPP, minus the glycine-rich loop. Unlike β-MPP, monomeric recombinant RPP was shown to have proteolytic activity on its own, cleaving basic synthetic peptides preferentially. RPP was able to cleave mitochondrial presequence peptides at specific sites in some cases, albeit at reduced efficiency compared with MPP. However, when tested on mitochondrial preproteins with short and long presequences,

respectively, RPP was inactive on its own. Processing of the short presequence only occurred when RPP was stoichiometrically mixed with yeast β-MPP, and it was demonstrated that β-MPP was involved in the catalytic activity and not RPP. Thus, RPP behaved like α-MPP as an activator of β-MPP. The long presequence was not processed by either RPP/β-MPP or RPP/α-MPP, and mutational studies on MPP indicated that this could be due to the absence of a glycine-rich loop on RPP (Kitada et al. 2007). Given the close relationship between mitochondria and *Rickettsia* (Andersson et al. 1998), these findings indicate that RPP may represent an ancestral form of both α-MPP and β-MPP, derived from the α-proteobacterial-like mitochondrial endosymbiont.

Homologues to β-MPP, both with conserved catalytic motifs, were discovered recently in the genomes of *T. vaginalis* and *G. intestinalis*. The *Giardia* β-MPP homologue was localized to mitosomes, and N-terminal sequencing of mitosomal IscU confirmed the cleavage site of its presequence at the position suggested by the PSORT prediction programme (Table 1) (Dolezal et al. 2005). The biochemical characterization of *Giardia* β-MPP subunit has showed that the protein functions as a monomer without the assistance of the a-subunit (Šmíd et al. 2008). The proteins seem to have co-evolved with the shorter mitosomal presequences and are not able to process presequences on hydrogenosomal or mitochondrial precursors (Šmíd et al. 2008). Similarly in *C. parvum*, a homologue to only β-MPP, but not α-MPP, has been reported (Abrahamsen et al. 2004; Henriquez et al. 2005).

In contrast, the hydrogenosomes of *T. vaginalis* contain typical dimeric MPP (Brown et al. 2007; Šmíd et al. 2008), which exhibits broader specificity as demonstrated on the efficient processing of presequences derived from mitosomal (*Giardia*) and mitochondrial (*S. cerevisiae*) precursor proteins (Šmíd et al. 2008). In *E. histolytica*, one presequence has been shown to be cleaved at a site predicted for MPP (Mai et al. 1999; Tovar et al. 1999), but no enzyme responsible for the processing has been identified yet. Similarly, no MPP homologue was found in *E. cuniculi* and other *Microsporidia*, but given the occurrence of presequence-independent protein import in microsporidia (Burri et al. 2006), they may have dispensed with processing peptidases during their reductive evolution. In ciliate *N. ovalis*, both MPP subunits have been identified suggesting conserved processing of precursor proteins in its hydrogenosomes (Boxma et al. 2005). In general, there have been no reports of MIP-like proteins nor of any R-10 motif on protein precursors in any of the hydrogenosomal or mitosomal species.

8.2 The Inner Membrane Protease

Anchored on the outer face of the inner membrane, the mitochondrial IMP complex consists of two proteases Imp1 and Imp2 and a regulatory subunit Som1 (Fig. 1). The two proteases have distinct specificities for IMS protein precursors. Some of the precursors contain bipartite presequences consisting of a matrix-targeting signal followed by an intermembrane space-sorting signal for sequential cleavage by

MPP and IMP. Imp1 and Imp2 show significant similarity to bacterial type I leader peptidases that cleave the N-terminal signal of precursors that traverse the bacterial membrane (Gakh et al. 2002).

The import route of mtG3PDH into microsporidian mitosomes seems to follow the stop-transfer pathway in *S. cerevisiae*, during which the translocation of mtG3PDH is stopped at the TIM23 complex, where the precursor remains in the membrane without release into the matrix (Esser et al. 2004). However, the processing step is different in *A. locustae* and *E. cuniculi* preproteins. In *A. locustae*, as in *S. cerevisiae,* the precursor seems to be processed by IMP that cleaves off the presequence at the position following the first transmembrane segment (Esser et al. 2004; Burri et al. 2006). In contrast, in *E. cuniculi*, the N-terminal domain is retained within the mature protein. *S. cerevisiae* IMP could process the *A. locustae* mtG3PDH precursor, and an IMP2 homologue is present in the *A. locustae* genome. Together, these data suggest that *A. locustae* has retained an IMP proteolytic processing pathway but that the related microsporidian species *E. cuniculi* may have discarded both MPP and IMP processing (Burri et al. 2006; Burri and Keeling 2007). Currently, there is no evidence for IMP processing in any of the other mitosomal or hydrogenosomal species.

9 Folding Newly Imported Soluble Proteins

Newly imported proteins enter mitochondria in an extended or only partly folded conformation. Two main chaperone systems have been characterized in mitochondria that fold these incoming proteins into a native state that permits them to perform their function. Mitochondria have inherited these efficient and intricate folding systems from their bacterial progenitor(s): one involving mt-Hsp70 and the other with Cpn60/Cpn10 or Hsp60/Hsp10 (Neupert 1997; Voos and Röttgers 2002).

Besides its role in preprotein translocation across the inner membrane through TIM23 and PAM, mt-Hsp70 can also act as a protein folding chaperone. Indeed, mt-Hsp70 in yeast mitochondria is either found in a membrane-associated complex with Tim44 and PAM or in a soluble state in association with co-chaperones Mdj1 and Mge1. Mdj1 is a highly conserved non-essential mitochondrial homologue of bacterial DnaJ and was shown not to be involved in translocation but to be important for protein folding in association with the homologues of GrpE and DnaK (Neupert 1997; Voos and Röttgers 2002). The manner in which the mt-Hsp70 chaperone functions is very similar to that of bacterial DnaK, and the system is likely to have been inherited from the bacterial progenitor of mitochondria (Hartl et al. 1994; Stuart et al. 1994; Szabo et al. 1994). As we have reported in Sect. 7, homologues to mt-Hsp70/DnaK have been found in the genomes of all mitosomal or hydrogenosomal species examined to date, and homologues to Mge1/GrpE have been found in *T. vaginalis* (Carlton et al. 2007), *G. intestinalis* (Martincová et al. 2015) and *C. parvum* (Abrahamsen et al. 2004). Homologues to Mdj1/DnaJ have been reported in *T. vaginalis*, *E. cuniculi* and *N. ovalis* (Katinka et al. 2001; Boxma

et al. 2005; Carlton et al. 2007). All the components of the DnaK-type machinery have been localized to *T. vaginalis* hydrogenosomes (Bozner 1997; Dyall et al. 2003; Rada et al. 2011; Schneider et al. 2011), suggesting that a similar protein folding mechanism occurs in these organelles.

The mitochondrial Cpn60/Cpn10 or Hsp60/Hsp10 chaperone system participates in the folding of the majority of newly imported matrix proteins (Neupert 1997; Voos and Röttgers 2002). This system functions downstream of the mt-Hsp70 system, but both systems are likely to cooperate in protein folding (Manning-Krieg et al. 1991). Cpn60 and Cpn10 derive from bacterial homologues GroEL and GroES, respectively, and phylogenetic and comparative analyses of both protein sequences show a robust relationship between the respective monophyletic mitochondrial groups and α-proteobacteria. Since the progenitor of mitochondria is likely to have been an ancestor of extant α-proteobacteria, these findings support the notion that Cpn60 and Cpn10 originate from the endosymbiont that gave rise to mitochondria (Gupta 2018). In bacteria, including α-proteobacteria, *groel* and *groes* genes are found on a single operon, such that the eukaryotic genes are likely to have a common origin (Gupta 2018). Much has been learnt about the mechanism of protein folding in bacteria through the structure of the bacterial GroEL/GroES complex. In *E. coli*, the GroEL proteins form a double-ring structure comprising two apposed heptameric rings that form a central cavity that binds protein folding intermediates of up to 50 kDa and facilitates folding to the native state. The chaperonin cavity switches from a binding to a folding state through conformational changes induced by ATP. This action is regulated by a saucer-shaped heptameric complex of GroES which modulates both the ATPase cycle and the conformation of GroEL monomers (Rye et al. 1997; Xu et al. 1997; Bukau and Horwich 1998). Both Cpn60 and Cpn10 are encoded by essential genes in yeast and are likely to function similarly to their bacterial homologues (Cheng et al. 1989; Rospert et al. 1993), but not all mitochondrial proteins require Cpn60 for folding (Rospert et al. 1996). Homologues to Cpn60 that show high affinity to mitochondrial Cpn60 have been found in *E. histolytica* (Clark and Roger 1995), *T. vaginalis* (Horner et al. 1996; Bui et al. 1996; Roger et al. 1996), *G. intestinalis* (Roger et al. 1998), *C. parvum* (Riordan et al. 2003; Putignani et al. 2004), *N. patriciarum* (van der Giezen et al. 2003) and *N. ovalis* (Boxma et al. 2005). These putative chaperones have been localized to either hydrogenosomes or mitosomes in *T. vaginalis* (Bui et al. 1996; Bozner 1997), *E. histolytica* (Mai et al. 1999; Tovar et al. 1999) , *G. intestinalis* (Regoes et al. 2005) (Regoes et al. 2005), *C. parvum* (Riordan et al. 2003; Putignani et al. 2004) and *N. patriciarum* (van der Giezen et al. 2003). So far, homologues to Cpn10 have been reported in the genomes of *T. vaginalis, G. intestinalis*, *E. histolytica and Cryptosporidium* species, but phylogenetic relationships with either mitochondrial or α-proteobacterial sequences could not be convincingly inferred (Bui et al. 1996; van der Giezen et al. 2005). Surprisingly, no homologue to either *cpn60* or *cpn10* was found in the genomes of microsporidia available to date. It may be that microsporidian mitosomal proteins do not require Cpn60/Cpn10 for protein folding, as has been noted for a subset of mitochondrial matrix proteins (Rospert et al. 1996). It is plausible that, for the sake of economy, the highly reduced microsporidian mitosomes utilize homologues to the

multifunctional mt-Hsp70 protein to both translocate and fold newly imported proteins and have dispensed with the energetically expensive Cpn60 machinery.

10 Perspectives

It is evident that protein import mechanisms are conserved between hydrogenosomes, mitosomes and mitochondria. Although no protein import pathway has been functionally deciphered for hydrogenosomes and mitosomes, we have started to get a glimpse of some putative mitochondrial-like components that may be involved in importing, processing and folding preproteins during biogenesis. The species that have been shown to harbour either mitosomes or hydrogenosomes and have had their genomes completely sequenced offer us an opportunity to examine their putative mitochondrial protein import complement. Mitosomes of *C. parvum* and hydrogenosomes of *T. vaginalis* potentially house the most mitochondrial-like components, though many of them have not as yet been localized. Their organelles could have mitochondrial-like SAM, TIM23 and PAM machines and mitochondrial-like preprotein processing and folding. On the contrary, the mitosomes of *E. histolytica* and *G. intestinalis* lack Tim17 family protein and Sam50, respectively, which are the key components present in all other eukaryotes. Does it mean *that G. intestinalis* develop an alternate mechanism to assemble β-barrel proteins in the outer membrane? Or did their β-barrel proteins evolve such way that they do not require SAM complex anymore? Did *E. histolytica* build the mitochondrial inner membrane translocase around different channel subunits?

We seem to be reaching the limits of how much we can assimilate and conclude from genome sequence analyses. These have been invaluable in identifying some putative protein translocases and chaperones. More sensitive searches like HMM may indeed deliver further putative candidates for mitochondrial-type translocases from the genome sequence databases of the hydrogenosomal and mitosomal species. However, we shall need to go back to the bench to demonstrate their localization and investigate their involvement in organellar protein trafficking. In addition to the above-mentioned unknowns, more general questions appear in the light of newly discovered protein transport and biogenesis pathways in model organisms. Specifically, what is the relationship of the mitochondria-related organelles to the endoplasmic reticulum and other endomembrane system organelles in terms of their biogenesis and dynamics? How do these protists control the number and the metabolic capacity of their organelles? Some of these questions can be answered with the limited set of tools we have at hand, but it is imperative to develop new techniques if we want to dissect these pathways.

Once these questions are answered, we shall be in a better position to formulate hypotheses on how these fantastic protein transport machines have evolved. By comparing protein import mechanisms and examining the structure of translocases between hydrogenosomal, mitosomal, mitochondrial, plastidic and bacterial systems, it is likely that we discover common principles for protein targeting. We can

ask further and broader impact questions. For instance, how do the intricacy of the protein targeting machines correlates with proteome size? How do targeting signals and translocases co-evolve? How hard is it for an endosymbiont to build a protein import machine?

References

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

Abrahamsen MS, Templeton TJ, Enomoto S, Abrahante JE, Zhu G, Lancto CA, Deng M, Liu C, Widmer G, Tzipori S, Buck GA, Xu P, Bankier AT, Dear PH, Konfortov BA, Spriggs HF, Iyer L, Anantharaman V, Aravind L, Kapur V (2004) Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. Science 304:441–445

Adams KL, Palmer JD (2003) Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Mol Phylogenet Evol 29:380–395

Adams KL, Daley DO, Qiu Y-L, Whelan J, Palmer JD (2000) Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. Nature 408:354

Alcock F, Webb CT, Dolezal P, Hewitt V, Shingu-Vasquez M, Likić VA, Traven A, Lithgow T (2012) A small Tim homohexamer in the relict mitochondrion of *Cryptosporidium*. Mol Biol Evol 29:113–122

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Alva V, Nam S-Z, Söding J, Lupas AN (2016) The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res 44:W410–W415

Andersson SGE, Zomorodipour A, Andersson JO, Sicheritz-Pontén T, Alsmark UCM, Podowski RM, Näslund AK, Eriksson A-S, Winkler HH, Kurland CG (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396:133

Andersson GE, Karlberg O, Canbäck B, Kurland CG (2003) On the origin of mitochondria: a genomics perspective. Philos Trans R Soc London Ser B Biol Sci 358:165–179

Arisue N, Sánchez LB, Weiss LM, Müller M, Hashimoto T (2002) Mitochondrial-type hsp70 genes of the amitochondriate protists, *Giardia intestinalis*, *Entamoeba histolytica* and two microsporidians. Parasitol Int 51:9–16

Baker A, Schatz G (1987) Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. Proc Natl Acad Sci U S A 84:3117–3121

Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer ELL, Studholme DJ, Yeats C, Eddy SR (2004) The Pfam protein families database. Nucleic Acids Res 32:D138–D141

Bausewein T, Mills DJ, Langer JD, Nitschke B, Nussberger S, Kühlbrandt W (2017) Cryo-EM Structure of the TOM Core Complex from *Neurospora crassa*. Cell 170:693–700.e7

Beasley EM, Müller S, Schatz G (1993) The signal that sorts yeast cytochrome b2 to the mitochondrial intermembrane space contains three distinct functional regions. EMBO J 12:2303–2311

Becker T, Pfannschmidt S, Guiard B, Stojanovski D, Milenkovic D, Kutik S, Pfanner N, Meisinger C, Wiedemann N (2008) Biogenesis of the mitochondrial TOM complex: Mim1 promotes insertion and assembly of signal-anchored receptors. J Biol Chem 283:120–127

Blobel G, Dobberstein B (1975a) Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J Cell Biol 67:835–851

Blobel G, Dobberstein B (1975b) Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. J Cell Biol 67:852–862

Bohnert M, Pfanner N, van der Laan M (2007) A dynamic machinery for import of mitochondrial precursor proteins. FEBS Lett 581:2802–2810

Bolliger L, Deloche O, Glick BS, Georgopoulos C, Jeno P, Kronidou N, Horst M, Morishima N, Schatz G (1994) A mitochondrial homolog of bacterial GrpE interacts with mitochondrial hsp70 and is essential for viability. EMBO J 13:1998–2006

Bolliger L, Junne T, Schatz G, Lithgow T (1995) Acidic receptor domains on both sides of the outer membrane mediate translocation of precursor proteins into yeast mitochondria. EMBO J 14:6318–6326

Bömer U, Meijer M, Guiard B, Dietmeier K, Pfanner N, Rassow J (1997) The sorting route of cytochrome b branches from the general mitochondrial import pathway at the preprotein translocase of the inner membrane. J Biol Chem 272:30439–30446

Bonnefoy N, Remacle C, Fox TD (2007) Genetic transformation of *Saccharomyces cerevisiae* and *Chlamydomonas reinhardtii* mitochondria. Methods Cell Biol 80:525–548

Boorstein WR, Ziegelhoffer T, Craig EA (1994) Molecular evolution of the HSP70 multigene family. J Mol Evol 38:1–17

Boxma B, de Graaf RM, van der Staay GWM, van Alen TA, Ricard G, Gabaldón T, van Hoek AHAM, Moon-van der Staay SY, Koopman WJH, van Hellemond JJ, Tielens AGM, Friedrich T, Veenhuis M, Huynen MA, Hackstein JHP (2005) An anaerobic mitochondrion that produces hydrogen. Nature 434:74

Bozner P (1997) Immunological detection and subcellular localization of Hsp70 and Hsp60 homologs in *Trichomonas vaginalis*. J Parasitol 83:224–229

Bracha R, Nuchamowitz Y, Mirelman D (2003) Transcriptional silencing of an amoebapore gene in *Entamoeba histolytica*: molecular analysis and effect on pathogenicity. Eukaryot Cell 2:295–305

Bradley PJ, Lahti CJ, Plumper E, Johnson PJ (1997) Targeting and translocation of proteins into the hydrogenosome of the protist Trichomonas: similarities with mitochondrial protein import. EMBO J 16:3484–3493

Brás XP, Zimorski V, Bolte K, Maier U-G, Martin WF, Gould SB (2013) Knockout of the abundant *Trichomonas vaginalis* hydrogenosomal membrane protein *Tv* HMP23 increases hydrogenosome size but induces no compensatory up-regulation of paralogous copies. FEBS Lett 587:1333–1339

Brix J, Dietmeier K, Pfanner N (1997) Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. J Biol Chem 272:20730–20735

Brix J, Ziegler GA, Dietmeier K, Schneider-Mergener J, Schulz GE, Pfanner N (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

Brondijk TH, Durand R, van der Giezen M, Gottschal JC, Prins RA, Fevre M (1996) scsB, a cDNA encoding the hydrogenosomal beta subunit of succinyl-CoA synthetase from the anaerobic fungus *Neocallimastix frontalis*. Mol Gen Genet 253:315–323

Brown MT, Goldstone HMH, Bastida-Corcuera F, Delgadillo-Correa MG, McArthur AG, Johnson PJ (2007) A functionally divergent hydrogenosomal peptidase with protomitochondrial ancestry. Mol Microbiol 64:1154–1163

Bui ET, Bradley PJ, Johnson PJ (1996) A common evolutionary origin for mitochondria and hydrogenosomes. Proc Natl Acad Sci U S A 93:9651–9656

Bui ET, Johnson PJ (1996) Identification and characterization of [Fe]-hydrogenases in the hydrogenosome of *Trichomonas vaginalis*. Mol Biochem Parasitol 76:305–310

Bukau B, Horwich AL (1998) The Hsp70 and Hsp60 chaperone machines. Cell 92:351–366

Burri L, Keeling PJ (2007) Protein targeting in parasites with cryptic mitochondria. Int J Parasitol 37:265–272

Burri L, Williams BAP, 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

Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, Wortman JR, Bidwell SL, Alsmark UCM, Besteiro S, Sicheritz-Ponten T, Noel CJ, Dacks JB, Foster PG, Simillion C, Van de Peer Y, Miranda-Saavedra D, Barton GJ, Westrop GD, Müller S, Dessi D, Fiori PL, Ren Q, Paulsen I, Zhang H, Bastida-Corcuera FD, Simoes-Barbosa A, Brown MT, Hayes RD, Mukherjee M, Okumura CY, Schneider R, Smith AJ, Vanacova S, Villalvazo M, Haas BJ, Pertea M, Feldblyum TV, Utterback TR, Shu C-L, Osoegawa K, de Jong PJ, Hrdy I, Horvathova L, Zubacova Z, Dolezal P, Malik S-B, Logsdon JM, Henze K, Gupta A, Wang CC, Dunne RL, Upcroft JA, Upcroft P, White O, Salzberg SL, Tang P, Chiu C-H, Lee Y-S, Embley TM, Coombs GH, Mottram JC, Tachezy J, Fraser-Liggett CM, Johnson PJ (2007) Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. Science 315:207–212

Carpenter ML, Cande WZ (2009) Using morpholinos for gene knockdown in Giardia intestinalis. Eukaryot Cell 8:916–919

Cautain B, Hill R, de Pedro N, Link W (2015) Components and regulation of nuclear transport processes. FEBS J 282:445–462

Cavalier-Smith T (1987) The simultaneous symbiotic origin of mitochondria, chloroplasts, and microbodies. Ann N Y Acad Sci 503:55–71

Chacinska A, Lind M, Frazier AE, Dudek J, Meisinger C, Geissler A, Sickmann A, Meyer HE, Truscott KN, Guiard B, Pfanner N, Rehling P (2005) Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. Cell 120:817–829

Chan KW, Slotboom D-J, Cox S, Embley TM, Fabre O, van der Giezen M, Harding M, Horner DS, Kunji ERS, León-Avila G, Tovar J (2005) A novel ADP/ATP transporter in the mitosome of the microaerophilic human parasite Entamoeba histolytica. Curr Biol 15:737–742

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

Cheng MY, Hartl F-U, Martin J, Pollock RA, Kalousek F, Neuper W, Hallberg EM, Hallberg RL, Horwich AL (1989) Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. Nature 337:620

Choi C, Liu Z, Adams KL (2006) Evolutionary transfers of mitochondrial genes to the nucleus in the *Populus* lineage and coexpression of nuclear and mitochondrial Sdh4 genes. New Phytol 172:429–439

Clark CG, Roger AJ (1995) Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. Proc Natl Acad Sci U S A 92:6518–6521

Clements A, Bursac D, Gatsos X, Perry AJ, Civciristov S, Celik N, Likic VA, Poggio S, Jacobs-Wagner C, Strugnell RA, Lithgow T (2009) The reducible complexity of a mitochondrial molecular machine. Proc Natl Acad Sci U S A 106:15791–15795

Connolly T, Gilmore R (1986) Formation of a functional ribosome-membrane junction during translocation requires the participation of a GTP-binding protein. J Cell Biol 103:2253–2261

Craig EA, Kramer J, Kosic-Smithers J (1987) SSC1, a member of the 70-kDa heat shock protein multigene family of *Saccharomyces cerevisiae*, is essential for growth. Proc Natl Acad Sci U S A 84:4156–4160

Curran SP, Leuenberger D, Leverich EP, Hwang DK, Beverly KN, Koehler CM (2004) The role of Hot13p and redox chemistry in the mitochondrial TIM22 import pathway. J Biol Chem 279:43744–43751

D'Silva PD, Schilke B, Walter W, Andrew A, Craig EA (2003) J protein cochaperone of the mitochondrial inner membrane required for protein import into the mitochondrial matrix. Proc Natl Acad Sci U S A 100:13839–13844

Dagley MJ, Dolezal P, Likic VA, Smid O, Purcell AW, Buchanan SK, Tachezy J, Lithgow T (2009) The protein import channel in the outer mitosomal membrane of *Giardia intestinalis*. Mol Biol Evol 26:1941–1947

Dalbey RE, Lively MO, Bron S, Van Dijl JM (1997) The chemistry and enzymology of the type I signal peptidases. Protein Sci 6:1129–1138

Daley DO, Clifton R, Whelan J (2002) Intracellular gene transfer: reduced hydrophobicity facilitates gene transfer for subunit 2 of cytochrome c oxidase. Proc Natl Acad Sci U S A 99:10510–10515

Dan M, Wang AL, Wang CC (2000) Inhibition of pyruvate-ferredoxin oxidoreductase gene expression in *Giardia lamblia* by a virus-mediated hammerhead ribozyme. Mol Microbiol 36:447–456

Davis AJ, Ryan KR, Jensen RE (1998) Tim23p contains separate and distinct signals for targeting to mitochondria and insertion into the inner membrane. Mol Biol Cell 9:2577–2593

Dekker PJT, Keil P, Rassow J, Maarse AC, Pfanner N, Meijer M (1993) Identification of MIM23, a putative component of the protein import machinery of the mitochondrial inner membrane. FEBS Lett 330:66–70

Delgadillo MG, Liston DR, Niazi K, Johnson PJ (1997) Transient and selectable transformation of the parasitic protist *Trichomonas vaginalis*. Proc Natl Acad Sci U S A 94:4716–4720

Dietmeier K, Honlinger A, Bomer U, Dekker PJ, Eckerskorn C, Lottspeich F, Kubrich M, Pfanner N (1997) Tom5 functionally links mitochondrial preprotein receptors to the general import pore. Nature 388:195–200

Dolezal P, Smid O, Rada P, Zubacova Z, Bursac D, Sutak R, Nebesarova J, Lithgow T, Tachezy J (2005) *Giardia* mitosomes and *trichomonad* hydrogenosomes share a common mode of protein targeting. Proc Natl Acad Sci U S A 102:10924–10929

Dolezal P, Likic V, Tachezy J, Lithgow T (2006) Evolution of the molecular machines for protein import into mitochondria. Science 313:314–318

Dyall SD, Koehler CM, Delgadillo-Correa MG, Bradley PJ, Plümper E, Leuenberger D, Turck CW, Johnson PJ (2000) Presence of a member of the mitochondrial carrier family in hydrogenosomes: conservation of membrane-targeting pathways between hydrogenosomes and mitochondria. Mol Cell Biol 20:2488–2497

Dyall SD, Lester DC, Schneider RE, Delgadillo-Correa MG, Plümper E, Martinez A, Koehler CM, Johnson PJ (2003) *Trichomonas vaginalis* Hmp35, a putative pore-forming hydrogenosomal membrane protein, can form a complex in yeast mitochondria. J Biol Chem 278:30548–30561

Dyall SD, Yan W, Delgadillo-Correa MG, Lunceford A, Loo JA, Clarke CF, Johnson PJ (2004) Non-mitochondrial complex I proteins in a hydrogenosomal oxidoreductase complex. Nature 431:1103–1107

Ebneter JA, Heusser SD, Schraner EM, Hehl AB, Faso C (2016) Cyst-Wall-Protein-1 is fundamental for Golgi-like organelle neogenesis and cyst-wall biosynthesis in *Giardia lamblia*. Nat Commun 7:13859

Eddy SR (1998) Profile hidden Markov models. Bioinformatics 14:755–763

Emtage JL, Jensen RE (1993) MAS6 encodes an essential inner membrane component of the yeast mitochondrial protein import pathway. J Cell Biol 122:1003–1012

Esser K, Jan P-S, Pratje E, Michaelis G (2004) The mitochondrial IMP peptidase of yeast: functional analysis of domains and identification of Gut2 as a new natural substrate. Mol Gen Genomics 271:616–626

Falah M, Gupta RS (1994) Cloning of the hsp70 (dnaK) genes from *Rhizobium meliloti* and *Pseudomonas cepacia*: phylogenetic analyses of mitochondrial origin based on a highly conserved protein sequence. J Bacteriol 176:7748–7753
Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37
Folsch H, Guiard B, Neupert W, Stuart RA (1996) Internal targeting signal of the BCS1 protein: a novel mechanism of import into mitochondria. EMBO J 15:479–487
Frazier AE, Dudek J, Guiard B, Voos W, Li Y, Lind M, Meisinger C, Geissler A, Sickmann A, Meyer HE, Bilanchone V, Cumsky MG, Truscott KN, Pfanner N, Rehling P (2004) Pam16 has an essential role in the mitochondrial protein import motor. Nat Struct Mol Biol 11:226
Freibert S-A, Goldberg AV, Hacker C, Molik S, Dean P, Williams TA, Nakjang S, Long S, Sendra K, Bill E, Heinz E, Hirt RP, Lucocq JM, Embley TM, Lill R (2017) Evolutionary conservation and in vitro reconstitution of microsporidian iron–sulfur cluster biosynthesis. Nat Commun 8:13932
Fukasawa Y, Oda T, Tomii K, Imai K (2017) Origin and evolutionary alteration of the mitochondrial import system in eukaryotic lineages. Mol Biol Evol 34:1574–1586
Gabaldón T, Huynen MA (2003) Reconstruction of the proto-mitochondrial metabolism. Science 301:609
Gakh O, Cavadini P, Isaya G (2002) Mitochondrial processing peptidases. Biochim Biophys Acta, Mol Cell Res 1592:63–77
García-Rodríguez LJ, Gay AC, Pon LA (2007) Puf3p, a Pumilio family RNA binding protein, localizes to mitochondria and regulates mitochondrial biogenesis and motility in budding yeast. J Cell Biol 176:197–207
Garg SG, Gould SB (2016) The role of charge in protein targeting evolution. Trends Cell Biol 26:894–905
Garg S, Stölting J, Zimorski V, Rada P, Tachezy J, Martin WF, Gould SB (2015) Conservation of transit peptide-independent protein import into the mitochondrial and hydrogenosomal matrix. Genome Biol Evol 7:2716–2726
Gavel Y, von Heijne G (1990) Cleavage-site motifs in mitochondrial targeting peptides. Protein Eng Des Sel 4:33–37
Geissler A, Chacinska A, Truscott KN, Wiedemann N, Brandner K, Sickmann A, Meyer HE, Meisinger C, Pfanner N, Rehling P (2002) The mitochondrial presequence translocase: an essential role of Tim50 in directing preproteins to the import channel. Cell 111:507–518
Geli V, Yang MJ, Suda K, Lustig A, Schatz G (1990) The MAS-encoded processing protease of yeast mitochondria. Overproduction and characterization of its two nonidentical subunits. J Biol Chem 265:19216–19222
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
Gentle IE, Perry AJ, Alcock FH, Likić VA, Dolezal P, Ng ET, Purcell AW, McConnville M, Naderer T, Chanez A-L, Charrière F, Aschinger C, Schneider A, Tokatlidis K, Lithgow T (2007) Conserved motifs reveal details of ancestry and structure in the small TIM chaperones of the mitochondrial intermembrane Space. Mol Biol Evol 24:1149–1160
Germot A, Philippe H, Le Guyader H (1996) Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. Proc Natl Acad Sci U S A 93:14614–14617
Germot A, Philippe H, Le Guyader H (1997) Evidence for loss of mitochondria in *Microsporidia* from a mitochondrial-type HSP70 in *Nosema locustae*. Mol Biochem Parasitol 87:159–168
Glaser E, Sjoling S, Tanudji M, Whelan J (1998) Mitochondrial protein import in plants. Signals, sorting, targeting, processing and regulation. Plant Mol Biol 38:311–338
Glick BS, Brandt A, Cunningham K, Müller S, Hallberg RL, Schatz G (1992) Cytochromes c1 and b2 are sorted to the intermembrane space of yeast mitochondria by a stop-transfer mechanism. Cell 69:809–822

Goldberg AV, Molik S, Tsaousis AD, Neumann K, Kuhnke G, Delbac F, Vivares CP, Hirt RP, Lill R, Embley TM (2008) Localization and functionality of microsporidian iron–sulphur cluster assembly proteins. Nature 452:624–628

Gray MW (2015) Mosaic nature of the mitochondrial proteome: Implications for the origin and evolution of mitochondria. Proc Natl Acad Sci U S A 112:10133–10138

Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. Science 283:1476–1481

Gupta RS (2006) The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes. FEMS Microbiol Rev 24:367–402

Gupta RS (2018) Evolution of the chaperonin families (HSP60, HSP 10 and TCP-1) of proteins and the origin of eukaryotic cells. Mol Microbiol 15:1–11

Gupta RS, Singh B (1994) Phylogenetic analysis of 70 kD heat shock protein sequences suggests a chimeric origin for the eukaryotic cell nucleus. Curr Biol 4:1104–1114

Hahne K, Haucke V, Ramage L, Schatz G (1994) Incomplete arrest in the outer membrane sorts NADH-cytochrome b5 reductase to two different submitochondrial compartments. Cell 79:829–839

Hartl F-U, Hlodan R, Langer T (1994) Molecular chaperones in protein folding: the art of avoiding sticky situations. Trends Biochem Sci 19:20–25

Hausler T, Stierhof YD, Blattner J, Clayton C (1997) Conservation of mitochondrial targeting sequence function in mitochondrial and hydrogenosomal proteins from the early-branching eukaryotes *Crithidia*, *Trypanosoma* and *Trichomonas*. Eur J Cell Biol 73:240–251

Heins L, Schmitz UK (1996) A receptor for protein import into potato mitochondria. Plant J 9:829–839

Heinz E, Lithgow T (2013) Back to basics: a revealing secondary reduction of the mitochondrial protein import pathway in diverse intracellular parasites. Biochim Biophys Acta, Mol Cell Res 1833:295–303

Henriquez FL, Richards TA, Roberts F, McLeod R, Roberts CW (2005) The unusual mitochondrial compartment of *Cryptosporidium parvum*. Trends Parasitol 21:68–74

Herrmann JM (2003) Converting bacteria to organelles: evolution of mitochondrial protein sorting. Trends Microbiol 11:74–79

Hill K, Model K, Ryan MT, Dietmeier K, Martin F, Wagner R, Pfanner N (1998) Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. Nature 395:516

Höhr AIC, Lindau C, Wirth C, Qiu J, Stroud DA, Kutik S, Guiard B, Hunte C, Becker T, Pfanner N, Wiedemann N (2018) Membrane protein insertion through a mitochondrial β-barrel gate. Science 359:eaah6834

Hönlinger A, Kübrich M, Moczko M, Gärtner F, Mallet L, Bussereau F, Eckerskorn C, Lottspeich F, Dietmeier K, Jacquet M (1995) The mitochondrial receptor complex: Mom22 is essential for cell viability and directly interacts with preproteins. Mol Cell Biol 15:3382–3389

Hoogenraad NJ, Ward LA, Ryan MT (2002) Import and assembly of proteins into mitochondria of mammalian cells. Biochim Biophys Acta, Mol Cell Res 1592:97–105

Hoppins SC, Nargang FE (2004) The Tim8-Tim13 complex of Neurospora crassa functions in the assembly of proteins into both mitochondrial membranes. J Biol Chem 279:12396–12405

Horie C, Suzuki H, Sakaguchi M, Mihara K (2002) Characterization of signal that directs C-tail–anchored proteins to mammalian mitochondrial outer membrane. Mol Biol Cell 13:1615–1625

Horner DS, Hirt RP, Kilvington S, Lloyd D, Embley TM (1996) Molecular data suggest an early acquisition of the mitochondrion endosymbiont. Proc Biol Sci 263:1053–1059

Horst M, Oppliger W, Rospert S, Schönfeld H, Schatz G, Azem A (1997) Sequential action of two hsp70 complexes during protein import into mitochondria. EMBO J 16:1842–1849

Hrdy I, Hirt RP, Dolezal P, Bardonová L, Foster PG, Tachezy J, Martin Embley T (2004) Trichomonas hydrogenosomes contain the NADH dehydrogenase module of mitochondrial complex I. Nature 432:618

Hrdy I, Müller M (1995a) Primary structure and eubacterial relationships of the pyruvate: ferredoxin oxidoreductase of the amitochondriate eukaryote *Trichomonas vaginalis*. J Mol Evol 41:388–396

Hrdy I, Müller M (1995b) Primary structure of the hydrogenosomal malic enzyme of *Trichomonas vaginalis* and its relationship to homologous enzymes. J Eukaryot Microbiol 42:593–603
Hulett JM, Lueder F, Chan NC, Perry AJ, Wolynec P, Likić VA, Gooley PR, Lithgow T (2008) The transmembrane segment of Tom20 is recognized by Mim1 for docking to the mitochondrial TOM complex. J Mol Biol 376:694–704
Hwang DK, Claypool SM, Leuenberger D, Tienson HL, Koehler CM (2007) Tim54p connects inner membrane assembly and proteolytic pathways in the mitochondrion. J Cell Biol 178:1161–1175
Ishikawa D, Yamamoto H, Tamura Y, Moritoh K, Endo T (2004) Two novel proteins in the mitochondrial outer membrane mediate β-barrel protein assembly. J Cell Biol 166:621–627
Janssen BD, Chen Y-P, Molgora BM, Wang SE, Simoes-Barbosa A, Johnson PJ (2018) CRISPR/Cas9-mediated gene modification and gene knock out in the human-infective parasite *Trichomonas vaginalis*. Sci Rep 8:270
Jarosch E, Rodel G, Schweyen RJ (1997) A soluble 12-kDa protein of the mitochondrial intermembrane space, Mrs11p, is essential for mitochondrial biogenesis and viability of yeast cells. Mol Gen Genet 255:157–165
Johnson PJ, Oliveira CE, Gorrell TE, Müller M (1990) Molecular analysis of the hydrogenosomal ferredoxin of the anaerobic protist *Trichomonas vaginalis*. Proc Natl Acad Sci U S A 87:6097–6101
Jores T, Klinger A, Groß LE, Kawano S, Flinner N, Duchardt-Ferner E, Wöhnert J, Kalbacher H, Endo T, Schleiff E, Rapaport D (2016) Characterization of the targeting signal in mitochondrial β-barrel proteins. Nat Commun 7:12036
Kang Y, Baker MJ, Liem M, Louber J, McKenzie M, Atukorala I, Ang C-S, Keerthikumar S, Mathivanan S, Stojanovski D (2016) Tim29 is a novel subunit of the human TIM22 translocase and is involved in complex assembly and stability. Elife 5
Katinka MD, Duprat S, Cornillot E, Méténier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivarès CP (2001) Genome sequence and gene compaction of the eukaryote parasite Encephalitozoon cuniculi. Nature 414:450
Kerscher O, Holder J, Srinivasan M, Leung RS, Jensen RE (1997) The Tim54p–Tim22p Complex Mediates Insertion of Proteins into the Mitochondrial Inner Membrane. J Cell Biol 139:1663–1675
Kerscher O, Sepuri NB, Jensen RE, Fox TD (2000) Tim18p is a new component of the Tim54p-Tim22p translocon in the mitochondrial inner membrane. Mol Biol Cell 11:103–116
Kitada S, Uchiyama T, Funatsu T, Kitada Y, Ogishima T, Ito A (2007) A protein from a parasitic microorganism, *Rickettsia prowazekii*, can cleave the signal sequences of proteins targeting mitochondria. J Bacteriol 189:844–850
Koehler CM (2004) New developments in mitochondrial assembly. Annu Rev Cell Dev Biol 20:309–335
Koehler CM, Murphy MP, Bally NA, Leuenberger D, Oppliger W, Dolfini L, Junne T, Schatz G, Or E (2000) Tim18p, a new subunit of the TIM22 complex that mediates insertion of imported proteins into the yeast mitochondrial inner membrane. Mol Cell Biol 20:1187–1193
Kornmann B, Currie E, Collins SR, Schuldiner M, Nunnari J, Weissman JS, Walter P (2009) An ER-mitochondria tethering complex revealed by a synthetic biology screen. Science 325:477–481
Kovermann P, Truscott KN, Guiard B, Rehling P, Sepuri NB, Müller H, Jensen RE, Wagner R, Pfanner N (2002) Tim22, the essential core of the mitochondrial protein insertion complex, forms a voltage-activated and signal-gated channel. Mol Cell 9:363–373
Kozany C, Mokranjac D, Sichting M, Neupert W, Hell K (2004) The J domain–related cochaperone Tim16 is a constituent of the mitochondrial TIM23 preprotein translocase. Nat Struct Mol Biol 11:234

Kozjak V, Wiedemann N, Milenkovic D, Lohaus C, Meyer HE, Guiard B, Meisinger C, Pfanner N (2003) An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane. J Biol Chem 278:48520–48523

Kumar A, Harrison PM, Cheung K-H, Lan N, Echols N, Bertone P, Miller P, Gerstein MB, Snyder M (2002) An integrated approach for finding overlooked genes in yeast. Nat Biotechnol 20:58–63

Künkele K, Heins S, Dembowski M, Nargang FE, Benz R, Thieffry M, Walz J, Lill R, Nussberger S, Neupert W (1998) The preprotein translocation channel of the outer membrane of mitochondria. Cell 93:1009–1019

LaGier MJ, Tachezy J, Stejskal F, Kutisova K, Keithly JS (2003) Mitochondrial-type iron-sulfur cluster biosynthesis genes (IscS and IscU) in the apicomplexan *Cryptosporidium parvum*. Microbiology 149:3519–3530

Lahti CJ, Bradley PJ, Johnson PJ (1994) Molecular characterization of the alpha-subunit of *Trichomonas vaginalis* hydrogenosomal succinyl CoA synthetase. Mol Biochem Parasitol 66:309–318

Lahti CJ, d'Oliveira CE, Johnson PJ (1992) Beta-succinyl-coenzyme A synthetase from Trichomonas vaginalis is a soluble hydrogenosomal protein with an amino-terminal sequence that resembles mitochondrial presequences. J Bacteriol 174:6822–6830

Land KM, Delgadillo-Correa MG, Tachezy J, Vanacova S, Hsieh CL, Sutak R, Johnson PJ (2003) Targeted gene replacement of a ferredoxin gene in *Trichomonas vaginalis* does not lead to metronidazole resistance. Mol Microbiol 51:115–122

Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH (2007) Classical nuclear localization signals: definition, function, and interaction with importin alpha. J Biol Chem 282:5101–5105

Länge S, Rozario C, Müller M (1994) Primary structure of the hydrogenosomal adenylate kinase of *Trichomonas vaginalis* and its phylogenetic relationships. Mol Biochem Parasitol 66:297–308

Lee CM, Sedman J, Neupert W, Stuart RA (1999) The DNA helicase, Hmi1p, is transported into mitochondria by a C-terminal cleavable targeting signal. J Biol Chem 274:20937–20942

Leger MM, Kolisko M, Kamikawa R, Stairs CW, Kume K, Čepička I, Silberman JD, Andersson JO, Xu F, Yabuki A, Eme L, Zhang Q, Takishita K, Inagaki Y, Simpson AGB, Hashimoto T, Roger AJ (2017) Organelles that illuminate the origins of *Trichomonas* hydrogenosomes and *Giardia* mitosomes. Nat Ecol Evol 1:0092

Li Y, Dudek J, Guiard B, Pfanner N, Rehling P, Voos W (2004) The presequence translocase-associated protein import motor of mitochondria. J Biol Chem 279:38047–38054

Likić VA, Perry A, Hulett J, Derby M, Traven A, Waller RF, Keeling PJ, Koehler CM, Curran SP, Gooley PR, Lithgow T (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

Likic VA, Dolezal P, Celik N, Dagley M, Lithgow T (2010) Using hidden Markov models to discover new protein transport machines. Methods Mol Biol 619:271–284

Lin L, Pan G, Li T, Dang X, Deng Y, Ma C, Chen J, Luo J, Zhou Z (2012) The protein import pore Tom40 in the microsporidian *Nosema bombycis*. J Eukaryot Microbiol 59:251–257

Linford AS, Moreno H, Good KR, Zhang H, Singh U, Petri WA (2009) Short hairpin RNA-mediated knockdown of protein expression in *Entamoeba histolytica*. BMC Microbiol 9:38

Lithgow T, Schneider A (2010) Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes. Philos Trans R Soc Lond Ser B Biol Sci 365:799–817

Liu S, Roellig DM, Guo Y, Li N, Frace MA, Tang K, Zhang L, Feng Y, Xiao L (2016) Evolution of mitosome metabolism and invasion-related proteins in *Cryptosporidium*. BMC Genomics 17:1006

Loftus B, Anderson I, Davies R, Alsmark UCM, Samuelson J, Amedeo P, Roncaglia P, Berriman M, Hirt RP, Mann BJ, Nozaki T, Suh B, Pop M, Duchene M, Ackers J, Tannich E, Leippe M, Hofer M, Bruchhaus I, Willhoeft U, Bhattacharya A, Chillingworth T, Churcher C, Hance Z, Harris B, Harris D, Jagels K, Moule S, Mungall K, Ormond D, Squares R,

Whitehead S, Quail MA, Rabbinowitsch E, Norbertczak H, Price C, Wang Z, Guillén N, Gilchrist C, Stroup SE, Bhattacharya S, Lohia A, Foster PG, Sicheritz-Ponten T, Weber C, Singh U, Mukherjee C, El-Sayed NM, Petri WA, Clark CG, Embley TM, Barrell B, Fraser CM, Hall N (2005) The genome of the protist parasite *Entamoeba histolytica*. Nature 433:865–868

Lucattini R, Likic VA, Lithgow T (2004) Bacterial proteins predisposed for targeting to mitochondria. Mol Biol Evol 21:652–658

Luciano P, Géli V (1996) The mitochondrial processing peptidase: function and specificity. Experientia 52:1077–1082

Maarse AC, Blom J, Grivell LA, Meijer M (1992) MPI1, an essential gene encoding a mitochondrial membrane protein, is possibly involved in protein import into yeast mitochondria. EMBO J 11:3619–3628

Maarse AC, Blom J, Keil P, Pfanner N, Meijer M (1994) Identification of the essential yeast protein MIM17, an integral mitochondrial inner membrane protein involved in protein import. FEBS Lett 349:215–221

Maćasev D, Whelan J, Newbigin E, Silva-Filho MC, Mulhern TD, Lithgow T (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

Maduke M, Roise D (1993) Import of a mitochondrial presequence into protein-free phospholipid vesicles. Science 260:364–367

Mai Z, Ghosh S, Frisardi M, Rosenthal B, Rogers R, Samuelson J (1999) Hsp60 is targeted to a cryptic mitochondrion-derived organelle in the microaerophilic protozoan parasite *Entamoeba histolytica*. Mol Cell Biol 19:2198–2205

Makiuchi T, Nozaki T (2014) Highly divergent mitochondrion-related organelles in anaerobic parasitic protozoa. Biochimie 100:3–17

Makiuchi T, Mi-ichi F, Nakada-Tsukui K, Nozaki T (2013) Novel TPR-containing subunit of TOM complex functions as cytosolic receptor for *Entamoeba* mitosomal transport. Sci Rep 3:1129

Makki A, Rada P, Žárský V, Kereïche S, Kováčik L, Novotný M, Jores T, Rapaport D, Tachezy J (2019) Triplet-pore structure of a highly divergent TOM complex of hydrogenosomes in *Trichomonas vaginalis*. PLoS Biol 17:32

Mani J, Desy S, Niemann M, Chanfon A, Oeljeklaus S, Pusnik M, Schmidt O, Gerbeth C, Meisinger C, Warscheid B, Schneider A (2015) Mitochondrial protein import receptors in Kinetoplastids reveal convergent evolution over large phylogenetic distances. Nat Commun 6:6646

Mani J, Meisinger C, Schneider A (2016) Peeping at TOMs—diverse entry gates to mitochondria provide insights into the evolution of eukaryotes. Mol Biol Evol 33:337–351

Manning-Krieg UC, Scherer PE, Schatz G (1991) Sequential action of mitochondrial chaperones in protein import into the matrix. EMBO J 10:3273–3280

Martijn J, Vosseberg J, Guy L, Offre P, Ettema TJG (2018) Deep mitochondrial origin outside the sampled alphaproteobacteria. Nature 557:101–105

Martin J, Langer T, Boteva R, Schramel A, Horwich AL, Hartl FU (1991) Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. Nature 352:36–42

Martin WF, Garg S, Zimorski V (2015) Endosymbiotic theories for eukaryote origin. Philos Trans R Soc B Biol Sci 370:20140330

Martincová E, Voleman L, Pyrih J, Žárský V, Vondráčková P, Kolísko M, Tachezy J, Doležal P (2015) Probing the biology of *Giardia intestinalis* mitosomes using in vivo enzymatic tagging. Mol Cell Biol 35:2864–2874

Martinez-Caballero S, Grigoriev SM, Herrmann JM, Campo ML, Kinnally KW (2007) Tim17p regulates the twin pore structure and voltage gating of the mitochondrial protein import complex TIM23. J Biol Chem 282:3584–3593

Marvin-Sikkema FD, Kraak MN, Veenhuis M, Gottschal JC, Prins RA (1993) The hydrogenosomal enzyme hydrogenase from the anaerobic fungus Neocallimastix sp. L2 is recognized by

antibodies, directed against the C-terminal microbody protein targeting signal SKL. Eur J Cell Biol 61:86–91
McArthur AG, Morrison HG, Nixon JE, Passamaneck NQ, Kim U, Hinkle G, Crocker MK, Holder ME, Farr R, Reich CI, Olsen GE, Aley SB, Adam RD, Gillin FD, Sogin ML (2000) The *Giardia* genome project database. FEMS Microbiol Lett 189:271–273
McFadden GI (1999) Endosymbiosis and evolution of the plant cell. Curr Opin Plant Biol 2:513–519
McInally S, Hagen K, Nosala C, Williams J, Nguyen K, Booker J, Jones K, Dawson SC (2018) Robust and stable transcriptional repression in *Giardia* using CRISPRi. Mol Biol Cell 30:119–130
Meier S, Neupert W, Herrmann JM (2005) Conserved N-terminal negative charges in the Tim17 subunit of the TIM23 translocase play a critical role in the import of preproteins into mitochondria. J Biol Chem 280:7777–7785
Meinecke M, Wagner R, Kovermann P, Guiard B, Mick DU, Hutu DP, Voos W, Truscott KN, Chacinska A, Pfanner N, Rehling P (2006) Tim50 maintains the permeability barrier of the mitochondrial inner membrane. Science 312:1523–1526
Meisinger C, Ryan MT, Hill K, Model K, Lim JH, Sickmann A, Müller H, Meyer HE, Wagner R, Pfanner N (2001) Protein import channel of the outer mitochondrial membrane: a highly stable Tom40-Tom22 core structure differentially interacts with preproteins, small tom proteins, and import receptors. Mol Cell Biol 21:2337–2348
Meisinger C, Rissler M, Chacinska A, Szklarz LKS, Milenkovic D, Kozjak V, Schönfisch B, Lohaus C, Meyer HE, Yaffe MP, Guiard B, Wiedemann N, Pfanner N (2004) The mitochondrial morphology protein Mdm10 functions in assembly of the preprotein translocase of the outer membrane. Dev Cell 7:61–71
Meisinger C, Pfannschmidt S, Rissler M, Milenkovic D, Becker T, Stojanovski D, Youngman MJ, Jensen RE, Chacinska A, Guiard B, Pfanner N, Wiedemann N (2007) The morphology proteins Mdm12/Mmm1 function in the major β-barrel assembly pathway of mitochondria. EMBO J 26:2229–2239
Ménétret J-F, Schaletzky J, Clemons WM, Osborne AR, Skånland SS, Denison C, Gygi SP, Kirkpatrick DS, Park E, Ludtke SJ, Rapoport TA, Akey CW (2007) Ribosome binding of a Single copy of the SecY complex: implications for protein translocation. Mol Cell 28:1083–1092
Mesecke N, Terziyska N, Kozany C, Baumann F, Neupert W, Hell K, Herrmann JM (2005) A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. Cell 121:1059–1069
Mi-ichi F, Abu Yousuf M, Nakada-Tsukui K, Nozaki T (2009) Mitosomes in *Entamoeba histolytica* contain a sulfate activation pathway. Proc Natl Acad Sci U S A 106:21731–21736
Milenkovic D, Kozjak V, Wiedemann N, Lohaus C, Meyer HE, Guiard B, Pfanner N, Meisinger C (2004) Sam35 of the mitochondrial protein sorting and assembly machinery is a peripheral outer membrane protein essential for cell viability. J Biol Chem 279:22781–22785
Model K, Meisinger C, Prinz T, Wiedemann N, Truscott KN, Pfanner N, Ryan MT (2001) Multistep assembly of the protein import channel of the mitochondrial outer membrane. Nat Struct Biol 8:361–370
Model K, Meisinger C, Kühlbrandt W (2008) Cryo-electron microscopy structure of a yeast mitochondrial preprotein translocase. J Mol Biol 383:1049–1057
Mokranjac D, Paschen SA, Kozany C, Prokisch H, Hoppins SC, Nargang FE, Neupert W, Hell K (2003a) Tim50, a novel component of the TIM23 preprotein translocase of mitochondria. EMBO J 22:816–825
Mokranjac D, Sichting M, Neupert W, Hell K (2003b) Tim14, a novel key component of the import motor of the TIM23 protein translocase of mitochondria. EMBO J 22:4945–4956
Morrison HG, Roger AJ, Nystul TG, Gillin FD, Sogin ML (2001) *Giardia lamblia* expresses a proteobacterial-like DnaK homolog. Mol Biol Evol 18:530–541

Müller A, Rassow J, Grimm J, Machuy N, Meyer TF, Rudel T (2002) VDAC and the bacterial porin PorB of Neisseria gonorrhoeae share mitochondrial import pathways. EMBO J 21:1916–1929
Mukherjee M, Brown MT, McArthur AG, Johnson PJ (2006a) Proteins of the glycine decarboxylase complex in the hydrogenosome of *Trichomonas vaginalis*. Eukaryot Cell 5:2062–2071
Mukherjee M, Sievers SA, Brown MT, Johnson PJ (2006b) Identification and biochemical characterization of serine hydroxymethyl transferase in the hydrogenosome of *Trichomonas vaginalis*. Eukaryot Cell 5:2072–2078
Murcha MW, Millar AH, Whelan J (2005a) The N-terminal cleavable extension of plant carrier proteins is responsible for efficient insertion into the inner mitochondrial membrane. J Mol Biol 351:16–25
Murcha MW, Rudhe C, Elhafez D, Adams KL, Daley DO, Whelan J (2005b) Adaptations required for mitochondrial import following mitochondrial to nucleus gene transfer of ribosomal protein S10. Plant Physiol 138:2134–2144
Neupert W (1997) Protein import into mitochondria. Annu Rev Biochem 66:863–917
Nixon JEJ, Wang A, Morrison HG, McArthur AG, Sogin ML, Loftus BJ, Samuelson J (2002) A spliceosomal intron in *Giardia lamblia*. Proc Natl Acad Sci U S A 99:3701–3705
Nomura H, Athauda SBP, Wada H, Maruyama Y, Takahashi K, Inoue H (2006) Identification and reverse genetic analysis of mitochondrial processing peptidase and the core protein of the cytochrome bc1 complex of *Caenorhabditis elegans*, a model parasitic nematode. J Biochem 139:967–979
Nunnari J, Fox TD, Walter P (1993) A mitochondrial protease with two catalytic subunits of nonoverlapping specificities. Science 262:1997–2004
Nyathi Y, Wilkinson BM, Pool MR (2013) Co-translational targeting and translocation of proteins to the endoplasmic reticulum. Biochim Biophys Acta, Mol Cell Res 1833:2392–2402
Osborne AR, Rapoport TA, van den Berg B (2005) Protein translocation by the Sec61/SecY channel. Annu Rev Cell Dev Biol 21:529–550
Park E, Rapoport TA (2012) Mechanisms of Sec61/SecY-mediated protein translocation across membranes. Annu Rev Biophys 41:21–40
Paschen SA, Waizenegger T, Stan T, Preuss M, Cyrklaff M, Hell K, Rapaport D, Neupert W (2003) Evolutionary conservation of biogenesis of beta-barrel membrane proteins. Nature 426:862–866
Paschen SA, Neupert W, Rapaport D (2005) Biogenesis of β-barrel membrane proteins of mitochondria. Trends Biochem Sci 30:575–582
Pemberton LF, Paschal BM (2005) Mechanisms of receptor-mediated nuclear import and nuclear export. Traffic 6:187–198
Perry AJ, Hulett JM, Likić VA, Lithgow T, Gooley PR (2006) Convergent evolution of receptors for protein import into mitochondria. Curr Biol 16:221–229
Peyretaillade E, Broussolle V, Peyret P, Metenier G, Gouy M, Vivares CP (1998) Microsporidia, amitochondrial protists, possess a 70-kDa heat shock protein gene of mitochondrial evolutionary origin. Mol Biol Evol 15:683–689
Pfanner N, Geissler A (2001) Versatility of the mitochondrial protein import machinery. Nat Rev Mol Cell Biol 2:339–349
Picotti P, Clément-Ziza M, Lam H, Campbell DS, Schmidt A, Deutsch EW, Röst H, Sun Z, Rinner O, Reiter L, Shen Q, Michaelson JJ, Frei A, Alberti S, Kusebauch U, Wollscheid B, Moritz RL, Beyer A, Aebersold R (2013) A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. Nature 494:266–270
Plümper E, Bradley PJ, Johnson PJ (2000) Competition and protease sensitivity assays provide evidence for the existence of a hydrogenosomal protein import machinery in *Trichomonas vaginalis*. Mol Biochem Parasitol 106:11–20
Putignani L, Tait A, Smith HV, Horner D, Tovar J, Tetley L, Wastling JM (2004) Characterization of a mitochondrion-like organelle in *Cryptosporidium parvum*. Parasitology 129:1–18
Pütz S, Dolezal P, Gelius-Dietrich G, Bohacova L, Tachezy J, Henze K (2006) Fe-hydrogenase maturases in the hydrogenosomes of *Trichomonas vaginalis*. Eukaryot Cell 5:579–586

Pyrihová E, Motyčková A, Voleman L, Wandyszewska N, Fišer R, Seydlová G, Roger A, Kolísko M, Doležal P (2018) A single Tim translocase in the mitosomes of *Giardia intestinalis* illustrates convergence of protein import machines in anaerobic eukaryotes. Genome Biol Evol 10:2813–2822

Qiu J, Wenz L-S, Zerbes RM, Oeljeklaus S, Bohnert M, Stroud DA, Wirth C, Ellenrieder L, Thornton N, Kutik S, Wiese S, Schulze-Specking A, Zufall N, Chacinska A, Guiard B, Hunte C, Warscheid B, van der Laan M, Pfanner N, Wiedemann N, Becker T (2013) Coupling of mitochondrial import and export translocases by receptor-mediated supercomplex formation. Cell 154:596–608

Rada P, Doležal P, Jedelský PL, Bursac D, Perry AJ, Šedinová M, Smíšková K, Novotný M, Beltrán NC, Hrdý I, Lithgow T, Tachezy J (2011) The core components of organelle biogenesis and membrane transport in the hydrogenosomes of *Trichomonas vaginalis*. PLoS One 6:e24428

Rada P, Makki AR, Zimorski V, Garg S, Hampl V, Hrdý I, Gould SB, Tachezy J (2015) N-terminal presequence-independent import of phosphofructokinase into hydrogenosomes of *Trichomonas vaginalis*. Eukaryot Cell 14:1264–1275

Rada P, Makki A, Žárský V, Tachezy J (2019) Targeting of tail-anchored proteins to *Trichomonas vaginalis* hydrogenosomes. Mol Microbiol

Rassow J, Dekker PJ, van Wilpe S, Meijer M, Soll J (1999) The preprotein translocase of the mitochondrial inner membrane: function and evolution. J Mol Biol 286:105–120

Rawlings ND, Barrett AJ (1995) Evolutionary families of metallopeptidases. Methods Enzymol 248:183–228

Regoes A, Egoes A, Zourmpanou D, León-Avila G, van der Giezen M, Tovar J, Hehl AB (2005) Protein import, replication, and inheritance of a vestigial mitochondrion. J Biol Chem 280:30557–30563

Rehling P, Model K, Brandner K, Kovermann P, Sickmann A, Meyer HE, Kühlbrandt W, Wagner R, Truscott KN, Pfanner N (2003) Protein insertion into the mitochondrial inner membrane by a twin-pore translocase. Science 299:1747–1751

Richards TA, van der Giezen M (2006) Evolution of the Isd11-IscS complex reveals a single alpha-proteobacterial endosymbiosis for all eukaryotes. Mol Biol Evol 23:1341–1344

Riordan CE, Ault JG, Langreth SG, Keithly JS (2003) *Cryptosporidium parvum* Cpn60 targets a relict organelle. Curr Genet 44:138–147

Roger AJ, Clark CG, Doolittle WF (1996) A possible mitochondrial gene in the early-branching amitochondriate protist *Trichomonas vaginalis*. Proc Natl Acad Sci U S A 93:14618–14622

Roger AJ, Svärd SG, Tovar J, Clark CG, Smith MW, Gillin FD, Sogin ML (1998) A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. Proc Natl Acad Sci U S A 95:229–234

Roger AJ, Muñoz-Gómez SA, Kamikawa R (2017) The origin and diversification of mitochondria. Curr Biol 27:R1177–R1192

Röhl T, Motzkus M, Soll J (1999) The outer envelope protein OEP24 from pea chloroplasts can functionally replace the mitochondrial VDAC in yeast. FEBS Lett 460:491–494

Roise D, Horvath SJ, Tomich JM, Richards JH, Schatz G (1986) A chemically synthesized pre-sequence of an imported mitochondrial protein can form an amphiphilic helix and perturb natural and artificial phospholipid bilayers. EMBO J 5:1327–1334

Rospert S, Junne T, Glick BS, Schatz G (1993) Cloning and disruption of the gene encoding yeast mitochondrial chaperonin 10, the homolog of *E. coli* groES. FEBS Lett 335:358–360

Rospert S, Looser R, Dubaquie Y, Matouschek A, Glick BS, Schatz G (1996) Hsp60-independent protein folding in the matrix of yeast mitochondria. EMBO J 15:764–774

Rout S, Zumthor JP, Schraner EM, Faso C, Hehl AB (2016) An interactome-centered protein discovery approach reveals novel components involved in mitosome function and homeostasis in *Giardia lamblia*. PLoS Pathog 12:e1006036

Ryan KR, Menold MM, Garrett S, Jensen RE (1994) SMS1, a high-copy suppressor of the yeast mas6 mutant, encodes an essential inner membrane protein required for mitochondrial protein import. Mol Biol Cell 5:529–538

Rye HS, Burston SG, Fenton WA, Beechem JM, Xu Z, Sigler PB, Horwich AL (1997) Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL. Nature 388:792–798

Sandoval P, León G, Gómez I, Carmona R, Figueroa P, Holuigue L, Araya A, Jordana X (2004) Transfer of RPS14 and RPL5 from the mitochondrion to the nucleus in grasses. Gene 324:139–147

Schatz G, Dobberstein B (1996) Common principles of protein translocation across membranes. Science 271:1519–1526

Schleyer M, Schmidt B, Neupert W (1982) Requirement of a membrane potential for the posttranslational transfer of proteins into mitochondria. Eur J Biochem 125:109–116

Schneider A, Behrens M, Scherer P, Pratje E, Michaelis G, Schatz G (1991) Inner membrane protease I, an enzyme mediating intramitochondrial protein sorting in yeast. EMBO J 10:247–254

Schneider RE, Brown MT, Shiflett AM, Dyall SD, Hayes RD, Xie Y, Loo JA, Johnson PJ (2011) The *Trichomonas vaginalis* hydrogenosome proteome is highly reduced relative to mitochondria, yet complex compared with mitosomes. Int J Parasitol 41:1421–1434

Shiota T, Imai K, Qiu J, Hewitt VL, Tan K, Shen H-H, Sakiyama N, Fukasawa Y, Hayat S, Kamiya M, Elofsson A, Tomii K, Horton P, Wiedemann N, Pfanner N, Lithgow T, Endo T (2015) Molecular architecture of the active mitochondrial protein gate. Science 349:1544–1548

Sirrenberg C, Bauer MF, Guiard B, Neupert W, Brunner M (1996) Import of carrier proteins into the mitochondrial inner membrane mediated by Tim22. Nature 384:582–585

Slamovits CH, Fast NM, Law JS, Keeling PJ (2004) Genome compaction and stability in microsporidian intracellular parasites. Curr Biol 14:891–896

Slapeta J, Keithly JS (2004) *Cryptosporidium parvum* mitochondrial-type HSP70 targets homologous and heterologous mitochondria. Eukaryot Cell 3:483–494

Šmíd O, Matušková A, Harris SR, Kučera T, Novotný M, Horváthová L, Hrdý I, Kutějová E, Hirt RP, Embley TM, Janata J, Tachezy J (2008) Reductive evolution of the mitochondrial processing peptidases of the unicellular parasites *Trichomonas vaginalis* and *Giardia intestinalis*. PLoS Pathog 4:e1000243

Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:W244–W248

Söllner T, Griffiths G, Pfaller R, Pfanner N, Neupert W (1989) MOM19, an import receptor for mitochondrial precursor proteins. Cell 59:1061–1070

Stewart M (2007) Molecular mechanism of the nuclear protein import cycle. Nat Rev Mol Cell Biol 8:195–208

Stojanovski D, Pfanner N, Wiedemann N (2007) Import of proteins into mitochondria. Methods Cell Biol 80:783–806

Stojanovski D, Bragoszewski P, Chacinska A (2012) The MIA pathway: a tight bond between protein transport and oxidative folding in mitochondria. Biochim Biophys Acta, Mol Cell Res 1823:1142–1150

Stuart R (2002) Insertion of proteins into the inner membrane of mitochondria: the role of the Oxa1 complex. Biochim Biophys Acta 1592:79–87

Stuart RA, Cyr DM, Craig EA, Neupert W (1994) Mitochondrial molecular chaperones: their role in protein translocation. Trends Biochem Sci 19:87–92

Sutak R, Dolezal P, Fiumera HL, Hrdy I, Dancis A, Delgadillo-Correa M, Johnson PJ, Müller M, Tachezy J (2004) Mitochondrial-type assembly of FeS centers in the hydrogenosomes of the amitochondriate eukaryote *Trichomonas vaginalis*. Proc Natl Acad Sci U S A 101:10368–10373

Szabo A, Langer T, Schröder H, Flanagan J, Bukau B, Hartl FU (1994) The ATP hydrolysis-dependent reaction cycle of the Escherichia coli Hsp70 system DnaK, DnaJ, and GrpE. Proc Natl Acad Sci U S A 91:10345–10349

Tachezy J, Sánchez LB, Müller M (2001) Mitochondrial type iron-sulfur cluster assembly in the amitochondriate eukaryotes *Trichomonas vaginalis* and *Giardia intestinalis*, as indicated by the phylogeny of IscS. Mol Biol Evol 18:1919–1928

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

Tovar J, León-Avila G, Sánchez LB, Sutak R, Tachezy J, van der Giezen M, Hernández M, Müller M, Lucocq JM (2003) Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. Nature 426:172–176

Truscott KN, Kovermann P, Geissler A, Merlin A, Meijer M, Driessen AJ, Rassow J, Pfanner N, Wagner R (2001) A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. Nat Struct Biol 8:1074–1082

Truscott KN, Voos W, Frazier AE, Lind M, Li Y, Geissler A, Dudek J, Müller H, Sickmann A, Meyer HE, Meisinger C, Guiard B, Rehling P, Pfanner N (2003) A J-protein is an essential subunit of the presequence translocase–associated protein import motor of mitochondria. J Cell Biol 163:707–713

Tsaousis AD, Gaston D, Stechmann A, Walker PB, Lithgow T, Roger AJ (2011) A functional Tom70 in the human parasite *Blastocystis* sp.: implications for the evolution of the mitochondrial import apparatus. Mol Biol Evol 28:781–791

van der Giezen M, Rechinger KB, Svendsen I, Durand R, Hirt RP, Fèvre M, Embley TM, Prins RA (1997) A mitochondrial-like targeting signal on the hydrogenosomal malic enzyme from the anaerobic fungus *Neocallimastix frontalis*: support for the hypothesis that hydrogenosomes are modified mitochondria. Mol Microbiol 23:11–21

van der Giezen M, Kiel JA, Sjollema KA, Prins RA (1998) The hydrogenosomal malic enzyme from the anaerobic fungus *Neocallimastix frontalis* is targeted to mitochondria of the methylotrophic yeast *Hansenula polymorpha*. Curr Genet 33:131–135

van der Giezen M, Slotboom DJ, Horner DS, Dyal PL, Harding M, Xue G-P, Embley TM, Kunji ERS (2002) Conserved properties of hydrogenosomal and mitochondrial ADP/ATP carriers: a common origin for both organelles. EMBO J 21:572–579

van der Giezen M, Birdsey GM, Horner DS, Lucocq J, Dyal PL, Benchimol M, Danpure CJ, Embley TM (2003) Fungal hydrogenosomes contain mitochondrial heat-shock proteins. Mol Biol Evol 20:1051–1061

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

van der Laan M, Chacinska A, Lind M, Perschil I, Sickmann A, Meyer HE, Guiard B, Meisinger C, Pfanner N, Rehling P (2005) Pam17 is required for architecture and translocation activity of the mitochondrial protein import motor. Mol Cell Biol 25:7449–7458

van der Laan M, Rissler M, Rehling P (2006a) Mitochondrial preprotein translocases as dynamic molecular machines. FEMS Yeast Res 6:849–861

van der Laan M, Wiedemann N, Mick DU, Guiard B, Rehling P, Pfanner N (2006b) A role for Tim21 in membrane-potential-dependent preprotein sorting in mitochondria. Curr Biol 16:2271–2276

van Wilpe S, Ryan MT, Hill K, Maarse AC, Meisinger C, Brix J, Dekker PJ, Moczko M, Wagner R, Meijer M, Guiard B, Hönlinger A, Pfanner N (1999) Tom22 is a multifunctional organizer of the mitochondrial preprotein translocase. Nature 401:485–489

Vinayak S, Pawlowic MC, Sateriale A, Brooks CF, Studstill CJ, Bar-Peled Y, Cipriano MJ, Striepen B (2015) Genetic modification of the diarrhoeal pathogen *Cryptosporidium parvum*. Nature 523:477

von Heijne G (1990) The signal peptide. J Membr Biol 115:195–201

von Heijne G, Steppuhn J, Herrmann RG (1989) Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem 180:535–545

Voncken FGJ, Boxma B, van Hoek AHAM, Akhmanova AS, Vogels GD, Huynen M, Veenhuis M, Hackstein JHP (2002) A hydrogenosomal [Fe]-hydrogenase from the anaerobic chytrid *Neocallimastix* sp. L2. Gene 284:103–112

Voos W, Röttgers K (2002) Molecular chaperones as essential mediators of mitochondrial biogenesis. Biochim Biophys Acta 1592:51–62
Voulhoux R, Bos MP, Geurtsen J, Mols M, Tommassen J (2003) Role of a highly conserved bacterial protein in outer membrane protein assembly. Science 299:262–265
Waizenegger T, Habib SJ, Lech M, Mokranjac D, Paschen SA, Hell K, Neupert W, Rapaport D (2004) Tob38, a novel essential component in the biogenesis of β-barrel proteins of mitochondria. EMBO Rep 5:704–709
Waizenegger T, Schmitt S, Zivkovic J, Neupert W, Rapaport D (2005) Mim1, a protein required for the assembly of the TOM complex of mitochondria. EMBO Rep 6:57–62
Waller RF, Jabbour C, Chan NC, Celik N, Likić VA, Mulhern TD, Lithgow T (2009) Evidence of a reduced and modified mitochondrial protein import apparatus in microsporidian mitosomes. Eukaryot Cell 8:19–26
Wenz L-S, Ellenrieder L, Qiu J, Bohnert M, Zufall N, van der Laan M, Pfanner N, Wiedemann N, Becker T (2015) Sam37 is crucial for formation of the mitochondrial TOM–SAM supercomplex, thereby promoting β-barrel biogenesis. J Cell Biol 210:1047–1054
Werhahn W, Niemeyer A, Jänsch L, Kruft V, Schmitz UK, Braun H (2001) Purification and characterization of the preprotein translocase of the outer mitochondrial membrane from *Arabidopsis*. Identification of multiple forms of TOM20. Plant Physiol 125:943–954
Wickner W, Schekman R (2005) Protein translocation across biological membranes. Science 310:1452–1456
Wiedemann N, Pfanner N (2017) Mitochondrial machineries for protein import and assembly. Annu Rev Biochem 86:685–714
Wiedemann N, Pfanner N, Ryan MT (2001) The three modules of ADP/ATP carrier cooperate in receptor recruitment and translocation into mitochondria. EMBO J 20:951–960
Wiedemann N, Kozjak V, Chacinska A, Schönfisch B, Rospert S, Ryan MT, Pfanner N, Meisinger C (2003) Machinery for protein sorting and assembly in the mitochondrial outer membrane. Nature 424:565–571
Wiedemann N, Truscott KN, Pfannschmidt S, Guiard B, Meisinger C, Pfanner N (2004) Biogenesis of the protein import channel Tom40 of the mitochondrial outer membrane. J Biol Chem 279:18188–18194
Williams BAP, Hirt RP, Lucocq JM, Embley TM (2002) A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. Nature 418:865–869
Wimley WC (2003) The versatile beta-barrel membrane protein. Curr Opin Struct Biol 13:404–411
Xu Z, Horwich AL, Sigler PB (1997) The crystal structure of the asymmetric GroEL–GroES–(ADP)7 chaperonin complex. Nature 388:741–750
Xu P, Widmer G, Wang Y, Ozaki LS, Alves JM, Serrano MG, Puiu D, Manque P, Akiyoshi D, Mackey AJ, Pearson WR, Dear PH, Bankier AT, Peterson DL, Abrahamsen MS, Kapur V, Tzipori S, Buck GA (2004) The genome of *Cryptosporidium hominis*. Nature 431:1107–1112
Yaffe MP, Ohta S, Schatz G (1985) A yeast mutant temperature-sensitive for mitochondrial assembly is deficient in a mitochondrial protease activity that cleaves imported precursor polypeptides. EMBO J 4:2069–2074
Yamamoto H, Esaki M, Kanamori T, Tamura Y, Nishikawa S, Endo T (2002) Tim50 is a subunit of the TIM23 complex that links protein translocation across the outer and inner mitochondrial membranes. Cell 111:519–528
Zara V, Dolce V, Capobianco L, Ferramosca A, Papatheodorou P, Rassow J, Palmieri F (2007) Biogenesis of eel liver citrate carrier (CIC): negative charges can substitute for positive charges in the presequence. J Mol Biol 365:958–967
Žárský V, Doležal P (2016) Evolution of the Tim17 protein family. Biol Direct 11:54
Zimorski V, Major P, Hoffmann K, Brás XP, Martin WF, Gould SB (2013) The N-terminal sequences of four major hydrogenosomal proteins are not essential for import into hydrogenosomes of *Trichomonas vaginalis*. J Eukaryot Microbiol 60:89–97