School of Doctoral Studies in Biological Sciences

University of south Bohemia in České Budějovice

Faculty of Science

Characterisation of novel mitochondrial proteins in *Trypanosoma brucei*

Ph.D. Thesis

Mgr. Vendula Rašková

Supervisor: Msc. Durante Ignacio Miguel, Ph.D.

Biology Centre, Institute of Parasitology, Czech Academy of Sciences, and

Faculty of Science, University of south Bohemia, České Budějovice, Czech Republic

České Budějovice 2023

This thesis should be cited as:

Rašková, V. (2023). Characterisation of novel mitochondrial proteins in *Trypanosoma brucei*. Ph.D. Thesis. University of south Bohemia, Faculty of Science, School of Doctoral Studies in Biological Sciences, České Budějovice, Czech Republic.

Annotation

This study analyses and characterises novel mitochondrial proteins in the parasitic protist *Trypanosoma brucei*. Applying phylogenetic analysis was described the evolutionary origin of ZapE protein in eukaryotes, using a newly developed proximity-dependent biotinylation approach (BioID2) we identified ZapE interaction partners like Oxa1. We also discovered a relationship when distribution of mitochondrial ZapE is restricted only to organisms with Oxa1, respiratory complexes, and a mitochondrial genome. TbPams were detected by phylogenetic analyses as orthologs of corresponding proteins in Opistokonts. We analyse the function of TbPam18 and TbPam16 in the replication of the mitochondrial DNA and determine, how the TMDs of TbPam18 and TbPam16 are essential for their functions. Finally, we evaluated a set of putative mitochondrial proteins of the heterolobosean *N. gruberi* defined by Localisation of Organelle Proteins by Isotope Tagging (LOPIT) and analyse the origin of mtFfh and mtFtsY.

Declaration

I hereby declare that I did all the work presented in this thesis by myself or in collaboration with co-authors of the presented papers and only using the cited literature.

České Budějovice. 13. 3. 2023   Vendula Rašková

Declaration [in Czech]

Prohlašuji, že svoji disertační práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své disertační práce, a to v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice. 13. 3. 2023

Vendula Rašková

This thesis originated from a partnership between the Faculty of Science, University of South Bohemia, and Institute of Parasitology, Biology Centre of the Czech Academy of Sciences, supporting doctoral studies in the Molecular and Cell biology and Genetics study program.

Financial support

Acknowledgements

among walls of laboratory in České Budějovice, but also Inka Škodová-Sveráková and her colleagues whom I met during internship in Bratislava.

I'm before the finish line of my entire PhD studies now and I couldn´t have done it on my own. Although their help might not be as obvious as the assistance and guidance of teachers and colleagues already mentioned, there are also other people I would like to give my thanks.

Above all, it is the influence of my "laboratory wife" Michaela Kunzová what shaped my academic path. Our friendship has begun by discussing books in the canteen and as is often the case, common hobbies and interests led to debates concerning our work projects and experiments, ultimately ending with sharing protocols and helping with presentations at conferences. I thank Míša for having my back and for being priceless emotional support during difficult moments and hardships every student encounters. I could always count on her to be there trying to cheer me up, make me laugh or cook some fabulous dish for me so I wouldn't go hungry during experiments.

I thank my family for always standing by my side even though the road I have taken was not an easy one. I really appreciate that you have never tried to talk me out of it nor questioned my decision.

List of papers and author's contribution
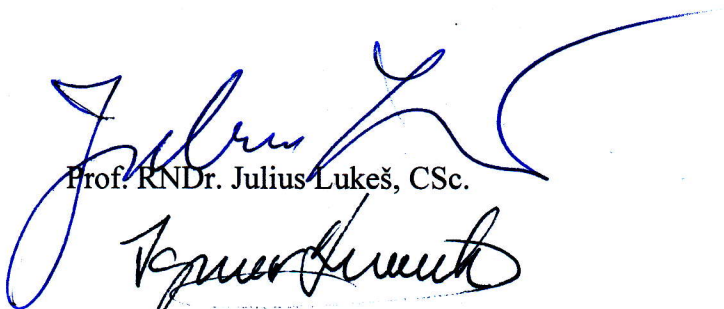
The thesis is based on the following papers (listed chronologically):

I.   Pyrih J.*, **Rašková V.***, Škodová-Sveráková I., Pánek T., Lukeš J. (2020) ZapE/Afg1 interacts with Oxa1 and its depletion causes a multifaceted phenotype. *PLoS ONE* 15: e0234918. [IF=3.752] DOI: 10.1371/journal.pone.0234918. *Equal contribution

VR participated in study design, experimental data generation, data analysis, data interpretation and manuscript preparation. Her contribution was 50%.

II. Pyrih J., Pánek T., Durante I.M., **Rašková V.**, Cimrhanzlová K., Kriegová E., Tsaousis A., Eliáš M., Lukeš J. (2021) Vestiges of the bacterial signal recognition particle-based protein targeting in mitochondria. *Molecular Biology and Evolution* 38: 3170–3187. DOI: 10.1093/molbev/msab090. [IF=8.800]
VR participated in in experimental procedures, data analysis and manuscript preparation. Her contribution was 15%.

III. von Känel C., Warscheid B., **Rašková V.**, Durante I.M., Oljeklaus S., Schneider A. Orthologies of presequence translocase-associated motor subunits are essential for kinetoplast DNA replication in procyclic form of *Trypanosoma brucei* (manuscript in preparation).
VR participated in study design, experimental data generation, data analysis, data interpretation and manuscript preparation. Her contribution was 30%.

IV. Cadena L.R., Hammond M., Svobodová M., Benz C., **Rašková V.**, Hashimi H., Durante I.M., Lukeš J. Novel protein complex involved in kinetoplast DNA replication and maintenance (manuscript in preparation).
VR participated in experimental data generation. Her contribution was 15%.

Co-authors agreement

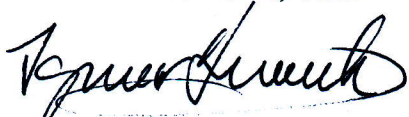Julius Lukeš, the co-supervisor of this Ph.D. thesis and co-author of three presented papers, fully acknowledges the contribution of Vendula Rašková.

Ignacio Durante, the supervisor of this Ph.D. thesis and co-author of three presented papers, fully acknowledges the contribution of Vendula Rašková.

Prof. RNDr. Julius Lukeš, CSc.

Msc. Durante Ignacio Miguel, Ph.D.

Contents

1. Introduction

1.1. The mitochondrion

The mitochondrion is an indispensable organelle of eukaryotes. It was originally acquired by the endosymbiosis of a relative of extant α-proteobacteria (Margulis, 1970). Until very recently the presence of mitochondria and related organelles in every studied eukaryote supported the view that mitochondria are essential cellular components, although at least one small group of eukaryotes represented by the protist *Monocercomonoides sp.* is based on available evidence considered as secondarily amitochondrial (Karnkowska *et al.,* 2016). While most mitochondria contain their own genome, some eukaryotes possess (highly) reduced forms of DNA-lacking mitochondria (Allen, 2015). Although there are numerous exceptions and also significant variation across eukaryotic lineages, the core components of the oxidative phosphorylation pathway constitute the majority of mitochondrial-encoded proteins (Embley and Martin, 2006).

Phylogenomic analyses have shown that the 'host' lineage of eukaryotes is most closely related to a newly discovered group of Archaea, known as the Asgards (Spang *et al.,* 2015; Zaremba-Niedzwiedzka *et al*. 2017). Molecular biology and genomic investigations of diverse protistan and multicellular lineages have further revealed that all known extant eukaryotes descend from a mitochondrion-containing ancestor (the last eukaryote common ancestor — LECA) that counterintuitively already had most of the cellular and genetic features of modern eukaryotes (Fig. 1) (Embley and Martin, 2006; Roger, 1999).

The timing of the mitochondrial endosymbiosis remains uncertain but postdates the first eukaryote common ancestor (FECA) and predates LECA. Although the complexity of the eukaryotic genome and proteome gradually increased during eukaryogenesis, the mitochondrial endosymbiont's genome was progressively reduced. This was likely caused by gradual incorporation of the host proteins as well as those acquired by horizontal gene transfer into the organelle (Roger *et al*., 2017).
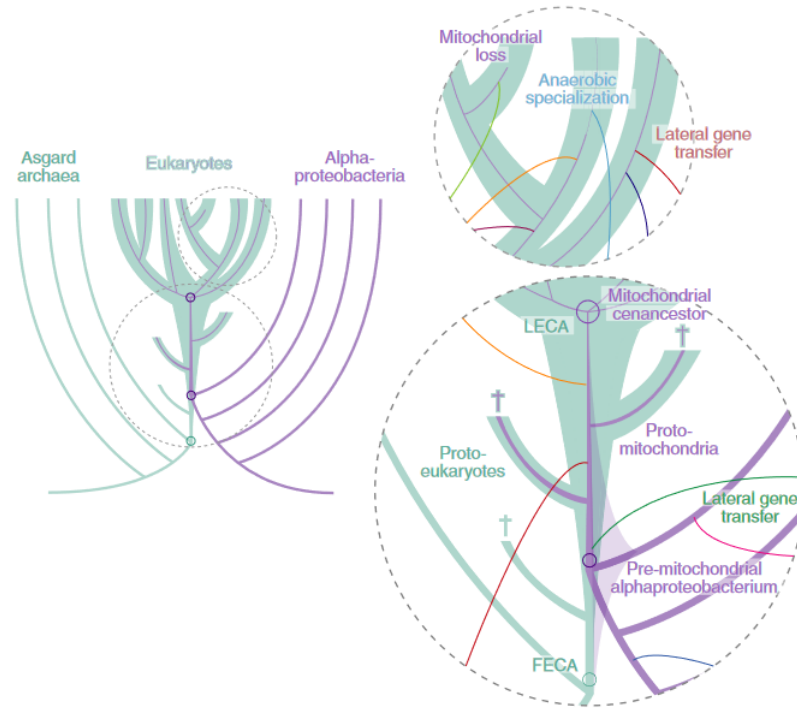
Figure 1: The origin and evolution of mitochondria and eukaryotes (Roger *et al*., 2017).

The mitochondrion is encircled by two membranes (the outer and inner membranes) which define the intermembrane space and matrix (Mannella, 1992). The inner membrane is impervious for most of the molecules and ions that must be actively transported into and out of the mitochondrion by different transporters. (Scheffler, 2007). The mitochondrion is a powerhouse of the cell, producing energy source for the key cellular processes: Oxidative phosphorylation (OXPHOS), the tricarboxylic acid cycle (TCA, also known as the Krebs cycle), β-oxidation, and synthesis of iron sulfur clusters.

OXPHOS pathway is present, in a very similar organization, in mitochondria from very distant eukaryotes. Moreover, the origin of the core protein components of the OXPHOS pathway can be traced back to the α-proteobacterial ancestor of mitochondria (Kurland and Anderson, 2000; Gabaldon and Huynen, 2003). OXPHOS is completely dependent on electron transport complexes (ETC) I, III and IV (Fig. 2) as they pump protons across the inner membrane into the intermembrane space against the proton gradient. The difference in charge on both sides of the membrane results in formation of the mitochondrial membrane potential (MMP). The MMP is used for protein import and together with the pH difference on both sides of the membrane generates proton motive force. As a result of these differences, protons move in favor of their concentration gradient providing the force to push protons through the FoF1 ATP synthase. This enzyme harnesses the proton motive force to drive the OXPHOS to generate ATP (Mitchell, 1961).
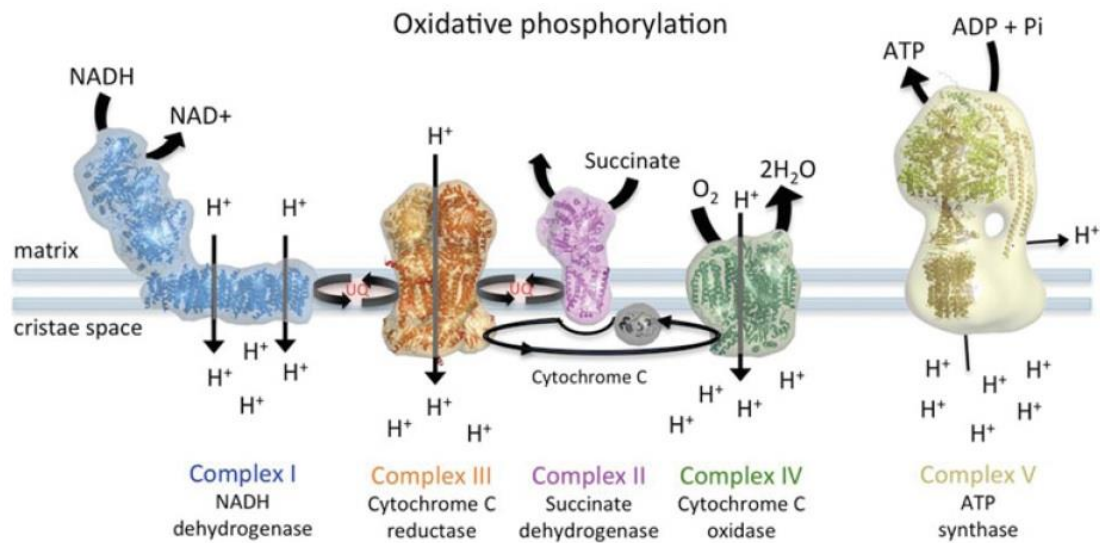
Figure 2: Complexes involved in oxidative phosphorylation in mitochondria (adapted from Davies and Daum, 2013).

The first and the largest from the ETC complexes that pumps protons across the inner mitochondrial membrane is complex I (NADH: ubiqinone oxidoreductase or NADH dehydrogenase, EC 7.1.1.2) (Morgan and Sazanov, 2008). Complex I binds reduced nicotinamide adenine dinucleotide (NADH) at flavin mononucleotide site and NADH gets oxidized to NAD. Then, the electrons from NADH are passed between seven iron-sulfur (Fe-S) clusters, and transferred to ubiquinone (Q), which is reduced to ubiquinol (QH2). This electron translocation is linked to the transport of protons from matrix to the intermembrane space. Subsequently, the reduced QH2 migrates through the inner mitochondrial membrane to complex III of ETC (Koolman and Roehm, 2005).

Complex II, also called succinate dehydrogenase (SDH or succinate-ubiquinone reductase, EC 1.3.5.1) (Lancaster, 2002; Koolman and Roehm, 2005) is an important enzymatic complex in both the Krebs cycle and the aerobic respiratory chains of mitochondria and prokaryotes (Hederstedt, 2003). In aerobic energy metabolism, electrons coming from succinate generated in the Krebs cycle are transferred to ubiquinone through complex II. Complex II catalysed the oxidation of succinate to fumarate and transfers its reducing equivalent to Q, an is also responsible for the reduction of fumarate. Fumarate serves as the terminal electron acceptor and electrons are transferred in the reverse direction (Tielens *et al*., 2002; Lancaster, 2002).

Complex III, i.e. ubiquinol-cytochrome c oxidoreductase or cytochrome c reductase (EC 7.1.1.8.) (Koolman and Roehm, 2005) is a homodimer that transfers electrons from ubiquinol to cytochrome c, so its oxidase QH2 back to Q. Each monomer, of this complex contains 11 subunits: 3 respiratory subunits, 2 core proteins and 6 low-molecular weight proteins. Three main, functionally important catalytic subunits with active redox centres are cytochrome b, cytochrome c 1, and the "Rieske" (2Fe-2S) protein (Iwata et al., 1998; Brzezinski et al., 2021; Zhang et al., 2011). The electrons obtained by the oxidation are passed onto cytochrome c. The reduced cytochrome c becomes separated from complex III and proceeds through the intermembrane space to complex IV. This redox reaction is coupled to the pumping four H^+ to the intermembrane space (Cramer et al., 2011).

Complex IV, also known as cytochrome c oxidase (EC 1.9.3.1) (Koolman and Roehm, 2005), is a member of the heme-copper oxidase family. This complex transfers electrons from cyt c to molecular oxygen, reducing it to water and meanwhile pumps two protons from the matrix side to the intermembrane space (Maldonado et al., 2021). The mammalian complex IV consists of 14 subunits and exists in two different states (assembled into supercomplexes or dispersed on mitochondrial inner membrane) (Zong et al, 2018). Variable number of accessory subunits, depending on the organism. The complex IV is composed of three conserved subunits (COX1, COX2, COX3) (Maldonado et al., 2021).

The proton gradient, which was created by ETC complexes I, III and IV is harvested for the generation of ATP. This universal cellular currency is produced in mitochondria mainly by F_oF_1 ATP synthase. Generally, the F_oF_1 ATP synthase is called complex V of the ETC. ATP is being formed from ADP and inorganic phosphate (Pi) during the process of OXPHOS (Walker, 2013). In this process, ATP synthase operate in the forward mode and utilize the proton motive force to produce ATP. Meanwhile, in the reverse mode, they act as ATP-consuming proton pumps contributing to the generation of electrochemical membrane potential (Cotter and Hill, 2003). ATP synthase consist of two major functional domain: the matrix-facing F1 and the membrane-embedded Fo. The F1 domain (F1-ATPase), is consists of a heterohexamer of α and β subunits and a central stalk (subunits γ, δ, and ε) ant its responsible for the phosphorylation. The Fo domain contains a motor, which generates rotation using the potential energy (Kuhlbrandt, 2019; Walker, 2013). Mitochondrial ATP synthases occur in dimers (Dudkina et al., 2005), this structure induces curvature of the inner mitochondrial membrane and governs cristae formation (Arnold et $al.,$ 1998; Davies et al., 2012).

1.1.2. FtsZ and bacterial cell division

A division of prokaryotic organisms is based on the Z-ring formation made up of the FtsZ protein and located in the inner site of the bacterial membrane (Margolin, 2005). Bacterial septation is a complex process, and dozens of essential and accessory proteins participate to assemble the division machinery, the so-called divisome complex. Most bacterial cells divide by binary fission, which is the splitting of a cell into two (roughly) equal progeny cells.

FtsZ is the key cytoskeletal protein in the bacterial cytokinesis machine. It forms the Z-ring under the membrane at the centre of the cell, and this Z-ring constricts to initiate cell division. In addition to FtsZ, about a dozen accessory proteins are essential for this critical mechanism in *E. coli* (reviewed by Dajkovic and Lutkenhaus, 2006). FtsZ is highly conserved across almost all bacterial and archaeal species (Bi and Lutkenhaus, 1991). An important clue regarding the role of FtsZ came from the comparison of its predicted amino acid sequence witheukaryotic tubulins, which show limited but significant homology (de Boer *et al.*, 1992; Lutkenhaus, 1993). This homology lies in regions that are conserved in each of the two protein families, most strikingly a glycine-rich loop having the sequence GGGTGTG in FtsZ versus GGGTGSG in tubulins. Based on these similarities a model has been proposed for the structure of the region around the GTP-binding motif (de Pereda, 1996). Its function lies in the polymerization on the inside of the cytoplasmic membrane in the central part of the cell, at the site of the future septum, where it forms an annular body (Bi and Lutkenhaus 1991). *In vitro*, FtsZ assembles into short, one-stranded protofilaments, averaging 30 subunits 125 nm in length (Chen and Erickson, 2005). In the bacterial cell, these are further assembled into a long, thin filamentous structure attached to the inner bacterial membrane. Normally, the filament forms a single Z-ring at the centre of the cell (Peters *et al.*, 2007; Thanedar and Margolin, 2004).

In order to properly divide one cell into two equivalent daughter cells, the FtsZ protein has to polymerase only in the central part of the cell and not at the poles. An exception is splitting during sporulation when polar polymerization is desirable (reviewed by Dajkovic *et al.*, 2008). Two additional proteins, FtsA and ZipA, are both essential for cell division and work together to anchor the Z-ring to the cytoplasmic membrane, bind to the carboxy-terminal tail of FtsZ. Moreover, they are required for the maturation of the Z-ring, including recruitment of downstream cell division proteins. ZipA also promotes bundling of FtsZ protofilaments *in vitro* (Herricks *et al., 2014)*. Spatial limitation of the FtsZ polymerization

in the middle of the cell is ensured by several Min proteins that prevent ring formation at the cell poles, and through the so-called night system that limits the ring structure to segregated daughter chromosomes. By combining the effects of these two systems, FtsZ polymerization is confined to the centre of the cell. The functional unit in this system in *E. coli* is the MinC protein, which interferes with the interaction of FtsZ proteins and prevents their polymerization (Dajkovic *et al.*, 2008).

1.1.3. Mitochondrial division

Almost all prokaryotes including cyanobacteria and α-proteobacteria divide by using the multimeric GTPase protein FtsZ (Arimura and Tsutsumi, 2002), while plastids carry homologs of the cyanobacterial-type FtsZ protein (TerBush *et al.*, 2013). Moreover, some members of amoebozoans (Gilson *et al.,* 2003), stramenopiles (Beech *et al.,* 2000), and the red alga *Cyanidioschyzon merolae* (Takahara *et al.,* 2000) encode mitochondrial FtsZ. Interestingly, animals and fungi (opisthokonts) and plants examined to date lack this protein (Bleazard *et al.,* 1999). The plant *Arabidopsis thaliana* (Arimura and Tsutsumi, 2002), the parabasalid *Trichomonas vaginalis* (Wexler-Cohen *et al.,* 2014), *Dictyostelium discoideum* (Wienke *et al.,* 1999), and *C. merolae* (Nishida *et al.*, 2003) has been found to contain the dynamin-related protein (Dnm1p/Drp1), which participates in the mitochondrial division in most eukaryotes (Leger *et al.,* 2015). Mitochondrial fission and fusion processes in those organisms are mediated by large guanosine triphosphatases (GTPases) in the dynamin family that are conserved between yeast, flies, and mammals and their machineries are regulated by proteolysis and posttranslational modifications (reviewed by Youle *et al.*, 2012).

The first mitochondrial fission protein to be identified and studied in the yeast *Saccharomyces cerevisiae* was Dnm1. This cytosolic protein wraps around and constricts the site of dividing mitochondria by its GTPase activity. It interacts with the adaptor protein Mdv1 and its paralog Caf4, which further interact with Fis1, a small mitochondrial outer membrane protein (Bui and Shaw, 2013). Human and animal mitochondrial fissions also are performed by Dnm1, homologous to Fis1 and Drp1, the latter of which contains regions homologous to the amino-terminal and carboxyl-terminal GTPase domains of dynamin (Scott and Youle, 2010). In yeast, adaptors between Dnm1 and Fis1 are present, orthologs of Mdv1/Caf4, but they are rare or absent in other eukaryotes. Fis1 homologs have major roles in the relocation of Drp1 from the cytosol to the mitochondrial fission sites. This interesting

protein forms a collar that progressively tightens around the mitochondrion, and this constriction leads to severing of the outer mitochondrial membrane and fission into two organelles (Otera *et al.*, 2010; Palmer *et al.*, 2011; Zhao *et al.*, 2011). In humans has been found another protein, which is involved in the regulation of fission, Mitochondrial fission protein (Mff). Like Fis1, Mff is anchored at the membrane by a C-terminal transmembrane domain and loss of mitochondrial fission factor blocks mitochondrial fission (Scott and Youle, 2010; Gandre-Babbe and van der Bliek, 2008).

Trypanosomes and related flagellates have only one identified mitochondrial fission dynamin like protein (TbDLP) that has probably a role in endocytosis, cytokinesis, and mitochondrial division, but nothing is known about the mechanism and/or regulation. There are two paralogs in *Trypanosoma brucei* called TbDLP1 and TbDLP2. Overexpression of TbDLP1 is able to rescue endocytosis and growth defects in bloodstream form of the parasite, while in the insect-dwelling procyclic form, both TbDLP proteins are indispensable (Benz *et al.,* 2017).

1.2. *Trypanosoma brucei*

1.2.1. General features of the model flagellate *Trypanosoma brucei*

T. brucei is a unicellular flagellated parasite discovered by Sir David Bruce in the blood of a cow in 1894 (Joubert *et al.*, 1993). Trypanosomes are monophyletic and it belongs to the genus *Trypanosoma*, group Kinetoplastea and supergroup Euglenozoa (Fig. 3) (Lukeš *et al*., 1997; Lukeš *et al.*, 2014). Phylogenetic relationships in the group have been maped based on glycosomal glyceraldehyde 3-phosphate dehydrogenase gene (Hamilton *et al.*, 2004) and, more importantly, on the nuclear small subunit (SSU, 18S) rRNA, that has been extensively used to analyse the phylogenetic relationship inside this group (Lukeš *et al.*, 1997; Kostygov *et al*., 2021; Simpson *et al.*, 2002).
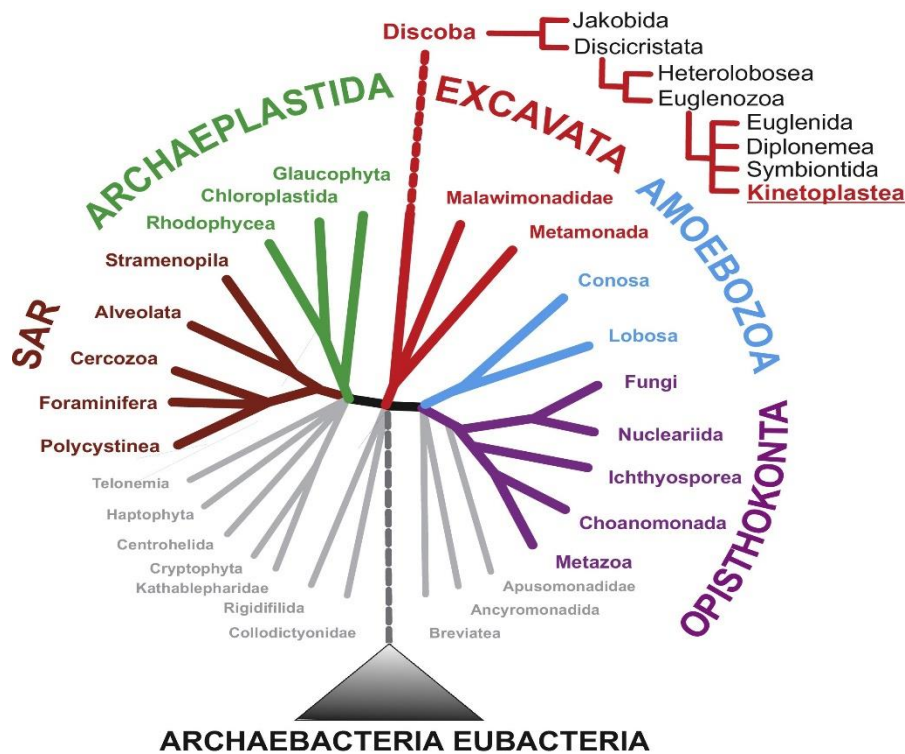
Figure 3: Phylogenetic position of *Kinetoplastea* within the tree of life (Lukeš *et al.*, 2014).

T. brucei in the causative agent of lethal sleeping sickness of human and livestock, which afflict the region of sub-Saharan Africa, with most affected people living in remote areas without access to quality healthcare (Lukeš *et al*., 2022). The most well-known diseases caused by trypanosomatids are Human African Trypanosomiasis (HAT, also known as sleeping sickness), Chagas disease, which affect people and a wide range of mammals mostly in central and South America and leishmaniases, inflicting humans in most tropical countries (Kennedy, 2019).

T. brucei earned its place among model organisms by its amenability to genetic methods and also its easy cultivation (Serricchio and Bütikofer, 2011). As a matter of fact, it is only the second organism on which new method for inhibition of gene expression - RNA interference was demonstrated (Ngo *et al.*, 1998). This method of post-transcriptional gene regulation is faster and easier than gene knock-outs and allows efficient functional analysis of especially essential genes (Wang *et al.*, 2000; Djikeng *et al.*, 2001). Other breakthrough discoveries concerning general biological processes associated with *T. brucei* are antigenic variation (Cross, 1977), *trans*-splicing (Sutton and Boothroyd, 1984), glycosyl-phosphatidylinositol anchoring (Ferguson *et al.*, 1988) and RNA editing (Blum *et al.*, 1990), and recently also

include stabilization of messenger RNAs (Viegas *et al.*, 2022). Indeed, all these processes were initially discovered in *T. brucei*, and only later were also found in a range of eukaryotes (Serricchio and Bütikofer, 2011).

1.2.2. Life cycle of *T. brucei*

T. brucei has a complex life cycle, which is caused by its dixenous nature. The transmission of this parasite between mammalian hosts is mediated by an insect vector – tsetse fly *Glossina* spp. (Fig. 4) (reviewed by Dyer *et al.*, 2013; Fenn and Matthews, 2007). The infected mammalian host is bitten by an insect, in the course of which the bloodstream stumpy form of *T. brucei* is taken up into the midgut, where the short stumpy forms differentiate into the proliferative procyclic trypomastigotes (PF). After establishing the midgut infection, the parasite relocates through the peritrophic membrane to the proventriculus. Here starts asymmetric division which generates two different cell type – the short and the long epimastigotes. Short epimastigote form eventually migrates into the salivary glands, where it attaches to the epithelium and starts new asymmetric cell division to generated the metacyclic form, which is no longer attached to the epithelium and is preadapted to survive in the mammalian host. Upon transmission, via blood feeding of the tsetse flies into a new host, the metacyclic form transforms into the proliferative long slender bloodstream form (BF), which replicates until the cells reach a certain density in this host (reviewed by Langousis and Hill, 2014).
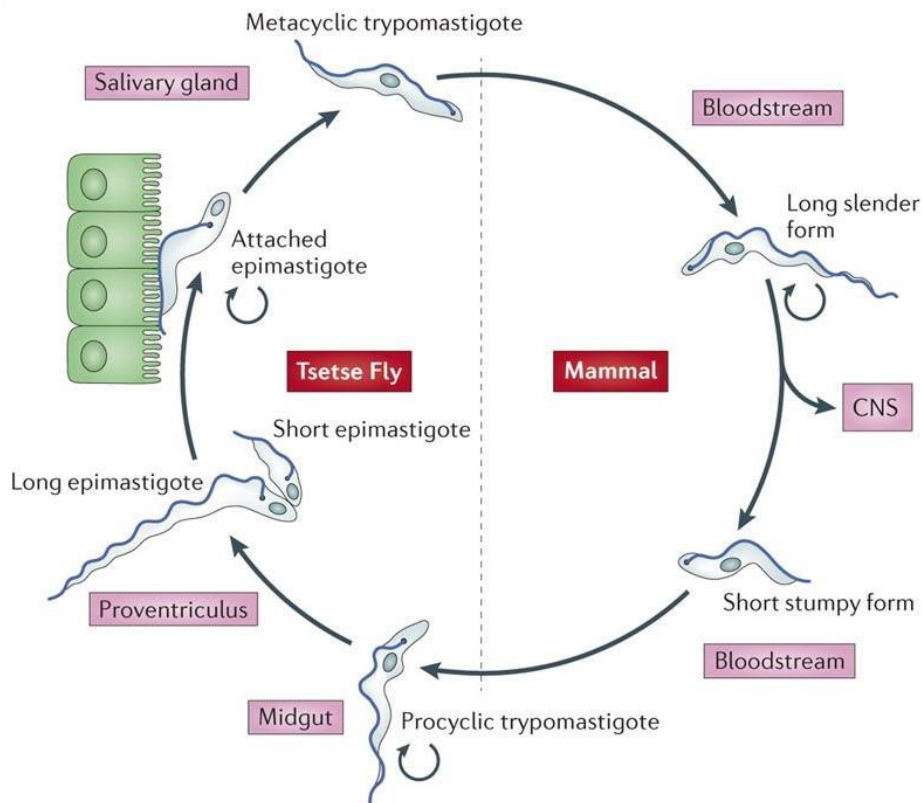
Figure 4: *T. brucei* life cycle (Langousis and Hill, 2014)

Although, *T. brucei* has several different life cycle stages, only two of them are easily cultivable under laboratory conditions, namely PF and BF. Therefore, most of the knowledge about this parasite is coming from these two life stages. Nonetheless, the information obtained is more than striking and intriguing as the PF and BF are extraordinarily different.

In case of the PF from the insect vector, its mitochondrion exists as a single organelle interlacing throughout the whole cell. The inner membrane is wrinkled into quite prominent cristae. Due to the lack of glucose in the insect host, the parasite uses amino acids (L-proline and L-threonine) as their principal carbon source (Smith *et al*., 2017).

The situation is strikingly different in the BF it has a morphologically reduced mitochondrion compared to that of the PF. Their mitochondrion assumes a tubular structure stretching only from the anterior to the posterior end of the cell (Barret *et al*., 2003; Jakob *et al*., 2016). Mitochondrion has only few extensions and loops (Vickerman, 1985). A recent study (Bílý *et al*, 2021) disproved the long-held view that cristae are practically absent in BF mitochondria (Vickerman, 1985). Surprisingly, cristae in BF occupied up to 15% of the organelle's total surface area (Bílý *et al*, 2021). This striking ultrastructural remodelling is accompanied by the equally dramatic changes in the mitochondrial metabolism. Since the blood of mammals is full of freely available glucose, the parasite utilizes it as its carbon source. The

ETC complexes III and IV are strongly downregulated. As a substitution of the respiration, which was in PF provided by complex IV, trypanosomes BF have upregulated expression of a plant-like alternative oxidase (TAO) (Smith *et al*., 2017).

1.3. Kinetoplast DNA

T. brucei and other trypanosomatids exhibit several unusual biological properties. One of the most striking features is the single, reticulated mitochondrion with unusual genome (Jensen and Englund, 2012). While the majority of α-proteobacterial ancestor-derived genes have been transferred from the mitochondrial genome to the nuclear genome. In *T. brucei*, the mitochondrial DNA, subsequently referred to as kinetoplast DNA (kDNA), is represented by a single network of maxicircles and minicircles (see below) (Verner *et al*., 2015). It is located as a densely packed disk-like structure at the posterior end of the mitochondrion, close to the basal body of the flagellum (Fig. 5) (Liu *et al*., 2005).

In fact, the kDNA is physically linked to the basal body of the single flagellum via a structure termed tripartite attachment complex (TAC) (Zhao *et al*., 2008). The TAC involves a set of unilateral filaments, located within the mitochondrial matrix, linking the basal bodies to a zone of differentiated outer and inner mitochondrial, the differentiated membrane, and the exclusion zone filaments in the cytosol (Ogbadoyi *et al.*, 2003). During the cell cycle, kDNA division follows the formation of a new flagellum (Woodward and Gull 1990). *T. brucei* strains that lost part of their kDNA are termed dyskinetoplastic (*Trypanosoma equiperdum*), while those which did not retain any kDNA are labelled as akinetoplastic (*Trypanosoma evansi*) (Lai *et al.*, 2008; Schnaufer *et al*., 2002).
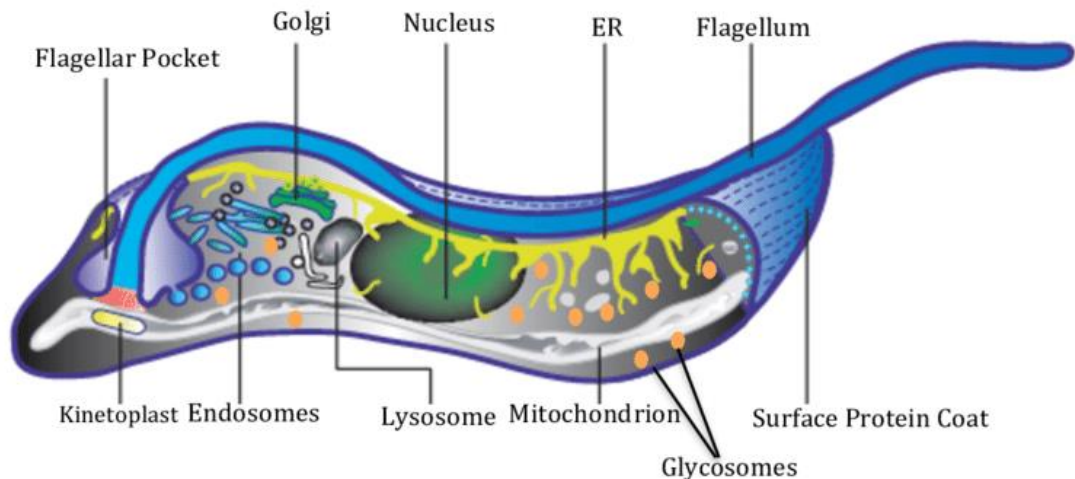
Figure 5: Schematic representation of the cellular organization of *T. brucei* (Dodson *et al*., 2011).

The kDNA is arranged into a single catenated network that consists of two types of DNA rings called minicircles and maxicircles. There are approximately 25 maxicircles, each of which is 23 kilobases (kb) in size, encoding two ribosomal RNA genes and 18 protein-coding genes, namely several subunits of cytochrome *c* oxidase (COI, COII, COIII), cytochrome *c* reductase (apocytochrome b), F_1F_0 ATPase (A6), and NADH dehydrogenase, as well as few unassigned reading frames, while tRNA genes are prominently absent (de la Cruz *et al*., 1984; Simpson *et al*., 1986). Most of the maxicircle-encoded transcripts require (extensive) RNA editing in the form of post-transcriptional additions and deletions of uridine residues (Hajduk and Ochsenreiter, 2010; Jensen and Englund, 2012; Povelones, 2014). Only properly edited transcripts are translatable on the mitochondrial ribosomes into functional proteins (Horvath et al., 2000). RNA editing is an extremely complex mechanism that requires hundreds of small RNA molecules called guide (g) RNAs, almost all of which are encoded on the minicircles, and dozens of dedicated nuclear-encoded proteins (Aphasizheva *et al*., 2021).

The minicircles are much more abundant than the maxicircles and make up the majority of the kDNA network. Several thousands of minicircles are each approximately 1.0 kb in size, highly heterogeneous in sequence, although their origin of replication is conserved (Ntambi and Englund, 1985). Each minicircle is catenated with three neighbouring minicircles, while maxicircles are probably linked with each other, constituting a network within the minicircle network (Fig. 6) (Chen *et al*., 1995). These mutually interlocked DNA circles are highly condensed into a disk-like structure (Shapiro, 1993). The replication of this complex kDNA network is a complex and only partially understood process, which may require the function of up to 150 proteins (Jensen and Englund, 2012; Verner *et al*., 2015).
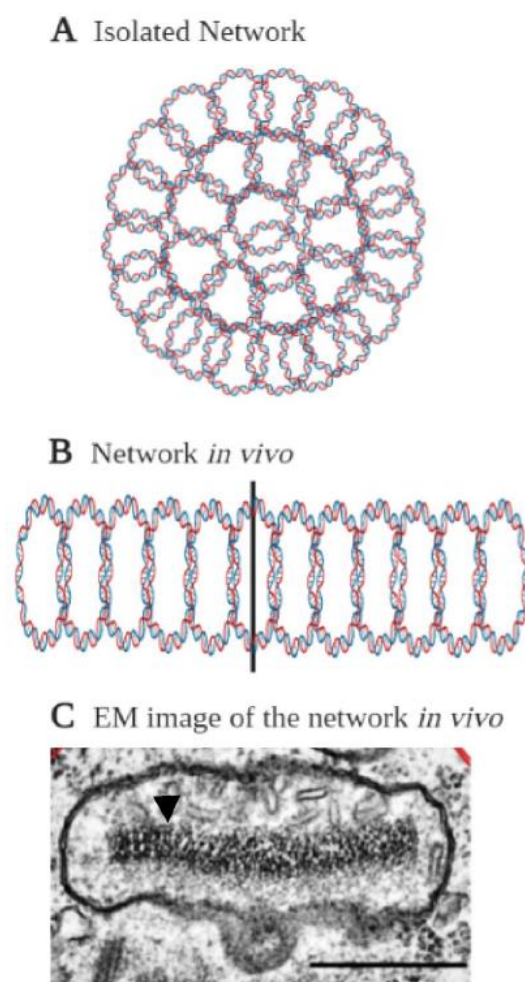
Figure 6: Organization of the kDNA network. A) Diagram of an isolated, uncondensed kDNA network; B) Diagram of a section of kDNA network, condensed in vivo into a disk-shaped structure. The vertical bar represents the axis of the disc. C) An electron micrograph of a thin section of the *T. brucei* mitochondrion. The black arrowhead marks the condensed kDNA disk within the mitochondrial matrix. Bar, 500 nm (from Yaffe *et al*., 2021).

The most recent findings show that the replication of kDNA starts with the release individual covalently closed minicircles into the region between the kDNA disk and the inner mitochondrial membrane, termed the kinetoflagellar zone (KFZ) (Englund, 1979; Drew and Englund, 2001). Subsequently, the replication continues with the release of the individual minicircles from the network into the KFZ by a topoisomerase II. Here DNA primase, two DNA polymerases (Klingbeil *et al.*, 2002), and universal minicircle sequence-binding protein (Abu-Elneel *et al.*, 2001) perform replication of the free minicircles (Sela and Shlomai, 2009). Replicated minicircles then migrate to the antipodal sites, which are protein complexes at opposing sites of the kDNA disc (Fig. 7). At these antipodal sites, most gaps between Okazaki fragments are repaired and minicircles are attached to the periphery of the kDNA network (Jensen and Englund, 2012; Melendy *et al.*, 1988; Ryan and Englund, 1989a; Ryan and Englund, 1989b). In contrast to minicircles, maxicircles are never released from the kDNA network and are replicated while being interlocked with minicircles and other maxicircles. This makes it harder to study the exact mechanism of maxicircle replication (Jensen and Englund, 2012; Drew and Englund, 2001; Carpenter and Englund, 1995). There is only one known protein, Pif2 helicase, which affects exclusively maxicircle replication (Liu *et al.*, 2009). Subsequently, this kinetoplast segregation process is mediated by movement apart of the flagellar basal bodies that are connected to the kDNA *via* the TAC structure (Ogbadoyi, 2003). The last step in kDNA separation is the cleavage of maxicircles that still connect the segregating networks (Jensen and Englund, 2012; Povelones, 2014; Gluenz *et al.*, 2011).
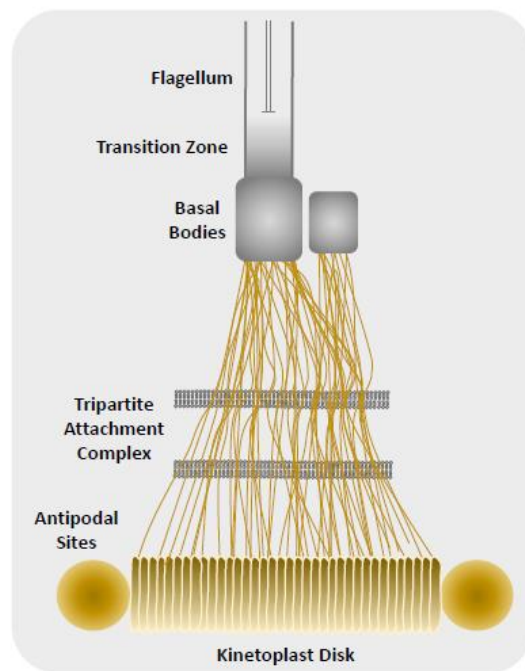
Figure 7: Schematized depiction of distinguished sub-compartments within the kinetoplast/TAC area (Pyrih *et al.*, under review).

1.4. Mitochondrial protein import

Mitochondrial protein import is executed by several protein complexes, which were thought to be highly conserved. Protein import has best been studied in *Saccharomyces cerevisiae* and much of this work can, in principle, be generalized to most eukaryotic lineages. Most mitochondrial proteins are synthesized on cytosolic ribosomes and post-translationally transported into the organelle (Pfanner *et al*., 2019; Hansen and Herrmann, 2019). The transport process is undertaken by the mitochondrial translocation machinery (Koehler, 2000).

For most of the mitochondrial proteome, the canonical import pathway involves an amphipathic α-helical N-terminal sequence that directs import through the Translocase of the Outer Membrane (TOM) complex pore directing it to the TIM17:23 inner membrane complex (Neupert and Herrmann, 2007). After entering the TOM complex, preproteins follow one of five mitochondrial import pathways, depending on their structure, function, and target destination (Fig. 8) (Kutik *et al.,* 2007; Schatz and Dobberstein, 1996).
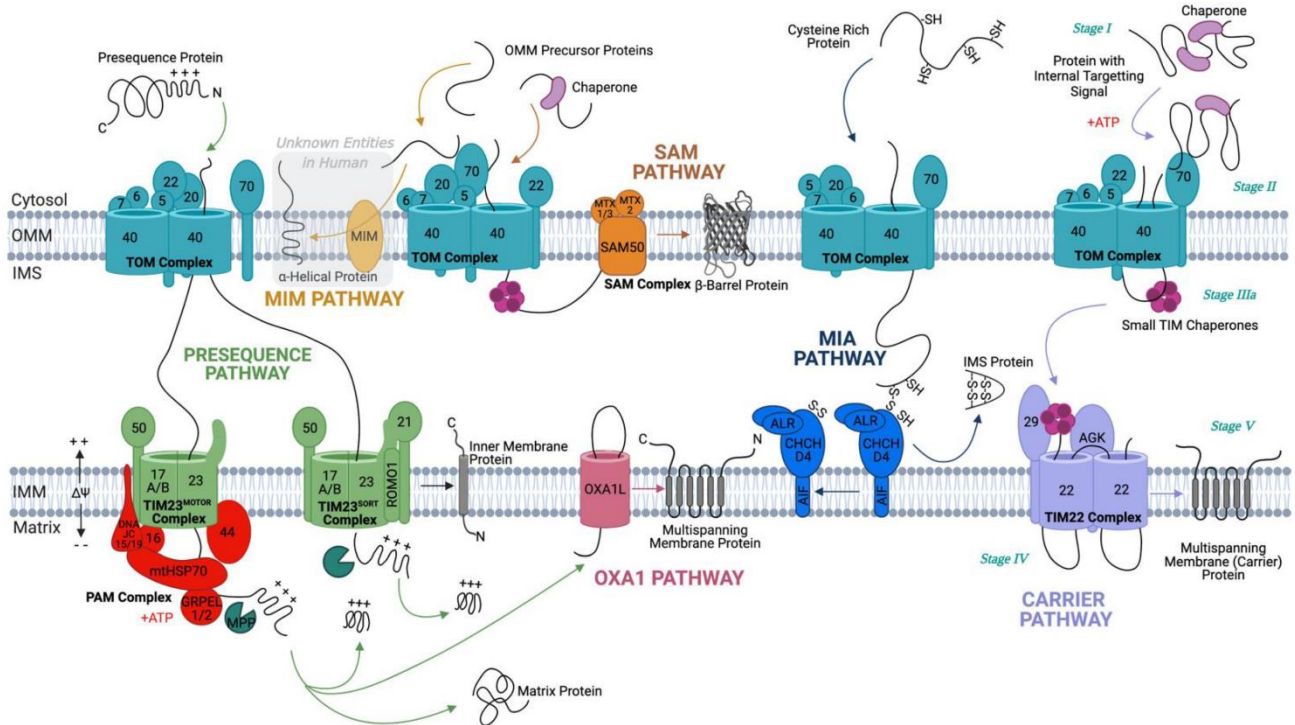
Figure 8: Overview of mitochondrial protein import pathways (Needs *et al*., 2021)

In *S. cerevisiae* the TOM complex comprises seven subunits, of which three are conserved across all eukaryotic lineages and can be traced to LECA (Mani *et al*., 2016; Maćasev *et al*., 2004; Mani *et al*., 2017). Its core five components (TOM40, TOM22, TOM7, TOM6, and TOM5) are further associated with additional TOM20 and TOM70 subunits (Chacinska *et al*., 2009; Tucker and Park, 2019). Precursor proteins containing an N-terminal presequence are passed directly from the TOM complex to the TIM23 complex (Harbauer *et al*., 2014; Chacinska *et al*., 2009). The N-terminal presequence is a cleavable region of 15 to 50 amino acids that precedes the mature protein, and which is rich in hydrophobic, hydroxylated, and basic residues, with an overrepresentation of arginine residues and a near absence of acidic residues, forming a positively charged, amphipathic α-helix (Neupert, 1997).

The yeast mitochondrial inner membrane has two heterooligomeric protein complexes with non-overlapping subunit composition, termed translocase of the inner membrane 23 (TIM23) and TIM22 (Harsman and Schneider, 2017; Wiedemann and Pfanner, 2017). The TIM22 complex mediates the insertion of mitochondrial carrier proteins into the inner mitochondrial membrane (Ferramosca and Zara, 2013). The TIM23 complex imports presequence-containing precursor proteins, which represent around 70% of all mitochondrial proteins, across or into the inner mitochondrial membrane (Hansen and Herrmann, 2019). The TIM23 complex is anchored to the inner mitochondrial membrane, and is composed of two sub-complexes. First is an integral membrane embedded core complex which in yeast contains three essential subunits (Tim17, Tim23, and Tim50) (Yamamoto *et al.*, 2002; Maarse *et al.*, 1994; Dekker *et al.*, 1993; Geissler *et al.*, 2002). Second one is the Presequence Assisted Motor (PAM) complex that powers the translocation of preproteins through the inner membrane channel (Neupert and Herrmann, 2007).

The core of the PAM complex is formed by the 70 kDa mitochondrial heat shock protein (mtHsp70) that drives the translocation and the unfolding of the preprotein by an ATP-dependent reaction. mtHsp70 is essential for cell viability, however four others essential co-chaperones are needed for the motor function of PAM (Voos *et al.*, 1999; Herrmann and Neupert, 2000). Mge1 catalyses the exchange of ADP for ATP allowing release and reload of the substrate (Liu *et al.*, 2003), while Pam18 stimulates the ATPase activity of mtHsp70 (Truscott *et al.*, 2003) and is connected to the PAM complex via Pam16 (Pais *et al.*, 2011). Finally, Pam17 stabilizes the whole complex (Sanjuán Szklarz *et al.*, 2005; van der Laan *et al.*, 2005).

Pam18 and Pam16 belong to the belong to the DnaJ and DnaJ-like families of proteins, respectively (Hennessy *et al.*, 2005). J-domain proteins contain a canonical tripeptide HPD motif of histidine, proline, and aspartate residues essential for stimulating the ATPase activity of Hsp70 (Truscott et al., 2003; Hennessy et al., 2005). Both Pam16 and Pam18 are anchored to the inner membrane *via* an N-terminal transmembrane domain (Mokranjac *et al.*, 2006) and are essential for yeast viability (Frazier *et al.*, 2003; Truscott *et al.*, 2003).

T. brucei has only a single TIM complex that can import presequence-containing proteins as well as mitochondrial carrier proteins associated with the PAM module. Furthermore, *T. brucei* contains orthologues of Pam18 and Pam16, termed TbPam18 and TbPam16, that are essential for normal growth and are not involved in mitochondrial protein import. During evolution of the trypanosomatid flagellates, Pam18 and Pam16 must have been replaced in the import motor by TbPam27 (Fig. 9) (von Känel *et al.*, 2020).
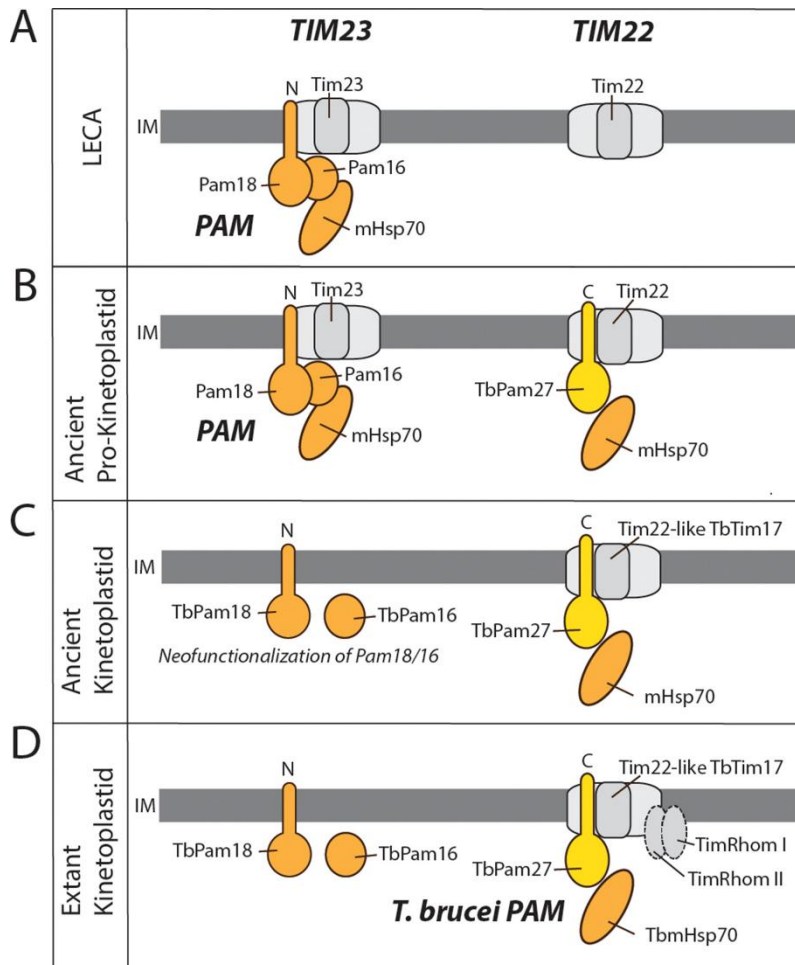
Figure 9: Evolutionary scenario explaining the homologue replacement observed in the *T. brucei* PAM. A) Ancestral situation predicted for LECA. B) Occurrence of the J domain-containing protein TbPam27 that may allow interaction with mHsp70. C) The TIM22 complex of the ancient kinetoplastid acquires the capability to translocate presequence-containing proteins. D) Disappearance of the trypanosomal TIM23 complex. TbPam18 and TbPam16 were retained because they acquired a new yet unknown function (taken von Känel *et al.*, 2020).

1.4.1. ZapE/Afg1/LACE

ZapE is a newly identified protein (alternatively Afg1 in yeast or LACE1 in humans) that is part of the FtsZ division machinery and occurs in the inner membrane of bacteria and mitochondria. Under normal growth condition is ZapE in *E. coli* non-essential. However, under increased temperature conditions, low-oxygen or ZapE overexpression, long cells are produced that are unable to divide. This elongated cell phenotype can be reversed by a single point mutation (K84A) in the overexpressed ZapE protein, which abolishes its ATPase activity. It is therefore plausible that the ATPase activity is directly linked to the observed division-lacking phenotype (Marteyn *et al.*, 2014). Afg1 (ATPase family gene 1) is a yeast mitochondrial ATPase, evolutionarily highly conserved protein with a mammalian homologue LACE1 (lactation elevated 1). Afg1 homologues ZapE is also a member of the AAA + ATPase family (ATPase associated with a variety of cellular activities) and contains an ATP/GTP binding P-loop (Marteyn *et al.*, 2014; Abrahams *et al.,* 2002; Saraste *et al.*, 1990).

ZapE orthologs mediate degradation of the mitochondrial-encoded subunits (COX1, COX2 and COX3) of the respiratory complex IV, however, the function of human and yeast homologs may be different, as it was proposed (Khalimonchuk *et al.,* 2007; Cesnekova *et al.*, 2016a). Furthermore, both organisms lack the FtsZ division system (Leger *et al.,* 2015). Yeast Afg1 deletion strain reduced activities of respiratory chain complexes III and IV. Loss of LACE1 leads to increased apoptotic resistance, whereas its overexpression results in increased apoptotic sensitivity. Moreover, LACE1 interact with p53 and was shown to mediate its translocation (Cesnekova *et al.*, 2016b). It is also noteworthy that ZapE was highly affected in the proteomic survey of the Oxa1 depletome in humans (Stiller *et al.,* 2016). Moreover, the functional link between Oxa1 and ZapE was also suggested in yeast, where only the mitochondrial-encoded subunits of respiratory complex IV were affected following the depletion of the latter protein (Germany *et al.*, 2018). The mechanism by which ZapE impacts the cell division pathway through its direct interaction with FtsZ is largely unknown (Marteyn *et al.*, 2014).

All in all, the phenotypes associated with ZapE in different organisms vary widely, and so far, they have not been integrated into a coherent picture. To shed light on the function(s) of the conserved ZapE protein, so far examined only in bacteria and opisthokonts, we have probed the function and interactions of its two paralogs, ZapE1 (XP_823041.1, Tb927.7.6930) and ZapE2 (XP_846313.1; Tb927.10.8070). According to the ATOM40 depletome-based

mitochondrial proteome (Peikert *et al.*, 2017) and TrypTag *in situ* tagging database (Dean *et al.,* 2015), both proteins are genuine components of the *T. brucei* mitochondrion.

2. Objectives

1. **Study of the functions of ZapE in *T. brucei*.**

    -   Phylogenetic analysis to determine the evolutionary origin of ZapE in eukaryotes.
    -   Identification of interacting proteins using BioID2.

2. **Analysis of the presence of homologs of the bacterial Ffh and FtsY proteins in various unrelated plastid-lacking unicellular eukaryotes.**

    -   Phylogenetic analysis of the origin of mtFfh and mtFtsY.
    -   Evaluating a set of putative mitochondrial proteins in the heterolobosean *N. gruberi*.

3. **Investigation of the role of TbPam18 and TbPam16 in *T. brucei*.**

    -   Identification and analysis of the associated function of TbPam18 and TbPam16.
    -   Determining how depletion of TbPam18 and TbPam16 affects the replication of kDNA, the activity of respiratory complexes and other mitochondrial functions.

4. **Novel protein complex involved in kinetoplast DNA replication and maintenance.**

    -   Screening the TrypTag localization repository and prioritizing 10 previously undescribed putative proteins displaying kinetoplast proximal enrichment (KEP).
    -   Providing a methodological pipeline for the identification of novel KP associated targets.

3. References

Abrahams B.S., Mak G.M., Berry M.L., Palmquist D.L., Saionz J.R., Tay A., Tan Y.H., Brenner S., Simpson E.M., Venkatesh B. (2002) Novel vertebrate genes and putative regulatory elements identified at kidney disease and NR2E1/fierce loci. *Genomics.* **80**: 45–53.

Abu-Elneel K., Robinson D.R., Drew M.E., Englund P.T., Shlomai J. (2001) Intramitochondrial localization of universal minicircle sequence-binding protein, a trypanosomatid protein that binds kinetoplast minicircle replication origin. *J Cell Biol*. **153**: 725–733.

Allen J.F. (2015) Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. *Proc Natl Acad Sci U S A*. **33:** 10231–8.

Arimura S. (2018) Fission and fFusion of pPlant mMitochondria, and gGenome mMaintenance. *Plant Physiol*. **176**: 152–161.

Arimura S., Tsutsumi N. (2002) A dynamin-like protein (ADL2b), rather than FtsZ, is involved in Arabidopsis mitochondrial division. *Proc Natl Acad Sci U S A*. **99**: 5727–5731.

Arnold I., Pfeiffer K., Neupert W., Stuart R.A., Schagger H. (1998) Yeast mitochondrial F_1F_o-ATP synthase exists as a dimer: identification of three dimer-specific subunits. *EMBO J*. **17**: 7170–7178.

Barrett M.P., Burchmore R.J., Stich A., Lazzari J.O., Frasch A.C., Cazzulo J.J., Krishna S. (2003) The trypanosomiases. *Lancet*. **362**: 1469–1480.

Beech P.L., Nheu T., Schultz T., Herbert S., Lithgow T., Gilson P.R., McFadden G.I. (2000) Mitochondrial FtsZ in a chromophyte alga. *Science*. **287**: 1276–9.

Benz C., Stříbrná E., Hashimi H., Lukeš J. (2017) Dynamin-like proteins in *Trypanosoma brucei*: A division of labour between two paralogs? *PLoS One*. **12**: e0177200.

Berrisford J.M., Sazanov L.A. (2009) Structural bBasis for the mMechanism of rRespiratory cComplex I. *J Biol Chem*. **284**: 29773–83.

Bi E.F., Lutkenhaus J. (1991) FtsZ ring structure associated with division in *Escherichia coli*. *Nature*. **354**: 161–164.

Bienen E.J., Saric M., Pollakis G., Grady R.W., Clarkson A.B.J. (1991) Mitochondrial development in *Trypanosoma brucei brucei* transitional bloodstream forms. *Mol Biochem Parasitol*. **45**: 185–192.

Bílý T., Sheikh S., Mallet A., Bastin P., Pérez-Morga D., Lukeš J., Hashimi H. (2021) Ultrastructural Changes of the Mitochondrion During the Life Cycle of *Trypanosoma brucei*. *J Eukaryot Microbiol*. **68**: e12846.

Bleazard W., McCaffery J.M., King E.J., Bale S., Mozdy A., Tieu Q., Nunnari J., Shaw J.M. (1999) The dynamin-related GTPase Dnm1 regulates mitochondrial fission in yeast. *Nat Cell Biol*. **1**: 298–304.

Blum B., Bakalara N., Simpson L. (1990) A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell*. **60**: 189–98.

Bohnert M., Rehling P., Guiard B., Herrmann J.M., Pfanner N., van der Laan M. (2010) Cooperation of stop-transfer and conservative sorting mechanisms in mitochondrial protein transport. *Curr. Biol*. **20**: 1227–1232.

Brzezinski P., Moe A., Ädelroth P. (2021) Structure and Mechanism of Respiratory III-IV Supercomplexes in Bioenergetic Membranes. *Chem Rev*. **121**: 9644–9673.

Bui H.T., Shaw J.M. (2013) Dynamin assembly strategies and adaptor proteins in mitochondrial fission. *Curr Biol*. **23**: 891–9.

Carpenter L.R., Englund P.T. (1995) Kinetoplast maxicircle DNA replication in *Crithidia fasciculata* and *Trypanosoma brucei*. *Mol Cell Biol.* **15**: 6794–6803.

Cesnekova J., Rodinova M., Hansikova H., Houstek J., Zeman J., Stiburek L. (2016a) The mammalian homologue of yeast Afg1 ATPase (lactation elevated 1) mediates degradation of nuclear-encoded complex IV subunits. *Biochem J*. **473**: 797–804.

Cesnekova J., Spacilova J., Hansikova H., Houstek J., Zeman J., Stiburek L. (2016b) LACE1 interacts with p53 and mediates its mitochondrial translocation and apoptosis. *Oncotarge*t. **7**: 47687–47698.

Chacinska A., Koehler C.M., Milenkovic D., Lithgow T., Pfanner N. (2009) Importing mitochondrial proteins: machineries and mechanisms. *Cell*. **138**: 628–644.

Chen J., Rauch C.A., White J.H., Englund P.T., Cozzarelli, N.R. (1995) The topology of thekinetoplast DNA network. *Cell.* **80**: 61–69.

Chen Y., Erickson H.P. (2005) Rapid in vitro assembly dynamics and subunit turnover of FtsZ demonstrated by fluorescence resonance energy transfer. *J Biol Chem*. **280**: 22549–54.

Cooper S., Wadsworth E.S., Ochsenreiter T., Ivens A., Savill N.J, Schnaufer A. (2019) Assembly and annotation of the mitochondrial minicircle genome of adifferentiation-competent strain of *Trypanosoma brucei*. *Nucleic Acids Res.* **47**: 11304–11325.

Cotter P.D.; Hill C. (2003) Surviving the acid test: responses of gram-positive bacteria to low pH. *Microbiol Mol Biol Rev*. **67**: 429–453.

Cramer W.A., Hasan S.S., Yamashita E. (2011) The Q cycle of cytochrome bc complexes: a structure perspective. *Biochim Biophys Acta*. **1807**: 788–802.

Cross G.A. (1977) Antigenic variation in trypanosomes. *Am J Trop Med Hyg*. **26**: 240–4.

Dajkovic A., Lan G., Sun S.X., Wirtz D., Lutkenhaus J. (2008) MinC sSpatially cControls bBacterial cCytokinesis by aAntagonizing the sScaffolding fFunction of FtsZ. *Current Biology*. **18**: 235–244.

Dajkovic A., Lutkenhaus J. (2006) Z ring as executor of bacterial cell division. *J Mol Microbiol Biotechnol*. **11**: 140–51.

Davies K.M., Anselmi C., Wittig I., Faraldo-Gomez J.D., Kuhlbrandt W. (2012) Structure of the yeast F_1F_o-ATP synthase dimer and its role in shaping the mitochondrial cristae. *Proc Natl Acad Sci U S A*. **109**: 13602–13607.

Davies K.M., Daum K. (2013) Role of cryo-ET in membrane bioenergetics research. *Biochem Soc Trans*. **41**: 1227–1234.

de Boer P., Crossley R., Rothfield L. (1992) The essential bacterial cell-division protein FtsZ is a GTPase. *Nature*. **359**: 254–56.

de la Cruz V., Neckelmann N., Simpson L. (1984) Sequences of six structural genes and several open reading frames in the kinetoplast maxicircle DNA of *Leishmania tarentolae. J Biol Chem*. **259**: 15136–15147.

de Pereda J.M., Leynadier D., Evangelio J.A., Chacon P., Andreu J.M. (1996) Tubulin secondary structure analysis, limited proteolysis sites, and homology to ftsZ. *Biochemistry*. **35**: 14203–15.

Dean S., Sunter J., Wheeler R.J., Hodkinson I., Gluenz E., Gull K. (2015) A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol*. **5**: 140197.

Dekker P.J.T., Keil P., Rassow J., Maarse A.C., 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.

Djikeng A., Shi H., Tschudi C., Ullu E. (2001) RNA interference in *Trypanosoma brucei*: cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs. *RNA*. **7**: 1522–1530.

Dodson H.C., Morris M.T., Morris J.C. (2011) Glycerol 3-phosphate alters *Trypanosoma brucei* hexokinase activity in response to environmental change. *J Biol Chem*. **286**: 33150–7.

Drew M.E., Englund, P.T. (2001) Intramitochondrial location and dynamics of *Crithidia fasciculata* kinetoplast minicircle replication intermediates. *J Cell Biol.* **153**: 735– 744.

Dudkina N.V., Heinemeyer J., Keegstra W., Boekema E.J., Braun H.P. (2005) Structure of dimeric ATP synthase from mitochondria: an angular association of monomers induces the strong curvature of the inner membrane. *FEBS Lett*. **579**: 5769–5772.

Dyer N.A., Rose C., Ejeh N.O., Acosta-Serrano A. (2013) Flying tryps: survival and maturation of trypanosomes in tsetse flies. *Trends Parasitol*. **29**: 188–96.

Embley T.M., Martin W. (2006) Eukaryotic evolution, changes and challenges. *Nature*. **440**: 623–630.

Englund P.T. (1979) Free minicircles of kinetoplast DNA in *Crithidia fasciculata*. *J. Biol. Chem.* **254**: 4895–4900.

Fenn K., Matthews K.R. (2007) The cell biology of *Trypanosoma brucei* differentiation. *Current Curr Opin Microbiol*. **10**: 539–546.

Ferguson M.A., Homans S.W., Dwek R.A., Rademacher T.W. (1988) The glycosylphosphatidylinositol membrane anchor of *Trypanosoma brucei* variant surface glycoprotein. *Biochem Soc Trans*. **16**: 265–8.

Ferramosca A., Zara V. (2013) Biogenesis of mitochondrial carrier proteins: Molecular mechanisms of import into mitochondria. *Biochim et Biophys Acta.* **1833**: 494–502.

Frazier A. E., Chacinska A., Truscott K. N., Guiard B., Pfanner N., Rehling P. (2003) Mitochondria use different mechanisms for transport of multispanning membrane proteins through the intermembrane space. *Mol Cell Biol*. **23**: 7818–7828.

Funes S., Nargang F.E., Neupert W., Herrmann J.M. (2004) The Oxa2 protein of *Neurospora crassa* plays a critical role in the biogenesis of cytochrome oxidase and defines a ubiquitous subbranch of the Oxa1/YidC/Alb3 protein family. *Mol Biol Cell*. **15**: 1853–1861.

Gabaldón T., Huynen M.A. (2003) Reconstruction of the proto-mitochondrial metabolism. *Science*. **301**: 609.

Gandre-Babbe S., van der Bliek A.M. (2008) The novel tail-anchored membrane protein Mff controls mitochondrial and peroxisomal fission in mammalian cells. *Mol Biol Cell*. **19**: 2402–2412.

Geissler A., Chacinska A., Truscott K.N., Wiedemann N., Brandner K., Sickmann A., Meyer H.E., 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.

Germany E.M., Zahayko N., Huebsch M.L., Fox J.L., Prahlad V., Khalimonchuk O. (2018) The AAA ATPase Afg preserves mitochondrial fidelity and cellular health by maintaining mitochondrial matrix proteostasis. *J Cell Sci*. **131**: jcs219956.

Gilson P.R., Yu X.C., Hereld D., Barth C., Savage A., Kiefel B.R., Lay S., Fisher P.R., Margolin W., Beech P.L. (2003) Two Dictyostelium orthologs of the prokaryotic cell division protein FtsZ localize to mitochondria and are required for the maintenance of normal mitochondrial morphology. *Eukaryot Cell*. **2**: 1315–26.

Gluenz E., Povelones M.L., Englund P.T., Gull K. (2011) The kinetoplast duplication cycle in *Trypanosoma brucei* is orchestrated by cytoskeleton-mediated cell morphogenesis. *Mol Cell Biol.* **31:** 1012–1021.

Hajduk, S., Ochsenreiter, T. (2010). RNA editing in kinetoplastids. *RNA Biol*. **7:** 229–236.

Hamilton P.B., Stevens J.R., Gaunt M.W., Gidley J. and Gibson W.C. (2004) Trypanosomes are monophyletic: evidence from genes for glyceraldehyde phosphate dehydrogenase and small subunit ribosomal RNA. *Int. J. Parasitol.* **34:** 1393–404.

Hansen K.G., Herrmann J.M. (2019) Transport of proteins into mitochondria. *Protein J*. **38**: 330–342.

Haque M.E., Elmore K.B., Tripathy A., Koc H., Koc E.C., Spremulli L.L. (2010) Properties of the C-terminal Tail of Human Mitochondrial Inner Membrane Protein Oxa1L and Its Interactions with Mammalian Mitochondrial Ribosomes. *J Biol Chem*. **285**: 28353–28362.

Harbauer A.B., Zahedi R.P., Sickmann A., Pfanner N., Meisinger C. (2014) The Protein Import Machinery of Mitochondria-A Regulatory Hub in Metabolism, Stress, and Disease. *Cell Metab*. **19**: 357–372.

Harsman A., Schneider A. (2017) Mitochondrial protein import in trypanosomes: Expect the unexpected. *Traffic.***18**: 96–109.

Hederstedt L. (2003) Complex II is complex too. *Science*. **299**: 671–672.

Hell K., Neupert W., Stuart R.A. (2001) Oxa1p acts as a general membrane insertion machinery for proteins encoded by mitochondrial DNA. *EMBO J*. **20**: 1281–1288.

Hennessy F., Nicoll W. S., Zimmermann R., Cheetham M. E., Blatch G. L. (2005). Not all J domains are created equal: Implications for the specificity of Hsp40–Hsp70 interactions. *Protein Sci*. **14:** 1697–1709.

Herricks J.R., Nguyen D., Margolin W. (2014) A thermosensitive defect in the ATP binding pocket of FtsA can be suppressed by allosteric changes in the dimer interface. *Mol Microbiol.* **94**: 713–27.

Herrmann J.M., Neupert W. (2000) What fuels polypeptide translocation? An energetical view on mitochondrial protein sorting. *Biochim Biophys Acta*. **1459**: 331–338.

Herrmann J.M., Neupert W., Stuart R.A. (1997) Insertion into the mitochondrial inner membrane of a polytopic protein, the nuclear-encoded Oxa1p. *Embo J*. **16**: 2217–2226.

Hjort K., Goldberg A.V., Tsaousis A.D., Hirt R.P., Embley T.M. (2010) Diversity and reductive evolution of mitochondria among microbial eukaryotes. *Philos Trans R Soc Lond B Biol Sci*. **365**: 713–27.

Iwata S., Lee J.W., Okada K., Lee J.K., Iwata M., Rasmussen B., Link T.A., Ramaswamy S., Jap B.K. (1998) Complete structure of the 11-subunit bovine mitochondrial cytochrome bc1 complex. *Science*. **281**: 64–71.

Jakob M., Hoffmann A., Amodeo S., Peitsch C., Zuber B., Ochsenreiter T. (2016) Mitochondrial growth during the cell cycle of *Trypanosoma brucei* bloodstream forms. *Sci. Rep*. **6**: 1–13.

Jensen, R.E., Englund, P.T. (2012) Network News: The replication of kinetoplast DNA. *Annu Rev Microbiol*. **66**: 473–491.

Joubert J.J., Schutte CH., Irons D.J., Fripp P.J. (1993) Ubombo and the site of David Bruce's discovery of *Trypanosoma brucei*. *Trans R Soc Trop Med Hyg*. **87**: 494–5.

Karnkowska A., Vacek V., Zubáčová Z., Treitli S.C., Petrželková R., Eme L., Novák L., Žárský V., Barlow L.D., Herman E.K., Soukal P., Hroudová M., Doležal P., Stairs C.W., Roger A.J., Eliáš M., Dacks J.B., Vlček Č., Hampl V. (2016) A Eukaryote without a Mitochondrial Organelle. *Curr Biol*. **26**: 1274–84.

Kennedy P.G.E. (2019) Update on human African trypanosomiasis (sleeping sickness). *J Neurol*. **266**: 2334–2337.

Khalimonchuk O., Bird A., Winge D.R. (2007) Evidence for a pro-oxidant intermediate in the assembly of cytochrome oxidase. *J Biol Chem*. **282**: 17442–17449.

Klingbeil M.M., Motyka S.A., Englund P.T. (2002) Multiple mitochondrial DNA polymerases in *Trypanosoma brucei*. *Mol Cell*. **10**: 175–186.

Koehler C.M. (2000) Protein translocation pathways of the mitochondrion. *FEBS Lett*. **476**: 27–31.

Kuhlbrandt W. (2019) Structure and mechanisms of F-type ATP synthases. *Annu. Rev. Biochem*. **88**: 515–549.

Kumazaki K., Chiba S., Takemoto M., Furukawa A., Nishiyama K., Sugano Y., Mori T., Dohmae N., Hirata K., Nakada-Nakura Y. (2014) Structural basis of Sec-independent membrane protein insertion by YidC. *Nature*. **509**: 516–520.

Kurland C.G., Andersson S.G. (2000) Origin and evolution of the mitochondrial proteome. *Microbiol Mol Biol Rev*. **64**: 786–820.

Kutik S., Guiard B., Meyer H.E., Wiedemann N., Pfanner N. (2007) Cooperation of translocase complexes in mitochondrial protein import. *J. Cell Biol*. **179**: 585–591.

Lancaster C.R.D. (2002) Succinate:quinone oxidoreductases: an overview. *Biochim Biophys Acta*. **1553**: 1–6.

Langousis G., Hill K.L. (2014) Motility and more: the flagellum of *Trypanosoma brucei*. *Nat Rev Microbiol*. **12**: 505–518.

Leger M.M., Petrů M., Žárský V., Eme L., Vlček Č., Harding T., Lang B.F., Eliáš M., Doležal P., Roger A.J. (2015) An ancestral bacterial division system is widespread in eukaryotic mitochondria. *Proc Natl Acad Sci U S A*. **112**: 10239–46.

Liu B., Liu Y., Motyka S.A., Agbo E.E., Englund P.T. (2005) Fellowship of the rings: the replication of kinetoplast DNA. *Trends Parasitol*. **21**: 363–369.

Liu Q., D'Silva P., Walter W., Marszalek J., Craig E. A. (2003) Regulated cycling of mitochondrial Hsp70 at the protein import channel. *Science*. **300:** 139–141.

Liu, B., Wang j., Yaffe N., Lindsay M., Zhao Z., Zick A., Shlomai J., Englund P.T. (2009) Trypanosomes have six mitochondrial DNA helicases with one controlling kinetoplast maxicircle replication. *Mol Cell*. **35**: 490–501.

Lukeš J., Jirků M., Doležel D., Kraľová I., Hollar L., Maslov D.A. (1997) Analysis of ribosomal RNA genes suggests that trypanosomes are monophyletic. *J Mo. Evol*. **44**: 521–527.

Lukeš J., Kachale A., Votýpka J., Butenko A., Field M.C. (2022) African trypanosome strategies for conquering new hosts and territories: the end of monophyly? *Trends Parasitol*. **38**: 724–736.

Lukeš J., Skalický T., Týč J., Votýpka J., Yurchenko V. (2014) Evolution of parasitism in kinetoplastid flagellates. *Mol Biochem Parasitol*. **195**: 115–22.

Lutkenhaus J. (1993) FtsZ ring in bacterial cytokinesis. *Mol Microbiol*. **9**: 403–9.

Maarse A.C., 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 M.C., Mulhern T.D., 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.

Maldonado M., Guo F., Letts J.A. (2021) Atomic structures of respiratory complex III$_2$, complex IV, and supercomplex III$_2$-IV from vascular plants. *Elife*. **10**: e62047.

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.

Mani J., Rout S., Desy S., Schneider A. (2017) Mitochondrial protein import - Functional analysis of the highly diverged Tom22 orthologue of *Trypanosoma brucei*. *Sci Rep*. **7:** 40738.

Mannella C.A. (1992) The 'ins' and 'outs' of mitochondrial membrane channels. *Trends Biochem Sci*. **17**: 315–20.

Margolin W. (2005) FtsZ and the division of prokaryotic cells and organelles. *Nat Rev Mol Cell Biol*. **6**: 862–71.

Margulis L. (1970) Origin of eukaryotic cells; evidence and research implications for a theory of the origin and evolution of microbial, plant, and animal cells on the Precambrian earth. New Haven: Yale University Press

Marteyn B.S., Karimova G., Fenton A.K., Gazi A.D., West N., Touqui L., Prevost M.C., Betton J.M., Poyraz O., Ladant D., Gerdes K., Sansonetti P.J., Tang C.M. (2014) ZapE is a novel cell division protein interacting with FtsZ and modulating the Z-ring dynamics. *MBio*. **5**: e00022-14.

Melendy, T., Sheline, C., Ray, D.S. (1988) Localization of a type II DNA topoisomerase to two sites at the periphery of the kinetoplast DNA of *Crithidia fasciculata*. *Cell*. **55**: 1083–1088.

Mitchell P. (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemiosmotic type of mechanism. *Nature*. **191**: 144–8.

Mokranjac D., Bourenkov G., Hell K., Neupert W., Groll M. (2006) Structure and function of Tim14 and Tim16, the J and J-like components of the mitochondrial protein import motor. *EMBO J.* **25**: 4675–4685.

Morgan D.J., Sazanov L.A. (2008) Three-dimensional structure of respiratory complex I from *Escherichia coli* in ice in the presence of nucleotides. *Biochim Biophys Acta.* **1777**: 711–718.

Nargang F.E., Preuss M., Neupert W., Herrmann J.M. (2002) The oxal protein forms a homooligomeric complex and is an essential part of the mitochondrial export translocase in *Neurospora crassa*. *J Biol Chem.* **277**: 12846–12853.

Neupert W. (1997) Protein import into mitochondria. *Annu Rev Biochem.* **66**: 863–917.

Neupert W., Herrmann J.M. (2007) Translocation of proteins into mitochondria. *Annu Rev Biochem.* **76**: 723–749.

Ngo H., Tschudi C., Gull K., Ullu E. (1998) Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A*. **95**: 14687–14692.

Nishida K., Takahara M., Miyagishima S.Y., Kuroiwa H., Matsuzaki M., Kuroiwa T. (2003) Dynamic recruitment of dynamin for final mitochondrial severance in a primitive red alga. *Proc Natl Acad Sci U S A*. **100**: 2146–51.

Ntambi J.M., Englund P.T. (1985) A gap at a unique location in newly replicated kinetoplast DNA minicircles from *Trypanosoma equiperdum. J Biol Chem.* **260**: 5574–9.

Ogbadoyi E.O., Robinson D.R., Gull K. (2003) A high-order transmembrane structural linkage is responsible for mitochondrial genome positioning and segregation by flagellar basal bodies in trypanosomes. *Mol Biol Cell.* **14**: 1769–1779.

Otera H., Wang C., Cleland M.M., Setoguchi K., Yokota S., Youle R.J., Mihara K. (2010) Mff is an essential factor for mitochondrial recruitment of Drp1 during mitochondrial fission in mammalian cells. *J Cell Biol.* **191**: 1141–58.

Pais J. E., Schilke B., Craig E. A. (2011) Reevaluation of the role of the Pam18:Pam16 interaction in translocation of proteins by the mitochondrial Hsp70-based import motor. *Mol Biol Cell.* **22**: 4740–4749.

Palmer C.S., Osellame L.D., Laine D., Koutsopoulos O.S., Frazier A.E., Ryan M.T. (2011) MiD49 and MiD51, new components of the mitochondrial fission machinery. *EMBO Rep*. **12**: 565–73.

Peikert C.D., Mani J., Morgenstern M., Käser S., Knapp B., Wenger C., Harsman A., Oeljeklaus S., Schneider A., Warscheid B. (2017) Charting organellar importomes by quantitative mass spectrometry. *Nat Commun*. **8**: 15272.

Peters P.C., Migocki M.D., Thoni C., Harry E.J. (2007) A new assembly pathway for the cytokinetic Z ring from a dynamic helical structure in vegetatively growing cells of *Bacillus subtilis*. *Mol Microbiol*. **64**: 487–99.

Pfanner N, Warscheid B, Wiedemann N. (2019) Mitochondrial proteins: from biogenesis to functional networks. *Nat Rev Mol Cell Biol*. **20**: 267–284.

Povelones M.L. (2014) Beyond replication: division and segregation of mitochondrial DNA in kinetoplastids. *Mol Biochem Parasitol*. **196**: 53–60.

Ramrath D.J.F., Niemann M., Leibundgut M., Bieri P., Prange C., Horn E.K., Leitner A., Boehringer D., Schneider A., Ban N. (2018) Evolutionary shift toward protein-based architecture in trypanosomal mitochondrial ribosomes. *Science*. **362**: eaau7735.

Roger A.J. (1999) Reconstructing early events in eukaryotic evolution. *Am. Nat*. **154**: 146–163.

Roger A.J., Muñoz-Gómez S.A., Kamikawa R. (2017) The Origin and Diversification of Mitochondria. *Curr Biol*. **27**: 1177–1192.

Ryan K.A., Englund P.T. (1989a) Replication of kinetoplast DNA in *Trypanosoma equiperdum*: Minicircle H strand fragments which map at specific locations. *J Biol Chem*. **264**: 823–830.

Ryan K.A., Englund, P.T. (1989b) Synthesis and processing of kinetoplast DNA minicircles in*Trypanosoma equiperdum*. *Mol Cell Biol*. **9**: 3212.

Sanjuán Szklarz L.K., Guiard B., Rissler M., Wiedemann N., Kozjak V., van der Laan M., Lohaus C., Marcus K., Meyer H.E., Chacinska A., Pfanner N., Meisinger C. (2005) Inactivation of the mitochondrial heat shock protein zim17 leads to aggregation of matrix hsp70s followed by pleiotropic effects on morphology and protein biogenesis. *J Mol Biol*. **351:** 206–218.

Saraste M., Sibbald P.R., Wittinghofer A. (1990) The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem. Sci*. **15**: 430–434.

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

Scheffler I.E. (2007) Mitochondria 2nd ed. Chichester, UK: John Wiley & Sons, Inc

Schnaufer A., Clark-Walker G.D., Steinberg A.G., Stuart K. (2005) The F1-ATP synthase complex in bloodstream stage trypanosomes has an unusual and essential function. *EMBO J*. **24**: 4029–40.

Scott I., Youle R.J. (2010) Mitochondrial fission and fusion. *Essays Biochem*. **47**: 85–98.

Sela D., Shlomai J. (2009) Regulation of UMSBP activities through redox-sensitive protein domains. *Nucleic Acids Res*. 37: 279–288.

Serricchio M., Bütikofer P. (2011) *Trypanosoma brucei*: a model micro-organism to study eukaryotic phospholipid biosynthesis. *FEBS Journal*. **278**: 1035–1046.

Shapiro T.A. (1993) Kinetoplast DNA maxicircles: networks within networks. *PNAS*. **90**: 7809–7813.

Simpson A.G.B., Lukeš J., and Roger A.J. (2002). The evolutionary history of kinetoplastids and their kinetoplasts. *Mol Biol Evol*. **19**: 2071–2083.

Simpson L. (1986) Kinetoplast DNA in trypanosomid flagellates. *Int Rev* Cytol. **99**: 119–179.

Smith T.K., Bringaud F., Nolan D.P., Figueiredo L.M. (2017) Metabolic reprogramming during the *Trypanosoma brucei* life cycle. *F1000Res*. **6**: F1000 Faculty Rev-683.

Spang A., Saw J.H., Jørgensen S.L., Zaremba-Niedzwiedzka K., Martijn J., Lind A.E., van Eijk R., Schleper C., Guy L., Ettema T.J.G. (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature*. **521**: 173–179.

Stiburek L., Fornuskova D., Wenchich L., Pejznochova M., Hansikova H., Zeman J. (2007) Knockdown of human Oxa1l impairs the biogenesis of F1Fo-ATP synthase and NADH: Ubiquinone oxidoreductase. *J. Mol. Biol*. **374**: 506–516.

Stiller S.B., Höpker J., Oeljeklaus S., Schütze C., Schrempp S.G., Vent-Schmidt J., Horvath S.E., Frazier A.E., Gebert N., van der Laan M., Bohnert M., Warscheid B., Pfanner N., Wiedemann N. (2016) Mitochondrial OXA translocase plays a major role in biogenesis of inner-membrane proteins. *Cell Metab*. **23**: 901–908.

Sutton R.E., Boothroyd J.C. (1986) Evidence for trans splicing in trypanosomes. *Cell*. **47**: 527–535.

Takahara M., Takahashi H., Matsunaga S., Miyagishima S., Takano H., Sakai A., Kawano S., Kuroiwa T. (2000) A putative mitochondrial ftsZ gene is present in the unicellular primitive red alga *Cyanidioschyzon merolae*. *Mol Gen Genet*. **264**: 452–60.

TerBush A.D., Yoshida Y., Osteryoung K.W. (2013) FtsZ in chloroplast division: structure, function and evolution. *Curr Opin Cell Biol*. **25**: 461–70.

Thanedar S., Margolin W. (2004) FtsZ exhibits rapid movement and oscillation waves in helix-like patterns in *Escherichia coli*. *Curr Biol*. **14**: 1167–73.

Tielens A.G., Rotte C., van Hellemond, J.J., Martin W. (2002) Mitochondria as we don't know them. *Trends Biochem. Sci*. **27**: 564–572.

Tielens A.G.M., Van Hellemond J.J. (1998) Differences in energy metabolism between Trypanosomatidae. *Trends Parasitol*. **14**: 265–271.

Truscott K. N., Voos W., Frazier A. E., Lind M., Li Y., Geissler A., Dudek J., Müller H., Sickmann A., Meyer H.E., 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.

Tucker K., Park E. (2019) Cryo-EM structure of the mitochondrial protein-import channel TOM complex at near-atomic resolution. *Nat Struct Mol Biol*. **26**: 1158–1166.

van der Laan M., Chacinska A., Lind M., Perschil I., Sickmann A., Meyer H. E., 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.

Verner, Z., Basu S., Benz C., Dixit S., Dobáková E., Faktorová D., Hashimi H., Horáková E., Huang Z., Paris Z., Peña-Diaz P., Ridlon L., Týč J., Wildridge D., Zíková A., Lukeš J. (2015) Malleable mitochondrion of *Trypanosoma brucei*. *Int Rev Cell Mol Biol*. **315**: 73–151.

Vickerman K. (1985) Developmental cycles and biology of pathogenic trypanosomes. *Br. Med. Bull*. **41**: 105–114.

Vickerman K., Tetley L., Hendry K.A.K., Turner C.M.R. (1988) Biology of African trypanosomes in the tsetse fly. *Biol Cell*. **64**: 109–119.

Viegas I.J., de Macedo J.P., Serra L., De Niz M., Temporão A., Silva Pereira S., Mirza A.H., Bergstrom E., Rodrigues J.A., Aresta-Branco F., Jaffrey S.R., Figueiredo L.M. (2022) N^6-methyladenosine in poly(A) tails stabilize VSG transcripts. *Nature*. **604**: 362–370.

Voos W., Martin H., Krimmer T., Pfanner N. (1999) Mechanisms of protein translocation into mitochondria. *Biochim. Biophys. Acta*. **1422**: 235–254.

Walker J.E. (2013) The ATP synthase: the understood, the uncertain and the unknown. *Biochem Soc Trans*. **41**: 1–16.

Walker J.E. (2013) The ATP synthase: The understood, the uncertain and the unknown. *Biochem. Soc. Trans*. **41**: 1–16.

Wang Z., Morris J.C., Drew M.E., Englund P.T. (2000) Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. *J Biol Chem*. **275**: 40174–9.

Wexler-Cohen Y., Stevens G.C., Barnoy E., van der Bliek A.M., Johnson P.J. (2014) A dynamin-related protein contributes to Trichomonas vaginalis hydrogenosomal fission. *FASEB J*. **28**: 1113–21.

Wiedemann N., Pfanner N. (2017) Mitochondrial Machineries for Protein Import and Assembly. *Annu Rev Bio-chem*. **86**: 685–714.

Wienke D.C., Knetsch M.L., Neuhaus E.M., Reedy M.C., Manstein D.J. (1999) Disruption of a dynamin homologue affects endocytosis, organelle morphology, and cytokinesis in Dictyostelium discoideum. *Mol Biol Cell*. **10**: 225–43.

Woodward R., Gull K. (1990) Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. *J Cell Sci.* **95**: 49–57.

Yaffe N., Rotem D., Soni A., Porath D., Shlomai J. (2021) Direct monitoring of the stepwise condensation of kinetoplast DNA networks. *Sci Rep.* **11**: 1501.

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.

Youle R.J., van der Bliek A.M. (2012) Mitochondrial Fission, Fusion, and Stress. *Science.* **337**: 1062–5.

Zaremba-Niedzwiedzka K., Caceres E.F., Saw J.H., Bäckström D., Juzokaite, L., Vancaester E., Seitz K.W., Anantharaman K., Starnawski P., Kjeldsen K.U., Stott M.B., Nunoura T., Banfield J.F., Schramm A., Baker B.J., Spang A., Ettema T.J. (2017) Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature.* **541**: 353–358.

Zhang Z., Huang L., Shulmeister V.M., Chi Y.I., Kim K.K., Hung L.W., Crofts A.R., Berry E.A., Kim S.H. (1998) Electron transfer by domain movement in cytochrome bc1. *Nature.* **392**: 677–84.

Zhao J., Liu T., Jin S., Wang X., Qu M., Uhlén P., Tomilin N., Shupliakov O., Lendahl U., Nistér M. (2011) Human MIEF1 recruits Drp1 to mitochondrial outer membranes and promotes mitochondrial fusion rather than fission. *EMBO J* **30**: 2762–2778.

Zhao Z., Lindsay M.E., Chowdhury A.R., Robinson D.R., Englund P.T. (2008) A link between the trypanosome mitochondrial DNA and flagellum, mediates genome segregation. *EMBO J.* **27**: 143–154.

Books:

ALBERTS, Bruce, Alexander JOHNSON, Julian LEWIS, Martin RAFF, Keith ROBERTS and Peter WALTER. Molecular biology of the cell. 5th ed. New York: Garland Science, c2008. ISBN 978-0-8153-4106-2.

BERG, Jeremy M., John L. TYMOCZKO, Lubert STRYER and Lubert STRYER. Biochemistry. 5th ed. New York: W.H. Freeman, c2002. ISBN 0-7167-3051-0.

KOOLMAN, Jan and Klaus-Heinrich RÖHM. Color atlas of biochemistry. 2nd ed., rev. and enl. New York: Thieme, c2005. Thieme flexibook. ISBN 3-13-100372-3.

45

4. Results

Chapter I.

ZapE/Afg1 interacts with Oxa1 and its depletion causes a multifaceted phenotype

PLoS ONE (2020) **15**: e0234918

PLOS ONE

RESEARCH ARTICLE

ZapE/Afg1 interacts with Oxa1 and its depletion causes a multifaceted phenotype

Jan Pyrih [1☉¤a]*, Vendula Rašková[1,2☉], Ingrid Škodová-Sveráková[1,3], Tomáš Pánek[4¤b], Julius Lukeš[1,2]*

1 Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice (Budweis), Czech Republic, **2** Faculty of Science, University of South Bohemia, České Budějovice (Budweis), Czech Republic, **3** Faculty of Sciences, Comenius University, Bratislava, Slovakia, **4** Faculty of Sciences, University of Ostrava, Ostrava, Czech Republic

☉ These authors contributed equally to this work.
¤a Current address: Faculty of Science, Charles University, Biocev, Vestec, Czech Republic
¤b Current address: Faculty of Science, Charles University, Prague, Czech Republic
* jula@paru.cas.cz (JL); jan.pyrih@gmail.com (JP)

Abstract

ZapE/Afg1 is a component of the inner cell membrane of some eubacteria and the inner mitochondrial membrane of eukaryotes. This protein is involved in FtsZ-dependent division of eubacteria. In the yeast and human mitochondrion, ZapE/Afg1 likely interacts with Oxa1 and facilitates the degradation of mitochondrion-encoded subunits of respiratory complexes. Furthermore, the depletion of ZapE increases resistance to apoptosis, decreases oxidative stress tolerance, and impacts mitochondrial protein homeostasis. It remains unclear whether ZapE is a multifunctional protein, or whether some of the described effects are just secondary phenotypes. Here, we have analyzed the functions of ZapE in *Trypanosoma brucei*, a parasitic protist, and an important model organism. Using a newly developed proximity-dependent biotinylation approach (BioID2), we have identified the inner mitochondrial membrane insertase Oxa1 among three putative interacting partners of ZapE, which is present in two paralogs. RNAi-mediated depletion of both ZapE paralogs likely affected the function of respiratory complexes I and IV. Consistently, we show that the distribution of mitochondrial ZapE is restricted only to organisms with Oxa1, respiratory complexes, and a mitochondrial genome. We propose that the evolutionarily conserved interaction of ZapE with Oxa1, which is required for proper insertion of many inner mitochondrial membrane proteins, is behind the multifaceted phenotype caused by the ablation of ZapE.

Introduction

Most eukaryotes retain a highly complex mitochondrion, which became an essential component of the cell by harboring the enzymatic machinery for ATP production and other important metabolic pathways [1]. Recent studies of mitochondria from non-model and neglected eukaryotic microorganisms revealed that their organelle contains several proteins or even complex metabolic pathways that are not present in the model eukaryotes. Mitochondrial FtsZ

division machinery and the Type II secretion system may serve as such recently documented examples [2,3]. At the same time, many conserved proteins found in the mitochondria across a broad range of eukaryotes including animals, yeast, and/or plants still await their functional characterization.

ZapE (alternatively Afg1 in yeast or LACE1 in humans) is an ATPase located in the inner membrane of the human and yeast mitochondria, and in the inner membrane of *Escherichia coli* [4,5]. The depletion of the human ZapE homolog triggers morphological changes of the mitochondria, eventually leading to their fragmentation [6]. A somewhat similar phenotype was observed in bacteria, where the affected cells became elongated following the up- or down-regulation of ZapE [5]. The available data is compatible with the view that in bacteria, ZapE is part of the FtsZ division machinery [5]. However, the function of human and yeast homologs may be different, as it was proposed that their ZapE orthologs mediate degradation of the mitochondrially-encoded subunits of the respiratory complex IV [4,6]. Furthermore, both organisms lack the FtsZ division system [3]. ZapE was shown to mediate the translocation of p53 and subsequent apoptosis in humans [7]. It is also noteworthy that ZapE was highly affected in the proteomic survey of the Oxa1 depletome in humans [8]. Moreover, the functional link between Oxa1 and ZapE was also suggested in yeast, where only the mitochondrially-encoded subunits of respiratory complex IV were affected after the depletion of ZapE. Finally, a novel role for this protein in maintaining mitochondrial matrix proteostasis was suggested [9]. All in all, the phenotypes associated with ZapE in different organisms vary widely, and so far, they have not been integrated into a coherent picture.

Trypanosoma brucei is both an important human pathogen causing African sleeping sickness and a model organism with highly developed molecular biology tools. It contains a single reticulated mitochondrion with its own genome represented by a network of mutually catenated DNA circles, termed kinetoplast DNA (kDNA) [10]. Transcripts of several kDNA-encoded genes become translatable only after they undergo extensive RNA editing of the uridine insertion/deletion type [11]. Another unique feature of the *T. brucei* mitochondrion is its capacity to undergo massive morphological and structural changes in the course of the parasite's life cycle, which involves vertebrate hosts and the tse-tse fly vector [12,13]. To shed light on the function(s) of the conserved ZapE protein, so far examined only in bacteria and opisthokonts, we have probed the function and interactions of its two paralogs, ZapE1 (XP_823041.1, Tb927.7.6930) and ZapE2 (XP_846313.1; Tb927.10.8070). According to the ATOM40 depletome-based mitoproteome [13] and Tryptag *in-situ* tagging database [14], both proteins are genuine components of the *T. brucei* mitochondrion.

A range of recently developed techniques that take advantage of proximity biotinylation, such as BioID, TurboID, and APEX, are particularly suitable for the studies of protein-protein interactions [15,16]. As compared to classical co-immunoprecipitation, their advantage is higher reproducibility and capacity to identify both stable complexes and transient interactions. Briefly, the protein of interest is fused with modified biotin ligase, which promiscuously biotinylates proteins in its proximity. However, until now BioID has been used in *T. brucei* only in just a handful of studies [17,18]. Here, we took advantage of the recently developed advanced biotin ligation-based approach named BioID2, which we have successfully adapted for *T. brucei*. In contrast to the classical BioID technique, BioID2 biotin ligase is smaller (26 kDa), more specific, and attached to the protein of interest by several nm-long linker, which improves protein folding and the biotinylation range [19]. When applied to *T. brucei*, BioID2 labeling produced a highly specific output. Consistent with human and yeast, Oxa1 was identified among three putative interaction partners of ZapE2. As *T. brucei* is very distantly related to opisthokonts, we hypothesize that interaction of ZapE and Oxa1 is conserved in eukaryotes.

48

This view is further supported by the co-occurrence of ZapE and Oxa1 genes and by the α-proteobacterial origin of eukaryotic ZapE, as documented by extensive phylogenetic analysis.

Results

Proximity-dependent biotinylation in *T. brucei*

In order to establish the *in situ* proximity biotinylation approach in *T. brucei*, the mNeongreen gene was replaced with the HA-tagged modified biotin ligase from *Aquifex aeolicus* (BioID2) in the pPOTv7_mNG-Blast plasmid [14]. To increase the range of proximity biotinylation [19], we have designed ~8 nm-long glycine-serine (GGGGS) repeat-containing linker by which biotin ligase becomes linked with the protein of interest (Fig 1A). This pPOT_BioID2 plasmid served as a template for PCR-based C-terminal *in situ* tagging of the ZapE1 and ZapE2 genes in *T. brucei*. IscU, a chaperone involved in iron-sulfur (Fe-S) cluster assembly in *T. brucei* [20], and ligase K-β (LigK-β), which is critical for the kDNA replication and is located in the antipodal sites of the *T. brucei* kDNA network [21], were used as controls. Similarly, the mitochondrial import signal of the IscU gene was *in situ*-tagged with biotin ligase to serve as an additional control. For the *in situ* tagging strategy, pPOT_BioID2 plasmid map, and the full sequence of BioID2 protein with linker see S1 Fig.

The mitochondrial localization of the BioID2-tagged proteins and their ability to biotinylate surrounding proteins was confirmed by direct fluorescent Alexa-488-streptavidin labelling of the procyclic stage of *T. brucei* (Fig 1B). This streptavidin-fluorescent conjugate detected specific biotinylation in the mitochondrial lumen of the ZapE1, ZapE2, and IscU cell lines. In the case of LigK-β, the labelling occurred in the two opposing antipodal sites of the kDNA disk (Fig 1B). The leader sequence of IscU was insufficient for the delivery of the BioID2 protein into the organelle, resulting in a predominantly cytosolic localization (Fig 1B).

Identification of mitochondrial proteins using BioID2

An intrinsic feature of BioID is that very abundant proteins become biotinylated not only because of their specific interaction with the bait protein but also due to their by-chance proximity. Therefore, we focused on proteins enriched or exclusively present in at least three out of four analyzed BioID2 samples, namely ZapE1, ZapE2, IscU, and LigK-β when compared to the IscU-leader negative control (Fig 2A). The mitochondrial import signal of IscU-containing BioID2 construct was used as a negative control for proximity-dependent biotinylation, as most of the fusion protein remained in the cytoplasm (Fig 1B). To identify mitochondrial proteins, purified mitochondria of the *T. brucei* procyclic stage were dissolved in 1% SDS-containing buffer and incubated at 80°C for 10 min, after which solubilized biotinylated proteins were affinity-purified by streptavidin-coated Dynabeads. For each protein, three independent biological replicates have been performed. The protein composition of each affinity purification was determined by label-free quantitative proteomics.

A comparative analysis revealed a set of 117 mitochondrial proteins (S1 Table). The identified set primarily contains strongly expressed mitochondrial proteins such as mtHsp70 and Cpn60, which due to their proximity to the BioID2-tagged protein apparently became randomly biotinylated. In total, 114 out of 117 proteins have previously been listed in the mitochondrial proteome or shown as located within the organelle *via* the *in situ* tagging approach [13,14]. The remaining three proteins most likely also reside in the organelle, as they are either strongly predicted to be targeted into it or are homologs of known mitochondrial proteins (S1 Table). In the IscU-leader control, a small fraction was delivered into the organelle, yet in this negative control the detected mitochondrial proteins were on average 60 times less abundant as compared to bait proteins datasets (S1 Table).

49

A.
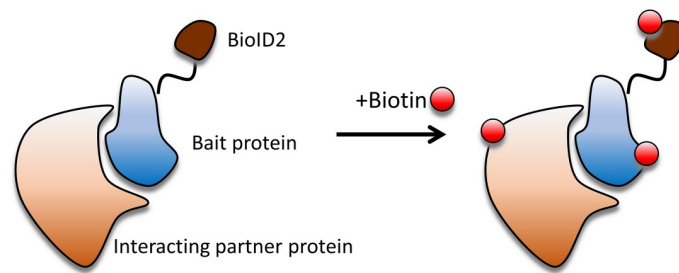
B.
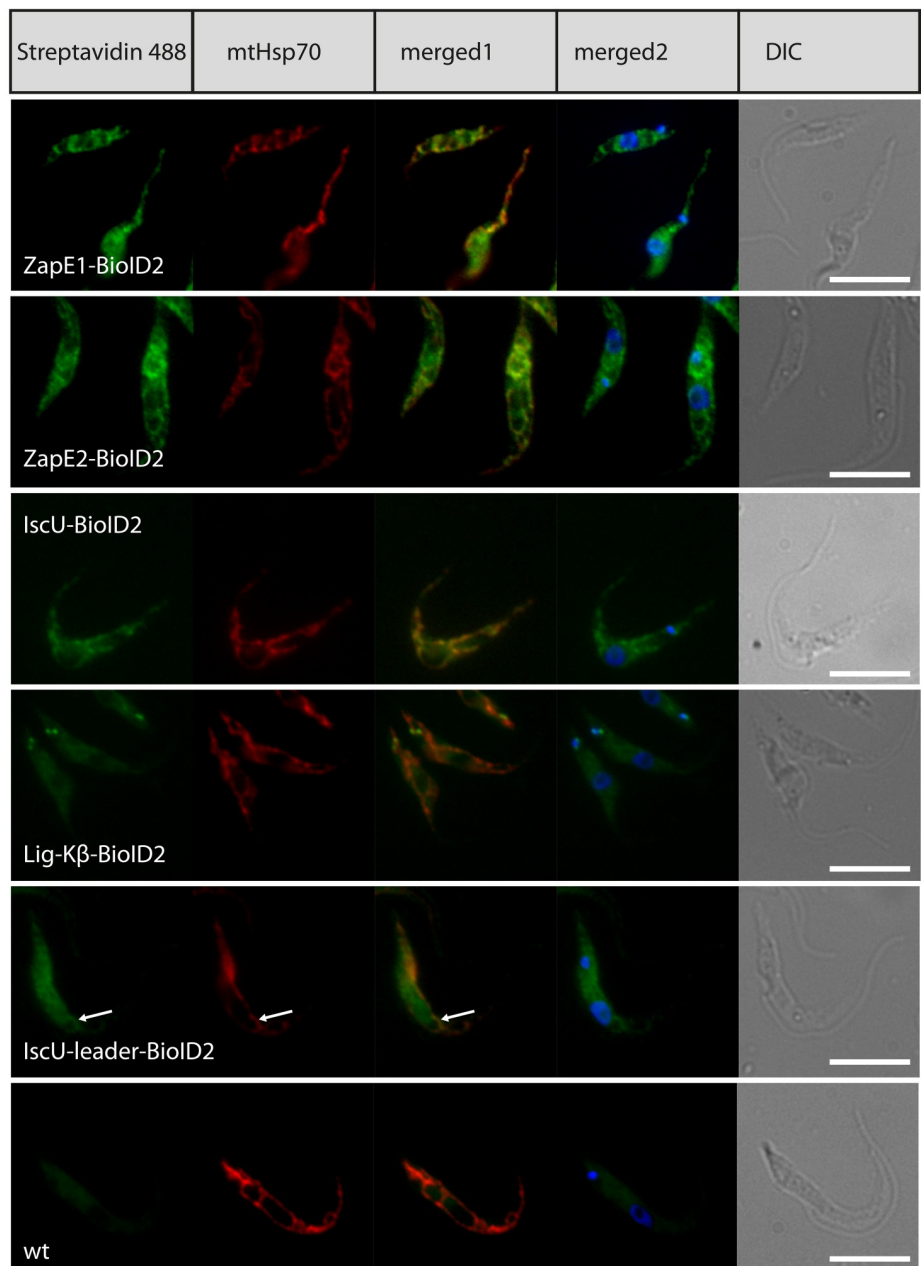
Fig 1. Mitochondrial proteins *in situ* tagged with BioID2. (**A**) A scheme of *in vivo* proximity-dependent biotinylation reaction induced by the addition of biotin. (**B**) Specific biotinylation for ZapE1, ZapE2, IscU, LigK-β, and IscU_Leader BioID2-HA fusion constructs. Fluorescently labeled Alexa-488 streptavidin detects biotinylated proteins in the mitochondrion (ZapE1, ZapE2, IscU), the antipodal kDNA sites (LigK-β) or predominantly in the cytoplasm (IscU_Leader BioID2 fusion). wt, wild type cells served as a control for background biotin staining. Monoclonal α-mHsp70 antibody was used as a mitochondrial marker. Arrow indicates the area in which partial mitochondrial localization of the IscU_Leader BioID2 fusion is visible. DNA was labeled with DAPI. DIC, differential interference contrast. Scale bars, 5 μm.

Identification of interacting proteins using BioID2

To distinguish between putative interacting partners and spurious interactions with abundant mitochondrial proteins, such as mtHsp70, other bait protein interactomes were used as controls. For instance, for LigK-β, purified IscU, ZapE1, and ZapE2 datasets were considered as negative controls. Proteins exclusively measured or more than 3x enriched as compared to the control datasets were considered as putative interacting partners (Fig 2A). By employing this strategy, however, we may filter out proteins that interact with two or more bait proteins. While this is unlikely for IscU and LigK-β, which have very different functions, the existence of a subset of overlapping interacting partners off ZapE1 and ZapE2 was plausible. For this reason, these two datasets were not treated reciprocally as a negative control but were compared to the IscU and LigK-β datasets only. As expected, the bait proteins invariably occupied the top position due to autobiotinylation by the fused BioID2 enzyme [19]. Along with the auto-biotinylated protein, less than 10 significantly enriched proteins were identified for each analyzed bait (Fig 2B). For ZapE1 and LigK-β, only a single putative interacting protein was identified. A full list of proteins enriched both above and below the chosen threshold is available in S2 Table. Oxa1 was identified exclusively in the ZapE2 interactome. The other two proteins likely interacting with ZapE2 are homologs of NADH:ubiquinone oxidoreductase complex assembly factor 3 (NDUFAF3) and DnaJ chaperone, both *bona fide* components of the *T. brucei* respiratory chain complex I [22]. While NDUFAF3 was identified as a putative Oxa1-interacting protein in humans [23], the function of DnaJ chaperone in complex I of *T. brucei* is unknown, as homologs of this protein were so far not encountered associated with complex I in other eukaryotes.

IscU is a core component of the Fe-S cluster assembly pathway [11]. Three identified putative interacting partners–frataxin, *Tb*Nfu1, and Tb927.11.15470 (possible homolog of methionyl-tRNA formyltransferase) were exclusively associated with the IscU bait protein, with six additional proteins more than 3x enriched in the IscU dataset as compared to the other bait proteins. While frataxin is a well-known interacting partner of IscU in various organisms [24], a possible role of *Tb*Nfu1 in the early steps of the Fe-S cluster assembly pathway is intriguing. *T. brucei* contains three essential Nfu paralogs in its mitochondrion [25], which implies their unique specialization. While *Tb*Nfu2 and *Tb*Nfu3 likely function in the late steps of the Fe-S cluster pathway, as they were able to compensate for the depletion of their yeast homologs, the roles of *Tb*Nfu1 [25] and Tb927.11.15470 remain unknown. Finally, only one protein, Tb927.8.4230, putatively interacting with LigK-β was identified.

Mitochondrial localization of putative interacting proteins

Most identified putative interacting proteins were either previously experimentally shown to localize to the mitochondrion of *T. brucei* [14] or are components of the mitochondrial ATOM40 depletome-based importome of the same organism [13]. Proteins for which intracellular distribution has not yet been convincingly shown were C-terminally *in situ*-tagged with the V5 tag and their mitochondrial localization was determined by immunofluorescence (Fig

51
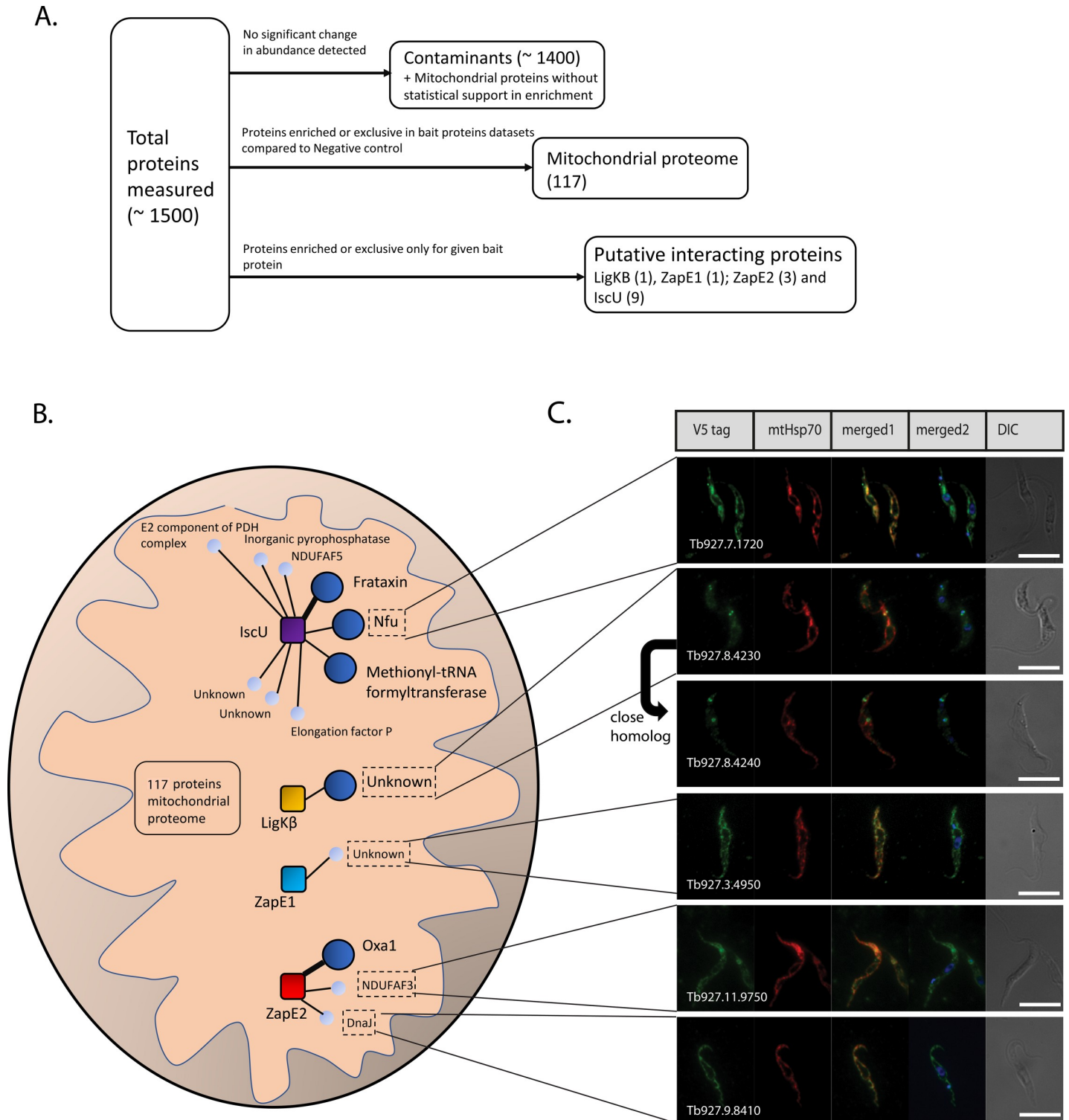
Fig 2. BioID2-labelled putative interacting proteins. (**A**) Scheme of how the proteomic data was processed to identify the mitochondrial proteome and putative interacting proteins. (**B**) Schematic representation displays identified putative interacting proteins. Each bait protein is represented by a square and is surrounded by identified putative interacting proteins visualized by a circle. The larger size of the circle for identified proteins highlights those exclusively identified only for a given bait protein or proteins enriched on average more than 100x as compared to other datasets. A thick line represents the expected interaction of the identified proteins based on data from other organisms. Proteins surrounded by the dashed rectangle are experimentally localized for the first time in this study. (**C**) Intracellular localization of proteins identified by the BioID2 technique. *In situ* C-terminally V5-tagged proteins were expressed in *T. brucei* and their localization was inspected using

52

immunofluorescence microscopy. Monoclonal α-V5 rabbit and α-mHsp70 mouse antibodies were used. DNA was stained with DAPI. DIC, differential interference contrast. Scale bars, 5 μm.

2C). Thus we demonstrate that Tb927.11.9750 protein, a *bona fide* component of complex I [22], is beyond reasonable doubt localized within the organelle (Fig 2C). Moreover, Tb927.8.4230, a 119 kDa acidic protein (pI 5.4) is specifically confined to the antipodal sites of the kDNA disk, which is consistent with the localization of its interactor, LigK-β. Interestingly, an additional protein (Tb927.8.4240) identified by a BLAST search, seems to be a homolog of Tb927.8.4230 (e-value of $3e^{-40}$ in BLAST against *T. brucei* protein database). Genes for these homologs reside on the same chromosome next to each other. Surprisingly, Tb927.8.4240 is not localized directly in the antipodal sites of the kDNA disk, but in their proximity (Fig 1C).

Depletion of ZapE proteins potentially affects the activity of complexes I and IV

Initially, we created single RNAi knockdowns for either ZapE1 or ZapE2 in procyclic cells. While complete elimination of the respective V5-tagged targeted proteins (S2 Fig) was achieved in both cell lines, this was not associated with any measurable growth phenotype (S2 Fig). The most plausible explanation is that these conserved, yet individually non-essential proteins can functionally substitute each other. To test this hypothesis, we created a cell line in which both ZapE paralogs were downregulated by RNAi. To do that, ~500 nt-long regions of the ZapE1 and ZapE2 genes were cloned into the RNAi pTrypSon-Phleo and pTrypSon-Blast plasmids, respectively [26]. We then transfected both plasmids one at a time into a cell line in which ZapE1 and ZapE2 were *in situ*-tagged with the V5 and HA tags, respectively. The efficient ablation of the tagged proteins following RNAi induction was monitored by western blotting. While the signal for V5-tagged ZapE1 protein was virtually eliminated, RNAi was somewhat less efficient in the case of HA-tagged ZapE2 (Fig 3A). Unexpectedly, even tandem depletion of both ZapE paralogues had no effect on cell growth under standard cultivation conditions (Fig 3B).

Next, activities of two respiratory chain complexes were measured using spectrophotometric assays in the double knockdown (Fig 3C). Complex IV was affected most significantly, as its activity almost doubled in the ZapE proteins-depleted cells. A statistically significant upregulation of complex I activity was also observed, although to a lesser extent than that of complex IV. Putative changes in the activities of complexes II and III upon RNAi induction remained below statistical significance (Fig 3C). Prior to the set of measurements, the activities of respiratory complexes II, III, and IV were validated using their specific inhibitors, namely malonate, antimycin A, and KCN, respectively, which caused their near complete inhibition.

Mitochondrial origin of ZapE in eukaryotes

To gain insight into the evolutionary origin of ZapE in eukaryotes, we carried out an extensive phylogenetic analysis, which includes eukaryotic as well as prokaryotic sequences (Fig 4). Outside eukaryotes, we identified ZapE only in two eubacterial groups, Actinobacteria and Proteobacteria. In our first analysis (Fig 4A), sequences formed four well-supported clades: i/ Actinobacteria (ultrafast bootstrap 100 and rapid bootstrap 100), ii/ β-proteobacteria (ultrafast bootstrap 94 and rapid bootstrap 89), iii/ γ-proteobacteria (ultrafast bootstrap 100 and rapid bootstrap 99), and iv/ α-proteobacterial and eukaryotic ZapE (support 100 and 94, respectively).

Eukaryotic sequences form a moderately supported group (ultrafast bootstrap 96, rapid bootstrap 69). The group of eukaryotic ZapE was formed in all preliminary analyses that

53
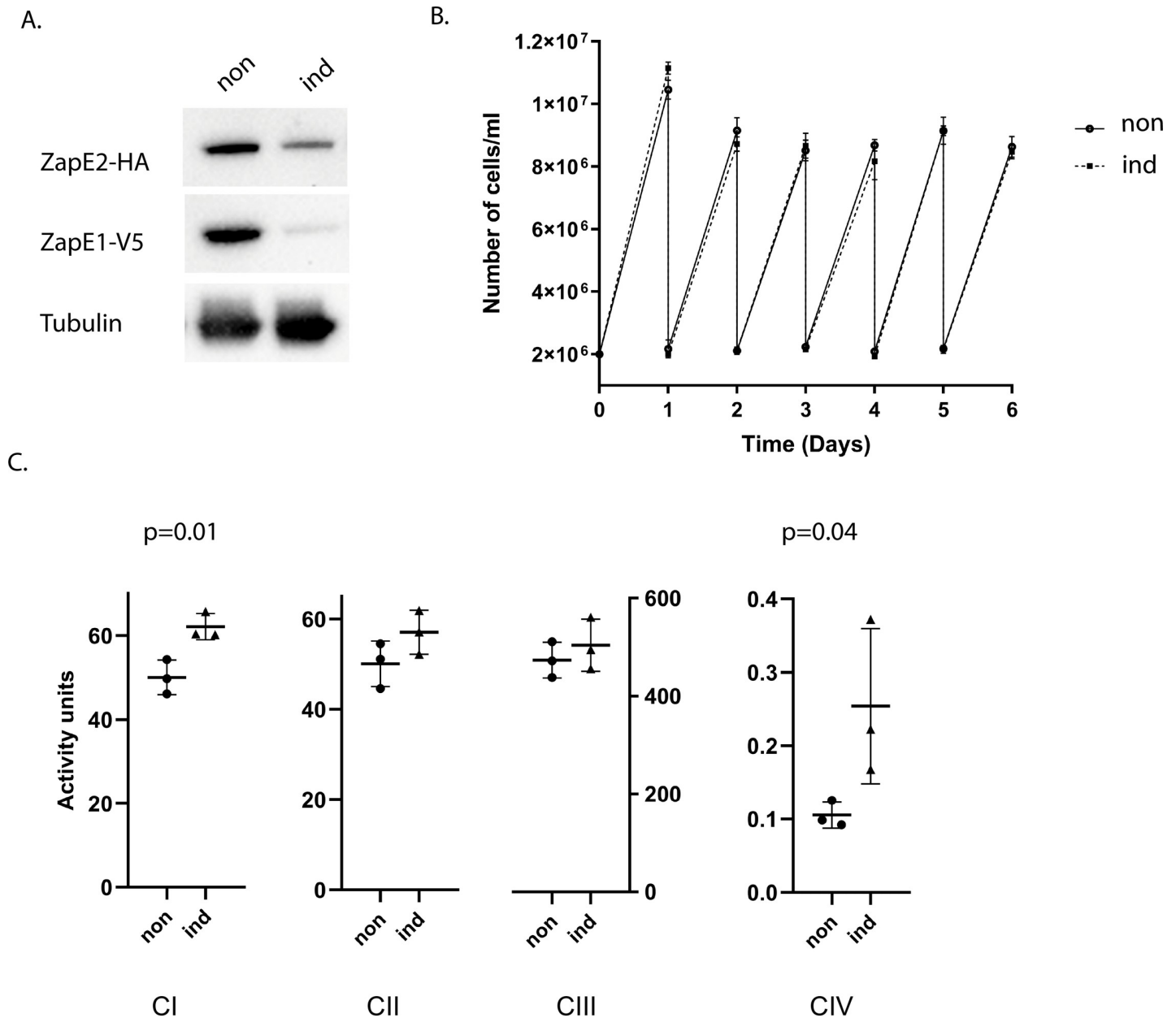
Fig 3. Downregulation of ZapE affects the activities of respiratory complexes. A double knock-down cell line was induced (ind) for six days with doxycycline. Noninduced (non) cells served as control. (**A**) ZapE1 and ZapE2 proteins were detected with α-V5 and α-HA antibodies, respectively, with α -tubulin antibody used as a loading control. (**B**) Cell densities of non-induced (circles) and RNAi-induced cells (squares) are indicated. The experiment was performed in biological triplicate. Error bars represent standard deviations. (**C**) Activities of two respiratory complexes are elevated in induced cells 6 days post-induction. The P-value of unpaired T-test is shown where a statistically significant difference was detected. Means from three independent biological replicates are displayed. Error bars represent standard deviation.

differed in taxon composition and were independently trimmed. Moreover, the same topology was recovered by both IQ-Tree and RAxML, two phylogenetic programs for maximum likelihood analysis. Interestingly, no other ZapE genes were identified outside Actinobacteria, Proteobacteria, and Eukaryota. We also did not find any ZapE in other proteobacterial groups, namely δ- and ε-proteobacteria. Several ZapE sequences present in genomic assemblies of some prokaryotes were shown to be contaminants by reciprocal BLAST or by manual inspection of

54

55
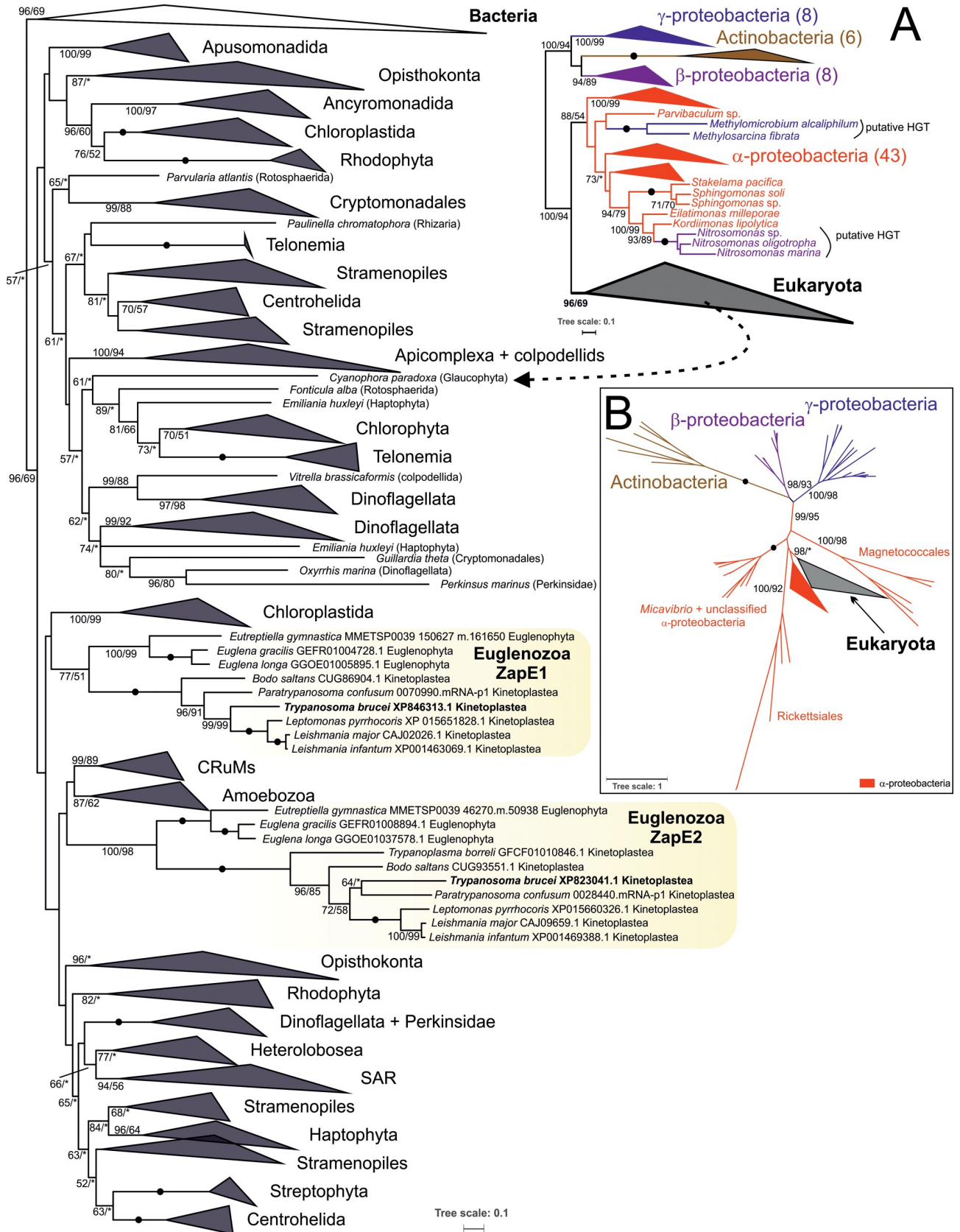
Fig 4. Phylogenetic analysis of ZapE. (A) A eukaryota-focused phylogenetic tree without long-branching α-proteobacterial sequences. The tree was arbitrary rooted using the midpoint-rooting method. Schematic representation showing the position of two proteobacterial clades branching within the clan of α-proteobacterial sequences is on the right, detailed phylogeny of eukaryotic ZapE on the left. A fully resolved tree is deposited as a supplementary S3 Fig. **(B)** Unrooted phylogenetic tree of ZapE that includes long-branching α-proteobacteria (Rickettsiales, Magnetococcales, *Micavibrio*) and shows a close affinity of eukaryotic ZapE to α-proteobacteria. Tree topologies are based on phylogenetic trees computed by the Maximum Likelihood method (LG4X model) in IQ-TREE. Branch supports were assessed by ultrafast bootstrap (N = 1000, IQ-TREE) and rapid bootstrap (N = 500, RAxML). Branch supports > 50% are indicated. A fully resolved tree is shown in S4 Fig.

flanking genomic regions. In all cases, the most probable source of contamination was α-, β- or γ-proteobacteria (S3 Table). Several sequences from *Nitrosomonas* spp. (β-proteobacteria), *Methylomicrobium alcaliphilum* and *Methylsarcina fibrata* (both γ-proteobacteria) that fell within the α-proteobacterial clade most likely represent horizontal gene transfer events (Fig 4A).

To better understand the origin of eukaryotic ZapE, we performed another phylogenetic analysis which also included extremely derived sequences of deep-branching α-proteobacteria (Rickettsiales, Magnetococcales, etc.). This analysis indicates that eukaryotic ZapE originated from Proteobacteria, most likely from a group branching within α-proteobacteria (Fig 4B). We hypothesize that ZapE entered the eukaryotic cell with the proto-mitochondrion and was transferred to the eukaryotic nucleus from the mitochondrial genome before modern eukaryotes diverged from their last common ancestor. Internal phylogeny of eukaryotic ZapE is not well resolved, although we encountered several instances of independent gene duplications—in Euglenozoa, the SAR supergroup, and Chloroplastida.

ZapE co-occurs with respiratory complexes, Oxa1 and mitochondrial genome in eukaryotes

ZapE modulates Z-ring stability during the FtsZ-dependent cell division in the γ-proteobacterium *E. coli*. It is not essential during growth under laboratory conditions, but in the anaerobic conditions or at temperatures over 37°C, the importance of ZapE becomes evident [5]. Although FtsZ-dependent division is present in most of the major groups of Eubacteria as well as in some Archaea, plastids, and mitochondria [27], our analyses show that the taxonomic distribution of ZapE is different. It is present only in a few eubacterial lineages but nearly all eukaryotes. More interestingly, eukaryotes often encode ZapE even when mitochondrial FtsZ-dependent division has been completely lost (Fig 5). The ZapE gene is present in eukaryotes with classical aerobic mitochondria, anaerobically functioning mitochondria (e.g. *Fasciola hepatica* and *Euglena gracilis*) and hydrogen-producing mitochondria (e.g. *Blastocystis hominis* and *Acanthamoeba castellanii*). On the other hand, it is absent in the obligate anaerobes equipped with extremely reduced derivates of mitochondria, such as hydrogenosomes and mitosomes (e.g. *Entamoeba histolytica* and *Giardia intestinalis*).

The presence of ZapE in eukaryotes depends on respiratory chain complexes, especially complex II, and correlates with the retention of the mitochondrial genome and Oxa1. *Cryptosporidium muris*, but not its close relative *C. parvum*, is the sole identified exception, as it retains the ZapE gene and respiratory complex II, while Oxa1 and most probably also the mitochondrial genome are absent [30] (Fig 5).

Discussion

For ZapE, a spectrum of seemingly unrelated functions ranging from apoptosis to cell division has been proposed [5–7,9]. Curiously, it was not yet addressed how a single protein may perform such a plethora of functions. Here we propose a hypothesis, which connects most of the observed phenotypes into a single putative role of the ZapE protein.

56

57
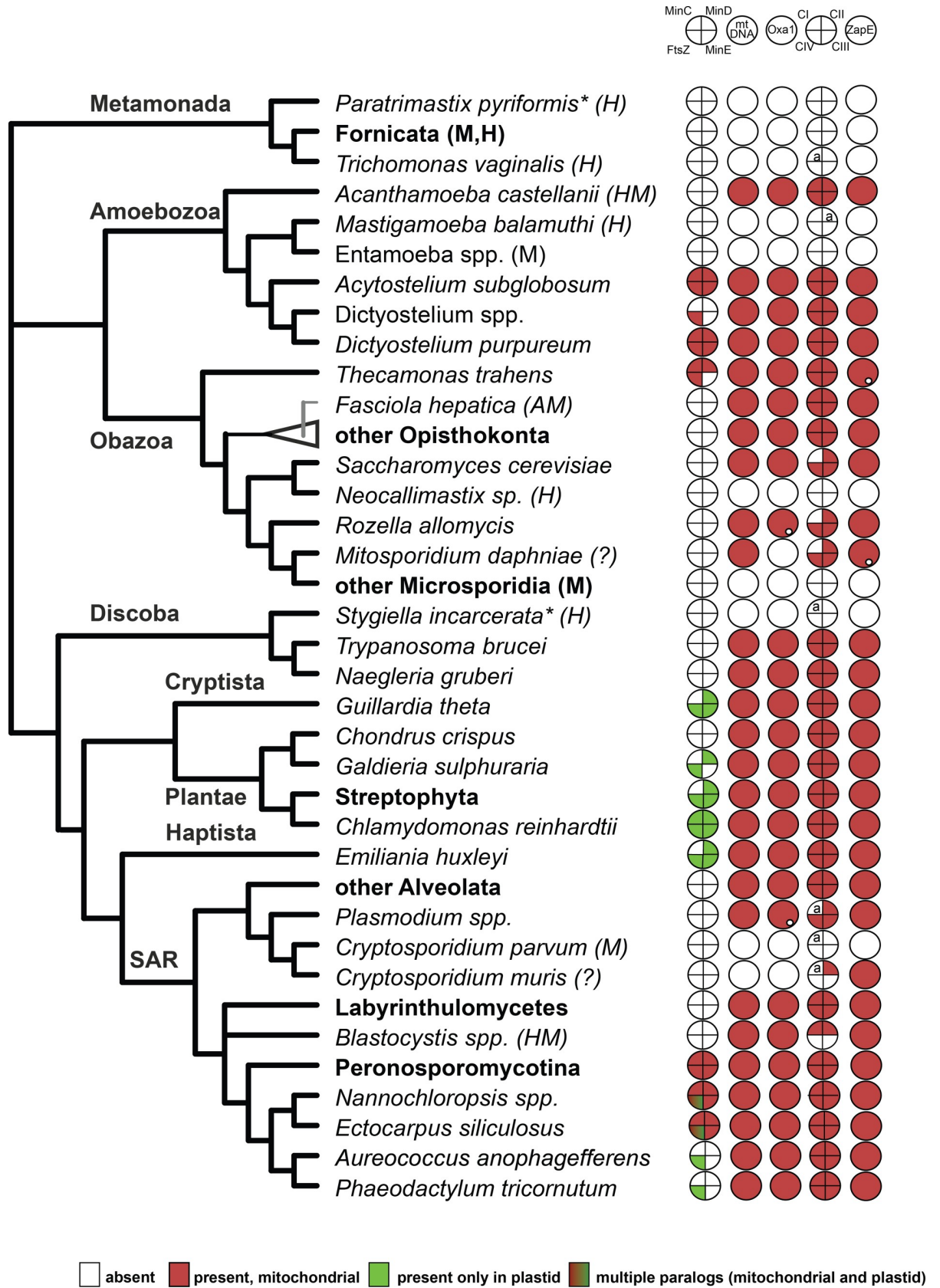
Fig 5. Co-occurrence of ZapE homologs with genes for respiratory complexes, Oxa1, and core FtsZ components. Prediction of mitochondrial/plastidial localization of given proteins based on Mitofates [28] and Deeploc [29] prediction tools is displayed. Small white dots indicate Oxa1 and ZapE proteins with unclear cellular localization. Data about subcellular localization and the presence of Min and FtsZ proteins were acquired from [3]. For the prediction of localization of respiratory complexes, phylogenetic affinity with previously localized proteins was considered. Alb3, plastid homolog of Oxa1, was excluded from the table. Asterisks (*) indicate results based on transcriptomic data; the letter "a" specifies that only some subunits of a given respiratory complex are present. AM, anaerobic mitochondrion; HM, hydrogen-producing mitochondrion; H, hydrogenosome; M, mitosome; (?), unclassified mitochondrion. Collapsed nodes representing higher taxonomic groups are in bold.

One of the three identified putative interaction partners of the ZapE2 paralog in *T. brucei* is Oxa1, a highly conserved protein required for both the insertion of mitochondrially-encoded subunits of the respiratory chain complexes [31], as well as of 28 nuclear-encoded proteins residing in the inner mitochondrial membrane, such as mitochondrial carriers, Atm1, and the sdh3 and sdh4 subunits of complex II [8]. We suggest that in eukaryotes Oxa1 and ZapE operate together, and that a depletion of the latter partner triggers a complex phenotype.

In yeast, only the mitochondrially-encoded subunits (cox1, cox2, and cox3) of complex IV were affected by the ablation of ZapE [4]. This observation can be nicely explained by a functional linkage between ZapE and Oxa1. Furthermore, ZapE was recently identified among the substrates of Oxa1, as its expression was dramatically decreased in the Oxa1 knock-down cells [8]. However, among 28 proteins identified in this survey, ZapE was the only one lacking a transmembrane domain, strongly indicating that it is not an integral protein of the inner mitochondrial membrane. This implies that ZapE is not a substrate of Oxa1 but rather its interaction partner. Finally, ZapE seems to be invariably present in the aerobic mitochondria (Fig 5), a key feature of which is the presence of numerous subunits of the respiratory chain complexes encoded by the mitochondrial genome [32]. Their protein products are inserted with the assistance of the membrane insertase Oxa1 into the corresponding complexes residing in the inner membrane [33]. In contrast, we have noted that ZapE is prominently absent in anaerobes lacking the mitochondrial genome, respiratory complexes, and Oxa1.

In *T. brucei* in which both ZapE paralogs were downregulated, the activities of Oxa1 substrates, in particular complexes I and IV, significantly increased. In yeast it was shown that following the depletion of ZapE, mitochondrially-encoded complex IV subunits cox1, cox2, and cox3 were also slightly elevated, which led the authors to conclude that ZapE facilitates their degradation [4]. As Oxa1 participates in the insertion of several subunits of the respiratory complexes into the inner mitochondrial membrane, we propose that ZapE may negatively regulate the function of Oxa1. We attribute the lack of growth phenotype to the fact that a small fraction of the target proteins escaped RNAi, but alternatively it is possible, that proteins might not be required for normal cell growth under tested conditions. Nevertheless, our findings are consistent with the connection between ZapE/Afg1 and the respiratory chain, which was so far observed only in yeast and humans [4,6].

We then investigated how the interaction of ZapE with Oxa1 could be linked with other described roles of the former protein, which was associated with the FtsZ-dependent bacterial division, as both its overexpression and ablation resulted in elongated bacterial cells [5]. The FtsZ-dependent division system, which comprises at least four components, was retained in the inner membranes of mitochondria and plastids of various eukaryotes, such as *Dictyostelium purpureum*, *Malawimonas californiana*, *Guillardia theta*, *Chondrus crispus* and *Chlorella variabilis* [3]. We explored a possible co-occurrence of ZapE with the FtsZ machinery, but none was found, so it is likely that ZapE participates in different processes. In phylogenetic trees, eukaryotic ZapE homologs form a sister group to α-proteobacterial ZapE, implying their mitochondrial origin, a conclusion further supported by the absence of ZapE in plastids and cyanobacteria (Fig 4). Therefore, the function of ZapE was either completely changed at an

58

early stage of the eukaryotic evolution or, alternatively, YidC (a homolog of mitochondrial Oxa1) is an interaction partner of ZapE also in bacteria. Another conclusion based on the phylogenetic analysis is that the two ZapE paralogs appeared before the radiation of kinetoplastid flagellates. Since the only identified putative interactor of ZapE1 is a protein of unknown function (Tb927.3.4950) with no homology outside the kinetoplastids, it does not provide any information about the function of ZapE1, which will possibly be specific for this group of highly derived protists. However, we cannot rule out a scenario, in which Oxa1 interacts with both paralogs, yet the interaction with ZapE1 is more transient or was disrupted by tagging.

In mammalian cells, the depletion of ZapE resulted in fragmented mitochondria [6], a phenotype reminiscent of that seen during deficiency in respiratory complexes [34]. A key component of the mitochondrial cristae shaping complex (MICOS) [35], Mic10 belongs to the substrates of Oxa1 [8]. Therefore, it is possible that the observed impact of ZapE on mitochondrial and bacterial morphology is just a consequence of the destabilization of respiratory chain components and/or the MICOS complex. Similarly, we can easily link ZapE and Oxa1 with apoptosis, as the respiratory chain complexes are its modulators [36], as well as with proteostasis since Oxa1 is responsible for the insertion of a subset of the mitochondrially- and nuclear-encoded inner membrane proteins [8].

To further validate the use of the BioID2 *in vivo* labelling technique, we turned our attention to IscU and LigK-β, which were used as controls. Frataxin and TbNfu1 (both components of the Fe-S cluster pathway) were among the three proteins exclusively associated with IscU. Quite surprising, however, is the absence of IscS, which was proposed to interact with IscU in trypanosomes [37]. This protein was strongly but equally measured in all bait protein datasets. It is possible that the fusion of IscU with BioID2 disrupted its structure and thus prevented its specific interaction with IscS.

For LigK-β, only Tb927.8.4230 was identified as a putative interacting protein, which was more than 100x enriched compared to other datasets. Importantly, this unknown function protein localized into the antipodal sites of the kDNA disk. Such a highly specific localization is telling. Due to its absence in the large non-catenated pro-kinetoplast DNA of the free-living *Bodo saltans* [38], it is plausible that this protein is involved in the catenation of circular DNAs, so-called minicircles, into the single kDNA network of *T. brucei*. Moreover, by a BLAST search, we found a 184 kDa homolog of Tb927.8.4230 named Tb927.8.4240. Reassuringly, the latter is also present in the proximity of the kDNA disk of trypanosomes, yet most likely does not reside in the antipodal sites (Fig 2B).

Additionally, we tested the suitability of BioID2 for the mapping of the mitochondrial proteins. We identified 117 proteins, all located to the organelle with high confidence. Thus, the BioID2 technique produces a very clean, albeit incomplete *T. brucei* mitochondrial proteome, which is estimated to contain ~1,200 proteins [13,14].

In conclusion, we have successfully established the BioID2 technique in *T. brucei*. The specificity of this technique was proven by a high purity of the mitochondrial protein dataset. Using this technique, we were able to demonstrate that Oxa1 is a putative interaction partner of the ZapE homolog also outside the Opisthokonta, which allowed us for the first time to identify common features of the ZapE-related phenotypes in two eukaryotic supergroups. More research is, however, needed to confirm the nature of the interaction between Oxa1 and ZapE proteins.

Experimental procedures

Preparation of cell lines

T. brucei procyclic stage SmOx cell line [39] was grown in SDM79 supplemented with 10% fetal bovine serum for most of the experiments. Alternatively, the glucose-poor SDM80

medium [40] was used for studies of the phenotype triggered by ZapE RNAi-based depletion. Various proteins were *in situ*-tagged by a recently developed PCR-based transfection protocol described elsewhere [14]. As a PCR template, different versions of the pPOT plasmids were used (S4 Table). For instance, proteins identified by the proteomic analysis and ZapE homologs were C-terminally V5-tagged using a previously modified pPOTv4_Hyg vector [41]. In order to *in situ* fuse various proteins with BioID2, pPOT7-Blast-mNG plasmid (kindly provided by Samuel Dean) was modified by replacing the mNG gene with the BioID2 gene so that the protein of interest is linked with the BioID2 protein by an 8 nm-long glycine-serine repeats-containing linker (S1 Fig). Furthermore, long hairpin RNAi constructs were assembled for the ZapE1 and ZapE2 double RNAi knock-down by cloning two 450 nt-long regions into the pTrypSon plasmid by the Gibson assembly protocol as described previously [26]. Note that the pTrypSon plasmid with long hairpin ZapE2 was further modified by exchanging the phleomycin resistance with that of blasticidin. In the double RNAi knock-down cells, ZapE1 and ZapE2 were *in situ*-tagged with V5 tag and HA tag, respectively, and then both proteins were targeted using the dedicated pTrypSon plasmids. This strategy resulted in a double knockdown cell line modified by four plasmids with four different resistances (DKD cell line). A similar strategy was employed to create single knockdown cell lines (S4 Table).

RNAi phenotype analysis

Cultures were grown in triplicate in the presence or absence of doxycycline. Cell density was counted using the Beckman Coulter Z2 Cell and Particle Counter every 24 hours and cells were subsequently diluted to 2×10^6 cells/ml, maintaining them in the exponential phase of growth. For Western blot analysis, monoclonal α-V5, α-HA (Life Technologies), α-tubulin antibodies (Sigma-Aldrich), and secondary HRP-conjugated α-mouse and α-rabbit IgG antibodies (Sigma-Aldrich) were used and the signal was visualized by Clarity Western ECL Blotting Substrate (Bio-Rad).

Enzymatic assays

For enzymatic assays, hypotonically isolated mitochondria from 5×10^8 cells were lysed on ice for 1 hour in 2% (w/v) dodecyl maltoside and 0.4 M aminocaproic acid. Upon centrifugation at 24.400 g for 30 min at 4°C supernatant was used for activity measurements. All measurements were carried in biological triplicates.

Activities of respiratory enzymes (complexes I–IV) were measured by a spectrophotometric approach described previously [42]. Briefly, NADH dehydrogenase activity (complex I) was measured in 1 ml of NDH buffer (50 mM KPi, pH 7.5; 1 mM EDTA, pH 8.5; 0.2 mM KCN) with the addition of 0.1 mM NADH. The reaction was started by the addition of 2 μM coenzyme Q_2 and followed at 340 nm for 3 min. For succinate dehydrogenase activity (complex II) mitochondrial lysate was pre-incubated with SDH buffer (25 mM KPi, pH 7.2; 5 mM MgCl$_2$; 20 mM sodium succinate) at 25°C for 10 min. Next, antimycin A, rotenone, KCN, and 2,6-dichlorophenolindophenol (DPIP) were separately added to a final concentration of 1.8 mM, 5 mM, 2 mM, and 50 μM, respectively. The reaction itself was started by the addition of coenzyme Q_2 to a final concentration of 65 μM and monitored at 600 nm for 5 min.

Cytochrome *c* reductase activity (complex III) was measured in 1 ml of QCR buffer (40 mM NaPi, pH 7.4; 0.5 mM EDTA; 20 mM sodium malonate; 50 μM cytochrome c; 0.005% [w/v] dodecyl maltoside). Simultaneously, 2 μl of mitochondrial lysate and 20 μM 2,3-dimethoxy-5-methyl-6-dodecyl-1,4-benzoquinol (DBH) were added. DBH was prepared by reduction of decylubiquinone as described elsewhere [43]. Reaction was followed at 550 nm for 1 min. Cytochrome *c* oxidase activity was measured in 1 ml of COX buffer (40 mM NaPi, pH 7.4; 0.5

60

mM EDTA; 20 μM cytochrome c; 30 μM ascorbic acid; 0.005% [w/v] dodecyl maltoside; solution was incubated overnight to oxidize surplus of ascorbic acid).

The mitochondrial lysate was added to the reaction buffer and monitored at 550 nm for 10 min. The unit (U) of appropriate activity is defined as an amount of enzyme required for the conversion of 1 nmol of NADH/min for NADH dehydrogenase; 1 nmol of DPIP/min for succinate dehydrogenase; 1 μmol of cytochrome *c* for both cytochrome *c* reductase and cytochrome *c* oxidase activities. Specific activities for all measurements were calculated as U per mg of mitochondrial proteins. To confirm the specificity of each enzymatic measurement, inhibitors of respiratory complexes were used. 1 mM sodium malonate, antimycin A at final concentration 300 ng/ml, and 100 μM KCN selectively inhibited the activities of complexes II, III, and IV, respectively. Inhibition of complex I was not tested since its contribution to overall NADH dehydrogenase activity within the cell is only partial [44].

Immunofluorescence

Cells were fixed with 4% paraformaldehyde in phosphate buffered saline (PBS) and then permeabilized with 0.1% Triton X-100 in PBS. After blocking (1% BSA; 0.03% Triton X-100 in PBS) monoclonal rabbit α-V5 (Sigma-Aldrich), and monoclonal mouse α-mtHsp70 (kindly provided by Alena Zíková) was used followed by α-rabbit Alexa Fluor 488 and α-mouse Alexa Fluor 555 staining (Life Technologies). DNA was stained with ProLong™Gold antifade reagent with 4',6-diamidine-2'-phenylindole dihydrochloride (DAPI) (Molecular Probes). The immunofluorescence assay was performed using a Zeiss microscope Axioplan 2 equipped with an Olympus DP73 digital camera.

Proteomic identification by BioID2

Prior to the experiment, SDM79 media was supplemented overnight with 100 μM biotin. Then mitochondrion-enriched fraction was obtained from 3×10^9 *T. brucei* procyclic cells by hypotonic lysis as described previously [42]. Mitochondria-enriched pellets were then resuspended in 1.8 ml of Boiling Buffer (1% SDS; 1 mM EDTA; 50 mM Tris; pH 7.4) for 10 min at 80°C. The dissolved samples were subsequently 10 times diluted in Incubation Buffer (150 mM NaCl; 5 mM EDTA; 1% Triton X-100; 50 mM Tris; pH 7.4) supplemented with Complete protease EDTA-free inhibitors (Roche). 0.5 mg of Dynabeads was added per sample (MyOne™ Streptavidin C1) and placed on a rocking platform for 2 hours at room temperature, following which the samples were stored overnight at 4°C. The Dynabeads were then washed three times separately with 1.5 ml of Boiling Buffer and 1.5 ml of Incubation Buffer. Dry Dynabeads were stored at -80°C prior to the proteomic analysis. The detailed protocol was deposited on protocol.io website (http://dx.doi.org/10.17504/protocols.io.bdrri556).

Mass spectrometry

Proteins immobilized on Dynabeads were trypsin-digested and nanoflow liquid chromatography was used for separation of resulting peptides. Next, analysis of samples by tandem mass spectrometry (nLC-MS2) on a Thermo Orbitrap Fusion (q-OT-qIT, Thermo Scientific) instrument was performed as described elsewhere [45]. The mass spectrometry data were analyzed and quantified using MaxQuant software (version 1.5.3.8) [46] with false discovery rate set to 1% for both proteins and peptides and we specified a minimum length of seven amino acids. The Andromeda search engine was used for the MS/MS spectra search against the *Trypanosoma brucei* (downloaded from Uniprot, November 2018). Enzyme specificity was set as C-terminal to arginine and lysine, also allowing cleavage at proline bonds and a maximum of two missed cleavages. Dithiomethylation of cysteine was selected as fixed modification and N-

61

terminal protein acetylation and methionine oxidation as variable modifications. The "match between runs" feature of MaxQuant was used to transfer identifications to other LC-MS/MS runs based on their masses and retention time (maximum deviation 0.7 min) and this was also used in quantification experiments. Then normalized intensity values were further processed using the Perseus software 1.5.2.4 [47]. Only proteins identified exclusively in each bait protein dataset or statistically significantly enriched were considered as putative interacting proteins. Exclusive identification is here defined as a situation where a given protein was measured in all three replicates of bait protein but was absent in all three control replicates. Statistically significant enrichment applies here for cases where a given protein was identified in all replicates for both, bait protein, and control, yet the difference between these groups was statistically significant. For statistical analysis, a Two-sample test with S0 parameter set to 0.1 was used as described elsewhere [47]. Mass spectrometry data have been deposited in the ProteomeXchange Consortium via the PRIDE [48] partner repository with the dataset identifier PXD014426.

Bioinformatic analyses

To identify ZapE homologs in both eukaryotes and prokaryotes, we prepared an initial dataset via the BLASTp algorithm with human and *E. coli* ZapE sequences as queries against NCBI non-redundant protein "nr" database (https://www.ncbi.nlm.nih.gov/). Fifty best BLASTp hits for each query were kept, merged to a single dataset, deduplicated, and aligned by MAFFT version 7 (ginsi algorithm). The alignment was subsequently used to build Hidden Markov Model for a sensitive homolog search by HMMER3 software [49] to get candidate ZapE proteins from 100 strategically sampled eukaryotic genomes and transcriptomes. In addition, we also searched taxonomically restricted NCBI non-redundant protein databases for Archaea and Eubacteria, the latter clustered at 70% threshold (https://toolkit.tuebingen.mpg.de; psi-blast). Presence in deep-branching α-proteobacteria was specifically tested by search in available metagenomic and genomic data. All candidate protein sequences were then aligned by MAFFT version 7 in auto mode [50] and trimmed manually. Only well-aligned complete or nearly complete sequences were retained for further analysis. To remove identical sequences and to decrease the number of sequences in overrepresented lineages, preliminary phylogenetic analysis was performed in IQ-TREE multicore version 1.6.10 [51] with 1000 ultrafast bootstraps under the LG4X substitution model suggested by ModelFinder [52]. Subsequently, several rounds of reciprocal BLASTp and phylogenetic analyses were performed to remove contaminants (S3 Table). The two versions of the final protein dataset (Fig 4A and 4B) were aligned by MAFFT version 7, G-INS-i method with BLOSUM30 scoring matrix, and unaligned level 0.3 or 0.6, respectively. Both were visually inspected and trimmed manually. Maximum-Likelihood phylogenetic analysis was inferred with IQ-TREE as described above. Branch supports were estimated by rapid bootstrapping in RAxML version 8.12.11 [53] (LG4X model, 500 replicates) and ultrafast bootstrapping in IQ-TREE with activated "bnni" option to reduce the risk of overestimating branch supports (1,000 replicates). Bootstrap replicates were mapped on the best IQ-TREE topology and visualized by CorelDRAW Home & Student Suite X8.

Supporting information

S1 Fig. Scheme of the pPOT7_BioID2 plasmid, *in situ* tagging strategy, and the full sequence of BioID2 protein with the linker. The pPOT7_BioID2 plasmid served as a template for long-primer PCR. These primers have 20 bp homology to the plasmid at its 3′ end and 80 bp homology to the *T. brucei* genomic sequence at its 5′ end. This 80 bp homology

regions facilitate specific recombination. The whole PCR product was then inserted in frame before the stop codon of the gene. The sequence of the BioID2 tag with GS linker, which replaced the stop codon of the gene, is also shown. PFR 3', Paraflagellar rod protein 2 terminator; ald 5' and 3', aldolase promotor and terminator; bsr, blasticidin S deaminase gene.
(TIF)

S2 Fig. Depletion of ZapE1 and ZapE2 protein expression and growth phenotype in single knockdown procyclic cell lines. Single knockdown ZapE1 and ZapE2 RNAi cells were treated with doxycycline for six days. (**A, C**) Protein levels detected by Western blot analysis. Non, without doxycycline; Ind, with doxycycline. α-tubulin antibody serves as a loading control. (**B, D**) Growth rates of induced and uninduced cell lines. The experiment was performed in biological triplicate. Error bars represent standard deviations.
(TIF)

S3 Fig. Fully resolved rooted Eukaryota-focused phylogenetic tree of ZapE without long-branching α-proteobacterial sequences.
(TIF)

S4 Fig. Fully resolved unrooted phylogenetic tree of ZapE that includes long-branching α-proteobacteria.
(PDF)

S1 Table. Mitochondrial proteins of *T. brucei* identified by BioID2 approach. Column E shows the probability that a given protein has the mitochondrial import signal detected with the Mitofates online prediction tool. Columns F and G display whether a given protein was previously experimentally localized in the mitochondrion (TrypTag) or was present in the Tom40-based depletome. Columns J to Q show enrichment of a given protein in the bait protein datasets compare to IscU_leader_BioID2 negative control dataset. +, statistical significance (Two-sample test) of enrichment in the previous column. Columns R to AF display a Log_2-transformed intensities for a given protein in a specific dataset. Lower case letters a, b, and c represent each replicate. Three proteins highlighted in blue were not present in the Tom40-based depletome, nor were they experimentally localized.
(XLSX)

S2 Table. Full list of proteins significantly enriched by BioID2 labelling. Columns B and C: predicted function and e-value based on BLASTp algorithm against the NCBI non-redundant protein "nr" database (https://www.ncbi.nlm.nih.gov/against) with a parameter to exclude kinetoplastids. Columns D to G: statistically significant enrichment of a given protein against other bait protein datasets. "-"demarks when enrichment could not be calculated (*e.g.* when a given protein was measured only for a bait protein). Column M: the probability that a protein has the mitochondrial import signal detected with the Mitofates online prediction tool. Columns N and O display whether the protein was previously experimentally localized in the mitochondrion (TrypTag) or was present in the Tom40-based depletome. Columns P to AD display logarithm of two of measured intensities for a given protein in a specific dataset. Lower case letters a, b, and c represent each replicate. Proteins enriched on average less than three times are highlighted in blue.
(XLSX)

S3 Table. Complete list of contaminant sequences, which were filtered out from the phylogenetic analysis. Reverse BLAST results for each gene are displayed in column D.
(XLSX)

63

S4 Table. Used plasmids and cell lines.
(XLSX)

S1 Raw Images.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Jan Pyrih, Julius Lukeš.

Data curation: Jan Pyrih.

Funding acquisition: Julius Lukeš.

Investigation: Jan Pyrih, Vendula Rašková, Ingrid Škodová-Sveráková, Tomáš Pánek.

Methodology: Jan Pyrih.

Resources: Julius Lukeš.

Supervision: Julius Lukeš.

Writing – original draft: Jan Pyrih, Julius Lukeš.

References

1. Friedman JR, Nunnari J. Mitochondrial form and function. Nature. 2014. pp. 335–343. https://doi.org/10.1038/nature12985 PMID: 24429632

2. Horváthová L, Žárský V, Pánek T, Derelle R, Pyrih J, Motyčková A, et al. Ancestral mitochondrial protein secretion machinery. bioRxiv. 2019. https://doi.org/10.1101/790865

3. Leger MM, Petrů M, Žárský V, Eme L, Vlček Č, Harding T, et al. An ancestral bacterial division system is widespread in eukaryotic mitochondria. Proc Natl Acad Sci. 2015; 112: 10239–10246. https://doi.org/10.1073/pnas.1421392112 PMID: 25831547

4. Khalimonchuk O, Bird A, Winge DR. Evidence for a pro-oxidant intermediate in the assembly of cytochrome oxidase. J Biol Chem. 2007; 282: 17442–17449. https://doi.org/10.1074/jbc.M702379200 PMID: 17430883

5. Marteyn BS, Karimova G, Fenton AK, Gazi AD, West N, Touqui L, et al. ZapE is a novel cell division protein interacting with FtsZ and modulating the Z-Ring dynamics. mBio. 2014; 5: e00022–14. https://doi.org/10.1128/mBio.00022-14 PMID: 24595368

6. Cesnekova J, Rodinova M, Hansikova H, Houstek J, Zeman J, Stiburek L. The mammalian homologue of yeast Afg1 ATPase (lactation elevated 1) mediates degradation of nuclear-encoded complex IV subunits. Biochem J. 2016; 473: 797–804. https://doi.org/10.1042/BJ20151029 PMID: 26759378

7. Cesnekova J, Spacilova J, Hansikova H, Houstek J, Zeman J, Stiburek L. LACE1 interacts with p53 and mediates its mitochondrial translocation and apoptosis. Oncotarget. 2016; 7: 47687–47698. https://doi.org/10.18632/oncotarget.9959 PMID: 27323408

8. Stiller SB, Höpker J, Oeljeklaus S, Schütze C, Schrempp SG, Vent-Schmidt J, et al. Mitochondrial OXA translocase plays a major role in biogenesis of inner-membrane proteins. Cell Metab. 2016; 23: 901–908. https://doi.org/10.1016/j.cmet.2016.04.005 PMID: 27166948

9. Germany EM, Zahayko N, Huebsch ML, Fox JL, Prahlad V, Khalimonchuk O. The AAA ATPase Afg1 preserves mitochondrial fidelity and cellular health by maintaining mitochondrial matrix proteostasis. J Cell Sci. 2018; 131: jcs219956. https://doi.org/10.1242/jcs.219956 PMID: 30301782

64

10. Read LK, Lukeš J, Hashimi H. Trypanosome RNA editing: The complexity of getting U in and taking U out. Wiley Interdisciplinary Reviews: RNA. 2016. pp. 33–51. https://doi.org/10.1002/wrna.1313 PMID: 26522170

11. Verner Z, Basu S, Benz C, Dixit S, Dobáková E, Faktorová D, et al. Malleable mitochondrion of *Trypanosoma brucei*. Int Rev Cell Mol Biol. 2015; 315: 73–151. https://doi.org/10.1016/bs.ircmb.2014.11.001 PMID: 25708462

12. Zíková A, Verner Z, Nenarokova A, Michels PAM, Lukeš J. A paradigm shift: The mitoproteomes of procyclic and bloodstream *Trypanosoma brucei* are comparably complex. PLoS Pathog. 2017; 13: e1006679. https://doi.org/10.1371/journal.ppat.1006679 PMID: 29267392

13. Peikert CD, Mani J, Morgenstern M, Käser S, Knapp B, Wenger C, et al. Charting organellar importomes by quantitative mass spectrometry. Nat Commun. 2017; 8: 15272. https://doi.org/10.1038/ncomms15272 PMID: 28485388

14. Dean S, Sunter J, Wheeler RJ, Hodkinson I, Gluenz E, Gull K. A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. Open Biol. 2015; 5: 140197. https://doi.org/10.1098/rsob.140197 PMID: 25567099

15. Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, et al. Efficient proximity labeling in living cells and organisms with TurboID. Nature Biotechnology. 2018. pp. 880–898. https://doi.org/10.1038/nbt.4201 PMID: 30125270

16. Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol. 2012; 196: 801–810. https://doi.org/10.1083/jcb.201112098 PMID: 22412018

17. Morriswood B, Havlicek K, Demmel L, Yavuz S, Sealey-Cardona M, Vidilaseris K, et al. Novel bilobe components in *Trypanosoma brucei* identified using proximity-dependent biotinylation. Eukaryot Cell. 2013; 12: 356–367. https://doi.org/10.1128/EC.00326-12 PMID: 23264645

18. Dang HQ, Zhou Q, Rowlett VW, Hu H, Lee KJ, Margolin W, et al. Proximity interactions among basal body components in *Trypanosoma brucei* identify novel regulators of basal body biogenesis and inheritance. MBio. 2017; 8: e02120–16. https://doi.org/10.1128/mBio.02120-16 PMID: 28049148

19. Kim DI, Jensen SC, Noble KA, Kc B, Roux KH, Motamedchaboki K, et al. An improved smaller biotin ligase for BioID proximity labeling. Mol Biol Cell. 2016; 27: 1188–1196. https://doi.org/10.1091/mbc.E15-12-0844 PMID: 26912792

20. Long S, Jirků M, Ayala FJ, Lukeš J. Mitochondrial localization of human frataxin is necessary but processing is not for rescuing frataxin deficiency in *Trypanosoma brucei*. Proc Natl Acad Sci. 2008; 105: 13468–13473. https://doi.org/10.1073/pnas.0806762105 PMID: 18768799

21. Downey N, Hines JC, Sinha KM, Ray DS. Mitochondrial DNA ligases of *Trypanosoma brucei*. Eukaryot Cell. 2005; 4: 765–774. https://doi.org/10.1128/EC.4.4.765-774.2005 PMID: 15821136

22. Acestor N, Zíková A, Dalley RA, Anupama A, Panigrahi AK, Stuart KD. *Trypanosoma brucei* mitochondrial respiratome: composition and organization in procyclic form. Mol Cell Proteomics. 2011; 10: M110.006908. https://doi.org/10.1074/mcp.M110.006908 PMID: 21610103

23. Saada A, Vogel RO, Hoefs SJ, van den Brand MA, Wessels HJ, Willems PH, et al. Mutations in NDUFAF3 (C3ORF60), encoding an NDUFAF4 (C6ORF66)-interacting Complex I assembly protein, cause fatal neonatal mitochondrial disease. Am J Hum Genet. 2009; 84: 718–727. https://doi.org/10.1016/j.ajhg.2009.04.020 PMID: 19463981

24. Cai K, Frederick RO, Tonelli M, Markley JL. Interactions of iron-bound frataxin with ISCU and ferredoxin on the cysteine desulfurase complex leading to Fe-S cluster assembly. J Inorg Biochem. 2018; 183: 107–116. https://doi.org/10.1016/j.jinorgbio.2018.03.007 PMID: 29576242

25. Benz C, Kovářová J, Králová-Hromadová I, Pierik AJ, Lukeš J. Roles of the Nfu Fe–S targeting factors in the trypanosome mitochondrion. Int J Parasitol. 2016; 46: 641–651. https://doi.org/10.1016/j.ijpara.2016.04.006 PMID: 27181928

26. McAllaster MR, Sinclair-Davis AN, Hilton NA, de Graffenried CL. A unified approach towards *Trypanosoma brucei* functional genomics using Gibson assembly. Mol Biochem Parasitol. 2016; 210: 13–21. https://doi.org/10.1016/j.molbiopara.2016.08.001 PMID: 27496178

27. Margolin W. FtsZ and the division of prokaryotic cells and organelles. Nat Rev Mol Cell Biol. 2005. pp. 862–871. https://doi.org/10.1038/nrm1745 PMID: 16227976

28. Fukasawa Y, Tsuji J, Fu S-C, Tomii K, Horton P, Imai K. MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. Mol Cell Proteomics. 2015; 14: 1113–1126. https://doi.org/10.1074/mcp.M114.043083 PMID: 25670805

29. Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics. 2017; 33: 3387–3395. https://doi.org/10.1093/bioinformatics/btx431 PMID: 29036616

65

30. Widmer G, Sullivan S. Genomics and population biology of Cryptosporidium species. Par Immunol. 2012. pp. 61–71. https://doi.org/10.1111/j.1365-3024.2011.01301.x PMID: 21595702

31. Kolli R, Soll J, Carrie C. Plant mitochondrial inner membrane protein insertion. Int J Mol Sci. 2018. pp. 1188–96. https://doi.org/10.3390/ijms19020641 PMID: 29495281

32. Smith DR, Keeling PJ. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. Proc Natl Acad Sci. 2015; 112: 10177–10184. https://doi.org/10.1073/pnas.1422049112 PMID: 25814499

33. Herrmann JM, Bonnefoy N. Protein export across the inner membrane of mitochondria: The nature of translocated domains determines the dependence on the Oxa1 translocase. J Biol Chem. 2004; 279: 2507–2512. https://doi.org/10.1074/jbc.M310468200 PMID: 14593095

34. Koopman WJH, Visch H-J, Verkaart S, van den Heuvel LWPJ, Smeitink JAM, Willems PHGM. Mitochondrial network complexity and pathological decrease in complex I activity are tightly correlated in isolated human complex I deficiency. Am J Physiol Physiol. 2005; 289: C881–C890. https://doi.org/10.1152/ajpcell.00104.2005 PMID: 15901599

35. Kaurov I, Vancová M, Schimanski B, Cadena LR, Heller J, Bílý T, et al. The diverged Trypanosome MICOS complex as a hub for mitochondrial cristae shaping and protein import. Curr Biol. 2018; 28: 3393–3407. https://doi.org/10.1016/j.cub.2018.09.008 PMID: 30415698

36. Kwong JQ, Henning MS, Starkov AA, Manfredi G. The mitochondrial respiratory chain is a modulator of apoptosis. J Cell Biol. 2007; 179: 1163–1177. https://doi.org/10.1083/jcb.200704059 PMID: 18086914

37. Paris Z, Changmai P, Rubio MAT, Zíková A, Stuart KD, Alfonzo JD, et al. The Fe/S cluster assembly protein Isd11 is essential for tRNA thiolation in *Trypanosoma brucei*. J Biol Chem. 2010; 285: 22394–22402. https://doi.org/10.1074/jbc.M109.083774 PMID: 20442400

38. Lukeš J, Guilbride DL, Votýpka J, Zíková A, Benne R, Englund PT. Kinetoplast DNA network: Evolution of an improbable structure. Eukaryot Cell. 2002. pp. 495–502. https://doi.org/10.1128/ec.1.4.495-502.2002 PMID: 12455998

39. Poon SK, Peacock L, Gibson W, Gull K, Kelly S. A modular and optimized single marker system for generating *Trypanosoma brucei* cell lines expressing T7 RNA polymerase and the tetracycline repressor. Open Biol. 2012; 2: 110037. https://doi.org/10.1098/rsob.110037 PMID: 22645659

40. Coustou V, Biran M, Breton M, Guegan F, Rivière L, Plazolles N, et al. Glucose-induced remodeling of intermediary and energy metabolism in procyclic *Trypanosoma brucei*. J Biol Chem. 2008; 283: 16343–16354. https://doi.org/10.1074/jbc.M709592200 PMID: 18430732

41. Peña-Diaz P, Vancová M, Resl C, Field MC, Lukeš J. A leucine aminopeptidase is involved in kinetoplast DNA segregation in *Trypanosoma brucei*. PLoS Pathog. 2017; 13: e1006310. https://doi.org/10.1371/journal.ppat.1006310 PMID: 28388690

42. Verner Z, Čermáková P, Škodová I, Kováčová B, Lukeš J, Horváth A. Comparative analysis of respiratory chain and oxidative phosphorylation in *Leishmania tarentolae*, *Crithidia fasciculata*, *Phytomonas serpens* and procyclic stage of *Trypanosoma brucei*. Mol Biochem Parasitol. 2014; 193: 55–65. https://doi.org/10.1016/j.molbiopara.2014.02.003 PMID: 24556248

43. Trumpower BL, Edwards CA. Purification of a reconstitutively active iron-sulfur protein (oxidation factor) from succinate. cytochrome c reductase complex of bovine heart mitochondria. J Biol Chem. 1979; 254: 8697–706. PMID: 224062

44. Verner Z, Čermáková P, Škodová I, Kriegová E, Horváth A, Lukeš J. Complex I (NADH:ubiquinone oxidoreductase) is active in but non-essential for procyclic *Trypanosoma brucei*. Mol Biochem Parasitol. 2011; 175: 196–200. https://doi.org/10.1016/j.molbiopara.2010.11.003 PMID: 21074578

45. Černá M, Kuntová B, Talacko P, Stopková R, Stopka P. Differential regulation of vaginal lipocalins (OBP, MUP) during the estrous cycle of the house mouse. Sci Rep. 2017; 7: 11674. https://doi.org/10.1038/s41598-017-12021-2 PMID: 28916783

46. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008; 26: 1367–1372. https://doi.org/10.1038/nbt.1511 PMID: 19029910

47. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature Methods. 2016. pp. 731–740. https://doi.org/10.1038/nmeth.3901 PMID: 27348712

48. 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. https://doi.org/10.1093/nar/gkv1145 PMID: 26527722

49. Eddy SR. A probabilistic model of local sequence alignment that simplifies statistical significance estimation. PLoS Comput Biol. 2008; 4: e1000069. https://doi.org/10.1371/journal.pcbi.1000069 PMID: 18516236

66

50. Katoh K, Standley DM. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. Bioinformatics. 2016; 32: 1933–1942. https://doi.org/10.1093/bioinformatics/btw108 PMID: 27153688

51. Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018; 35: 518–522. https://doi.org/10.1093/molbev/msx281 PMID: 29077904

52. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat Methods. 2017; 14: 587–589. https://doi.org/10.1038/nmeth.4285 PMID: 28481363

53. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30: 1312–1313. https://doi.org/10.1093/bioinformatics/btu033 PMID: 24451623

67

Supplementary Materials
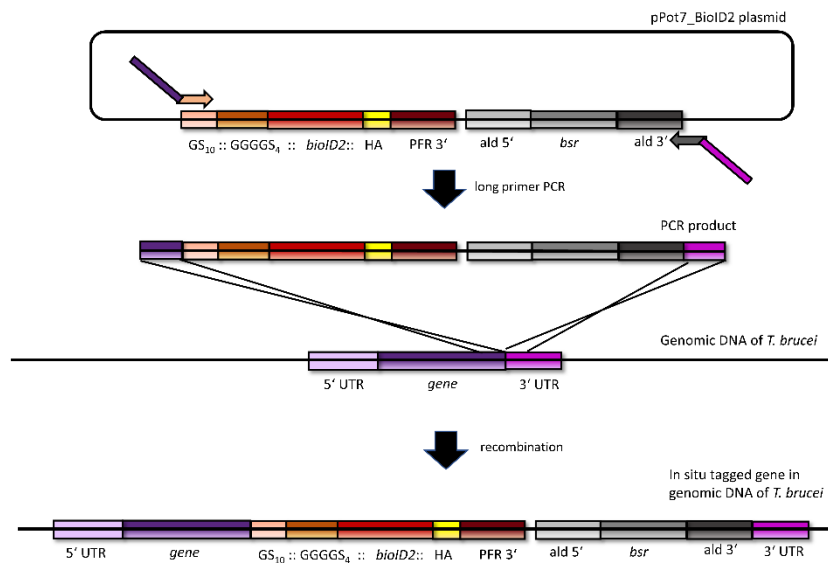
Exact sequence of linker + BioID2 tag which was exchanged instead of stop codon of the bait proteins.

GSGSGSGSGSGSKLDIGGGSGGGGSGGGGSGGGGSDIFKNLIWLKEVDSTQERL
KEWNVSYGTALVADRQTKGRGGLGRKWLSQEGGLYFSFLLNPKEFENLLQLPLV
LGLSVSEALEEITEIPFSLKWPNDVYFQEKKVSGVLCELSKDKLIVGIGINVNQREIP
EEIKDRATTLYEITGKDWDRKEVLLKVLKRISENLKKFKEKSFKEFKGKIESKMLYLG
EEVKLLGEGKITGKLVGLSEKGGALILTEEGIKEILSGEFSLRRS**YPYDVPDYA***

Linker formed of glycine-serine repeats :: BioID2 :: **HA tag**

S1 Fig: Scheme of the pPOT7_BioID2 plasmid, *in situ* tagging strategy, and the full sequence of BioID2 protein with the linker. The pPOT7_BioID2 plasmid served as a template for long-primer PCR. These primers have 20 bp homology to the plasmid at its 3′ end and 80 bp homology to the *T. brucei* genomic sequence at its 5′ end. This 80 bp homology regions facilitate specific recombination. The whole PCR product was then inserted in frame before the stop codon of the gene. The sequence of the BioID2 tag with GS linker, which replaced the stop codon of the gene, is also shown. PFR 3', Paraflagellar rod protein 2 terminator; ald 5' and 3', aldolase promotor and terminator; bsr, blasticidin S deaminase gene.

A.

B.

ZapE1 KD

C.

D.

ZapE2 KD

S2 Fig: Depletion of ZapE1 and ZapE2 protein expression and growth phenotype in single knockdown procyclic cell lines. Single knockdown ZapE1 and ZapE2 RNAi cells were treated with doxycycline for six days. (**A, C**) Protein levels detected by Western blot analysis. Non, without doxycycline; Ind, with doxycycline. α-tubulin antibody serves as a loading control. (**B, D**) Growth rates of induced and uninduced cell lines. The experiment was performed in biological triplicate. Error bars represent standard deviations.

S3 Fig: Fully resolved rooted Eukaryota-focused phylogenetic tree of ZapE without long-branching α-proteobacterial sequences.

0.3

S4 Fig: Fully resolved unrooted phylogenetic tree of ZapE that includes long-branching α-proteobacteria.

Chapter Summary

We have analysed the functions of ZapE in procyclic stage of *Trypanosoma brucei*, using a newly developed proximity-dependent biotinylation approach (BioID2), with the following outcomes:

1. Identification of ZapE-related phenotypes in two eukaryotic supergroups.

2. Identification of the inner mitochondrial membrane insertase Oxa1 among three putative interacting partners of ZapE, which is present in two paralogs.

3. The distribution of mitochondrial ZapE is restricted only to organisms with Oxa1, respiratory complexes, and a mitochondrial genome.

4. RNAi-mediated depletion of both ZapE paralogs likely affected the function of respiratory complexes I and IV.

5. The BioID2 technique was successfully established in *T. brucei*.

Chapter II.

Vestiges of the bacterial signal recognition particle-based protein targeting in mitochondria

Vestiges of the Bacterial Signal Recognition Particle-Based Protein Targeting in Mitochondria

Jan Pyrih [ORCID],*[†,1,2] Tomáš Pánek,[3,4] Ignacio Miguel Durante,[1] Vendula Rašková,[1,5] Kristýna Cimrhanzlová,[‡,1,5] Eva Kriegová,[§,1] Anastasios D. Tsaousis,[2] Marek Eliáš,*[,3] and Julius Lukeš*[,1,5]

[1]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice (Budweis), Czech Republic

[2]Laboratory of Molecular and Evolutionary Parasitology, RAPID Group, School of Biosciences, University of Kent, Canterbury, United Kingdom

[3]Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava, Czech Republic

[4]Department of Zoology, Faculty of Science, Charles University, Prague, Czech Republic

[5]Faculty of Sciences, University of South Bohemia, České Budějovice (Budweis), Czech Republic

[†]Present address: Faculty of Science, Charles University, BIOCEV, Vestec, Czech Republic
[‡]Present address: Whitebridge, Lake Macquarie, NSW, Australia
[§]Present address: Institute of Entomology, Biology Centre, Czech Academy of Sciences, České Budějovice, Czech Republic
*Corresponding authors: E-mails: jan.pyrih@gmail.com; marek.elias@osu.cz; jula@paru.cas.cz.
Associate editor: Maria C. Ávila-Arcos

Abstract

The main bacterial pathway for inserting proteins into the plasma membrane relies on the signal recognition particle (SRP), composed of the Ffh protein and an associated RNA component, and the SRP-docking protein FtsY. Eukaryotes use an equivalent system of archaeal origin to deliver proteins into the endoplasmic reticulum, whereas a bacteria-derived SRP and FtsY function in the plastid. Here we report on the presence of homologs of the bacterial Ffh and FtsY proteins in various unrelated plastid-lacking unicellular eukaryotes, namely Heterolobosea, Alveida, *Goniomonas*, and Hemimastigophora. The monophyly of novel eukaryotic Ffh and FtsY groups, predicted mitochondrial localization experimentally confirmed for *Naegleria gruberi*, and a strong alphaproteobacterial affinity of the Ffh group, collectively suggest that they constitute parts of an ancestral mitochondrial signal peptide-based protein-targeting system inherited from the last eukaryotic common ancestor, but lost from the majority of extant eukaryotes. The ability of putative signal peptides, predicted in a subset of mitochondrial-encoded *N. gruberi* proteins, to target a reporter fluorescent protein into the endoplasmic reticulum of *Trypanosoma brucei*, likely through their interaction with the cytosolic SRP, provided further support for this notion. We also illustrate that known mitochondrial ribosome-interacting proteins implicated in membrane protein targeting in opisthokonts (Mba1, Mdm38, and Mrx15) are broadly conserved in eukaryotes and nonredundant with the mitochondrial SRP system. Finally, we identified a novel mitochondrial protein (MAP67) present in diverse eukaryotes and related to the signal peptide-binding domain of Ffh, which may well be a hitherto unrecognized component of the mitochondrial membrane protein-targeting machinery.

Key words: evolution, Ffh, FtsY, LECA, mitochondrion, protein targeting, protists, signal recognition particle.

Introduction

The mitochondrion evolved from an endosymbiont belonging to alphaproteobacteria (Roger et al. 2017; Martijn et al. 2018) and as a cellular component has transitioned into particularly varied forms in different branches of the eukaryotic tree. The key factors underpinning mitochondrial diversity in the extant eukaryotes are lineage-specific innovations and acquisitions, paralleled to a varying degree by losses of ancestral traits. Although mitochondria of conventional model organisms are rather canonical organelles, extremes are found among lesser-known unicellular eukaryotes (Smith and Keeling 2015; Leger et al. 2019; Gray et al. 2020). An example

of an especially pronounced lineage-specific elaboration is provided by the kinetoplastid and diplonemid flagellates with baroquely complex structure and functions of their mitochondrial genomes and transcriptomes (Lukeš et al. 2018; Aphasizheva et al. 2020; Kaur et al. 2020). On the other hand, simplifications have dominated the mitochondrial adaptations of obligate anaerobes, which resulted in organelles without a genome and sometimes even without a function in energy metabolism (Leger et al. 2017; Santos et al. 2018). One such lineage, represented by the oxymonad *Monocercomonoides exilis*, has lost the mitochondrion completely (Karnkowska et al. 2016, 2019).

Somewhat less conspicuous are cases of extraordinary mitochondrial primitiveness, namely the retention of ancestral traits lost by the organelles of most other eukaryotes or at least the commonly studied ones. Some protist groups contain mitochondrial genomes (mitogenomes) that have retained genes relocated to the nuclear genome or completely lost in most other taxa (Kamikawa et al. 2016; Janouškovec et al. 2017). Perhaps the most spectacular example are jakobids with their mitogenomes still encoding subunits of the eubacterial-type RNA polymerase (Burger et al. 2013; Yabuki et al. 2018). Other primitive traits became apparent only with analyses of mitochondrial components encoded by the nuclear genome. The bacterial cytokinetic protein FtsZ present in mitochondria of various protists (Beech et al. 2000; Kiefel et al. 2004), some of which have even kept the regulatory Min system (Leger et al. 2015), is an obvious example. Another case is the recent discovery of a mitochondrial system that involves elements of the bacterial type II secretion system (Horváthová et al. 2021), which was most likely present in the last common eukaryotic ancestor (LECA), yet with the exception for a few little studied protist groups, it was lost in most modern lineages.

Altogether, a picture is emerging that the mitochondrion in the LECA was much more "bacterial" than would be inferred from comparing mitochondria of commonly studied eukaryotes. Here, we present evidence for a hitherto unnoticed bacterial piece of the mitochondrial puzzle that we uncovered while analyzing the mitochondrial proteome of the heterolobosean *Naegleria gruberi*, a free-living amoeboflagellate closely related to the "brain-eating" human pathogen *N. fowleri* (Fritz-Laylin et al. 2010). This piece relates to the mechanism of membrane protein targeting, briefly introduced in the following paragraphs to provide a background for the presentation of our findings.

We are here primarily concerned with mechanisms mediating protein insertion into the bacterial plasma membrane and its evolutionary equivalents, the mitochondrial inner membrane (MIM) and the thylakoid membrane in plastids. In bacteria, most plasma membrane proteins reach their destination via a cotranslational mechanism dependent on two critical components, the signal recognition particle (SRP) and its receptor protein FtsY (Saraogi and Shan 2014; Steinberg et al. 2018). Being composed of the Ffh protein and an RNA component (called 4.5S RNA or 6S RNA, depending on the taxon), SRP recognizes hydrophobic N-terminal signal peptides of nascent proteins as they emerge from the translating ribosome. Peripherally associated with the plasma membrane, FtsY interacts with the SRP, tethering the ribosome-nascent chain complex to the membrane (fig. 1A). This enables docking of this complex to the SecYEG translocation channel, which mediates the integration of the nascent peptide chain into the membrane. An important element of the system is the membrane protein YidC, which functions either in conjunction with the SecYEG channel or as an independent insertase, depending on the substrate (Saraogi and Shan 2014; Steinberg et al. 2018).

It is noteworthy that all eukaryotes share a cytoplasmic signal peptide-driven pathway of cotranslational protein targeting to the endoplasmic reticulum (ER) (fig. 1B). This is clearly an evolutionary derivative of the SRP-based system for plasma membrane protein targeting that operated in the archaeal ancestors of eukaryotes (Zwieb and Bhuiyan 2010; Akopian et al. 2013). The eukaryotic SRP consists of SRP54, a homolog of the archaeal Ffh, and an associated RNA component (7SL RNA). The ER-associated SRP receptor is composed of two subunits, one of which (SRα) evolved from the archaeal FtsY. Furthermore, the plastid-bearing eukaryotes also encode Ffh and FtsY homologs closely related to the eubacterial proteins, which are localized to plastids, organelles derived from an endosymbiotic cyanobacterium (Ponce-Toledo et al. 2017), and function as parts of an SRP machinery mediating cotranslational targeting of membrane proteins encoded by the plastid genome (Ziehe et al. 2017). Many algal groups still possess homologs of all the key components of the system, including cpSRP54 (derived from the cyanobacterial Ffh), SRP RNA (specified by the *ffs* gene still residing in the plastid genome), and cpFtsY, whereas some plants and algae have lost the RNA component (Träger et al. 2012; Ševčíková et al. 2019). The plastid SRP system functions in two modes (fig. 1C): cotranslationally in cooperation with plastid equivalents of SecYEG and YidC (the plastid homolog of the later protein is called Alb3) or posttranslationally, bringing the nucleus-encoded photosynthetic antenna proteins to the Alb3 insertase for their integration into the thylakoid membrane (Ziehe et al. 2017, 2018).

Mitochondria have their own YidC homologs called Oxa1 and Oxa2 (alternatively termed Cox18), which mediate the insertion of both mitochondrial- and nuclear-encoded proteins into the MIM (Oxa1) or are involved in cytochrome *c* oxidase biogenesis (Oxa2) (Hennon et al. 2015). Furthermore, the core subunits of the IMP protease complex in the MIM, which is needed for proteolytic processing of several subunits of the respiratory chain, are related to the bacterial signal peptidase and thus seem to be a rudiment of the original signal peptide-mediated targeting pathway present in the bacterial ancestor of the mitochondrion (Behrens et al. 1991; Gakh et al. 2002). Interestingly, mitogenomes of certain jakobids encode a homolog of the SecY protein (Lang et al. 1997; Burger et al. 2013). If the whole SecYEG complex is present in the mitochondria of these protists as was suggested previously (Tong et al. 2011), it would represent a case of exceptional retention of another ancestral trait related to the SRP-dependent targeting pathway.

However, no mitochondrial equivalent of the SRP system has been reported to date, and a systematic search for its components by bioinformatic analyses of available eukaryotic genomes failed to identify any mitochondrial homologs of Ffh or FtsY (Glick and Von Heijne 1996; Funes et al. 2013). In fact, there is no place for Ffh or FtsY in the paradigmatic view of mitochondrial translation established primarily by studies on yeast and human (fig. 1D), according to which the mitoribosome is stably tethered to the MIM to ensure cotranslational integration of the proteins into the membrane (Ott and Herrmann 2010). In this system, the mitoribosome-membrane association relies on its interaction with a C-terminal extension of Oxa1 and several proteins that seem to be
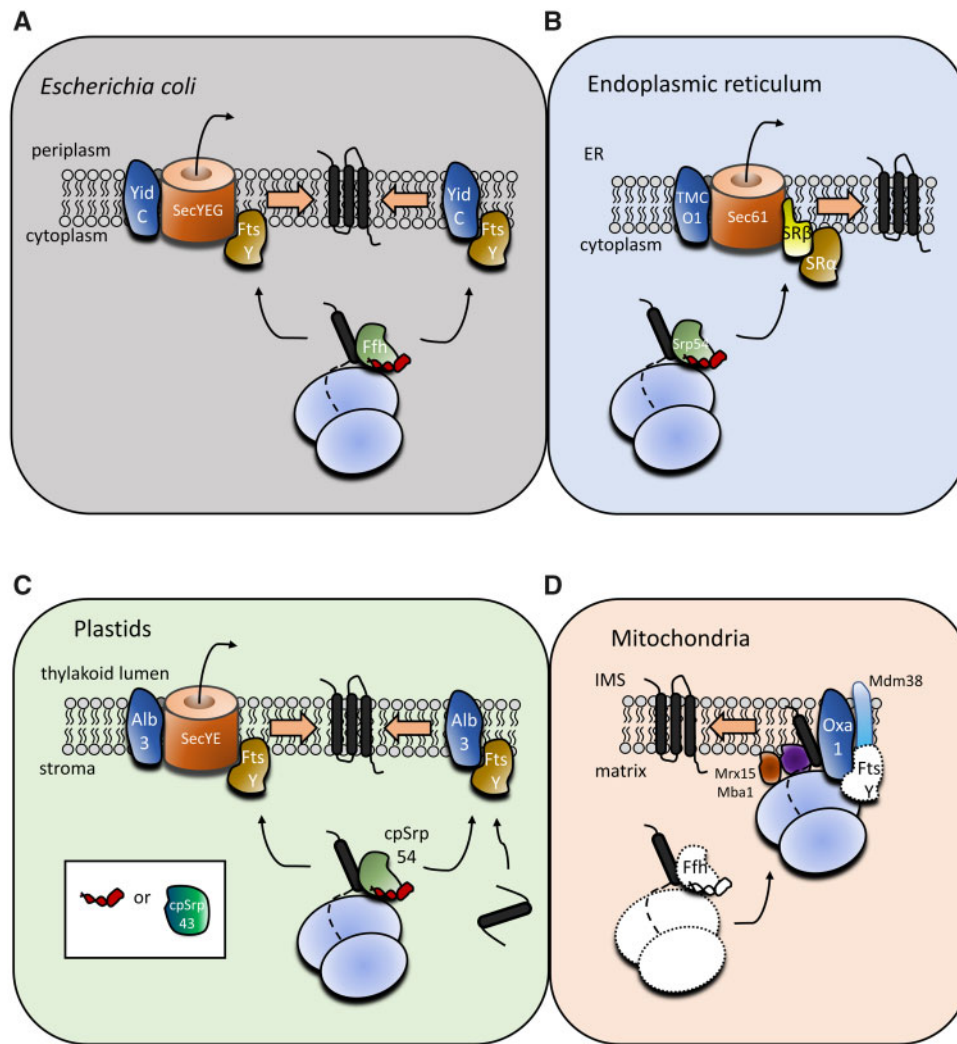
Fig. 1. Simplified cartoon representation of evolutionarily related protein targeting systems in bacteria (exemplified by *Escherichia coli*) (*A*), eukaryotic endoplasmic reticulum (*B*), primary plastids (*C*), and mitochondria (exemplified by *Saccharomyces cerevisiae*) (*D*). Homologous components across the systems are rendered in the same color. The small item in red corresponds to the noncoding 4.5S RNA that together with the protein Ffh or its differently named homologs constitutes the SRP. The inset in part *C* indicates that in some plastids, 4.5S RNA is missing and replaced by the novel protein cpSRP43. The model for the mitochondrion (including the names of the proteins) is based primarily on the situation in yeast mitochondria and it is not certain to what extent it is valid for eukaryotes as a whole. White filling and dotted outlines in part *D* indicate ancestral bacterial features, that is, free (membrane-unbound) translating mitoribosomes and components of the SRP pathway presumably present in the proto-mitochondrion yet presently unknown from mitochondria of extant eukaryotes.

evolutionary innovations of the mitochondrion, including Mba1, Mdm38, and Mrx15 (Ott and Herrmann 2010; Funes et al. 2013; Möller-Hergt et al. 2018). Data available from other eukaryotic models suggest that the mechanism of cotranslational insertion of mitochondrial membrane proteins may be generally similar across distantly related taxa (Christian and Spremulli 2012; Kolli et al. 2018a), although a more detailed comparison is lacking.

Funes et al. (2013) suggested that the SRP system was initially present in mitochondria but became dispensable upon the loss of genes encoding soluble proteins from the mitogenome and was eventually lost due to the emergence of alternative mechanisms for stable association of the mitoribosome with the MIM. They further speculated that some protist lineages with mitogenomes still encoding hydrophilic proteins might represent an intermediate evolutionary stage

with the SRP system possibly retained. Here, we demonstrate that this is indeed the case.

Results

Naegleria gruberi Possesses Mitochondrial Homologs of Ffh and FtsY

While examining a set of putative mitochondrial proteins of the heterolobosean *N. gruberi* defined by Localisation of Organelle Proteins by Isotope Tagging (LOPIT)-based proteomic analysis of cellular fractions (supplementary fig. S1, Supplementary Material online; for details, see Horváthová et al. 2021), we found two proteins, further referred to as *Ng*Ffh and *Ng*FtsY, more similar to the bacterial Ffh and FtsY proteins than to their eukaryotic homologs SRP54 and SRα. Comparison of the existing respective gene models

(Fritz-Laylin et al. 2010) with the genome sequence of *N. fowleri* revealed that both are inaccurate, not only due to incorrectly delimited coding sequence (CDS) but in the case of mtFfh also due to a genome assembly issue (supplementary fig. S2, Supplementary Material online). Amendments to both gene models were confirmed by real time-polymerase chain reaction (PCR) amplification of the 5′ end of the respective transcripts and verified by proteomic data (supplementary fig. S1, Supplementary Material online). The corrected protein sequences (supplementary table S1, Supplementary Material online) were evaluated by multiple protein-targeting prediction tools, which suggested the presence of a mitochondrial presequence in both proteins (supplementary table S2, Supplementary Material online), consistent with their identification in the putative mitochondrial proteome.

Next, the mitochondrial localization of *Ng*Ffh and *Ng*FtsY was tested in the heterologous system of the euglenozoan *Trypanosoma brucei*. Both the N-terminal region (supplementary fig. S1B, Supplementary Material online) and the complete CDSs of *Ng*Ffh and *Ng*FtsY were inserted upstream of the V5-tagged fluorescent mNeonGreen gene and integrated into the rDNA locus of *T. brucei*. Expressed fusion proteins were detected by immunofluorescence with an α-V5 antibody, which in all cases demonstrated colocalization with a mitochondrion-specific marker (α-mtHsp70 antibody) labeling the single reticulated mitochondrion of *T. brucei* (fig. 2). This indicates that the predicted mitochondrial presequences of the *N. gruberi* Ffh and FtsY proteins are recognized by the *T. brucei* mitochondrial protein import machinery, providing further evidence for the presence of homologs of the bacterial Ffh and FtsY proteins in the mitochondrion of *N. gruberi*.

Mitochondrial Ffh and FtsY Have Been Retained in Several Protist Lineages

To gain insights into the evolutionary origin of *Ng*Ffh and *Ng*FtsY, we carried out an exhaustive search for homologous genes in other eukaryotes. After excluding Ffh- and FtsY-related sequences most likely representing bacterial contaminants in the eukaryotic genome and transcriptome assemblies, our phylogenetic analysis revealed a broader set of sequences related to *Ng*Ffh and *Ng*FtsY (fig. 3). These sequences are authentic and not bacterial contaminants, as the corresponding genes (when available in genome assemblies) contain spliceosomal introns or are parts of genomic scaffolds containing unambiguous eukaryotic genes (supplementary table S1, Supplementary Material online). In addition, various prediction algorithms suggested mitochondrial localization for these proteins (supplementary table S2, Supplementary Material online). In case of the Ffh homologs, the putative mitochondrial presequences are apparent as an N-terminal extension missing in bacterial proteins (supplementary fig. S3, Supplementary Material online), whereas the FtsY sequences are insufficiently conserved at the N-terminus to allow such a comparison. As the mitochondrial localization seems to be a common feature of the Ffh and FtsY homologs beyond the experimentally investigated ones in *N. gruberi*, we here denote them mtFfh and mtFtsY, respectively.

Based on the current sampling of the eukaryotic diversity, mtFfh and mtFtsY are restricted to four distantly related eukaryotic lineages, namely Heterolobosea, Hemimastigophora, Alveida, and the genus *Goniomonas* from the supergroup Cryptista. In Heterolobosea, both mtFfh and mtFtsY were found in all species for which sufficiently complete sequence data are available (supplementary table S1, Supplementary Material online), indicating a widespread occurrence of the mitochondrial SRP system in this group (fig. 4). For Hemimastigophora, a recently recognized deep-branching eukaryote lineage (Lax et al. 2018), single-cell transcriptome assemblies yielded both mtFfh and mtFtsY homologs in *Hemimastix kukwesjijk* but only a mtFfh homolog in *Spironema* cf. *multiciliatum*, which most likely reflects an incomplete representation of the gene repertoire in the latter species. Alveida is another recently identified deep-branching lineage containing *Ancoracysta twista* (Janouškovec et al. 2017; Cavalier-Smith et al. 2018) and the isolate Colp-4b (Tikhonenkov DV, personal communication). Although both mtFfh and mtFtsY were found in the transcriptome assembly of Colp-4b, the assembly of *A. twista* contained only the former gene, yet a careful examination of the unassembled raw RNAseq reads allowed us to assemble a partial sequence that falls into the mtFtsY clade (fig. 3C and supplementary table S1, Supplementary Material online). Finally, both mtFfh and mtFtsY sequences were recovered from the genome and/or transcriptome assemblies available for two deeply diverged representatives of the genus *Goniomonas*, *G. avonlea* and *G. pacifica* (supplementary table S1, Supplementary Material online). Interestingly, no mtFfh and mtFtsY candidates were found in other members of Cryptista with genome-scale data available, including diverse algal species of the Cryptophyceae class and the heterotrophic flagellates *Palpitomonas bilix* and *Roombia truncata*.

Phylogenetic analyses resolved mtFfh and mtFtsY sequences as novel clades within the signal recognition–associated GTPase family nested among bacterial sequences but unrelated to the previously known plastid homologs cpSRP54 and cpFtsY (fig. 3). There is strong evidence for the monophyly of mtFfh, with the clade receiving maximal ultrafast bootstrap values in a broad analysis of the whole signal recognition–associated GTPase family (fig. 3A) and an analysis restricted to Ffh/SRP54 sequences (fig. 3B). For the later data set, we also calculated real nonparametric bootstrap values, providing 95% support for the mtFfh clade. The internal topology of the mtFfh clade is generally congruent with the relationships among and within the four major organismal lineages (fig. 3B), which is consistent with vertical inheritance of mtFfh in eukaryotes. Furthermore, the mtFfh clade forms a sister group to sequences from alphaproteobacteria (fig. 3B) or is even nested within them (fig. 3A). In the later case, it branches as a sister group to an Ffh homolog from the uncultivated alphaproteobacterium "MarineAlpha2." Importantly, since the relationship of mtFfh and alphaproteobacterial Ffh is strongly supported in both analyses (fig. 3A and B), mtFfh most likely evolved from Ffh of the alphaproteobacterial ancestor of the mitochondrion.

Presumably due to a more divergent nature of mtFtsY reflected by relatively long branches in the phylogenetic trees
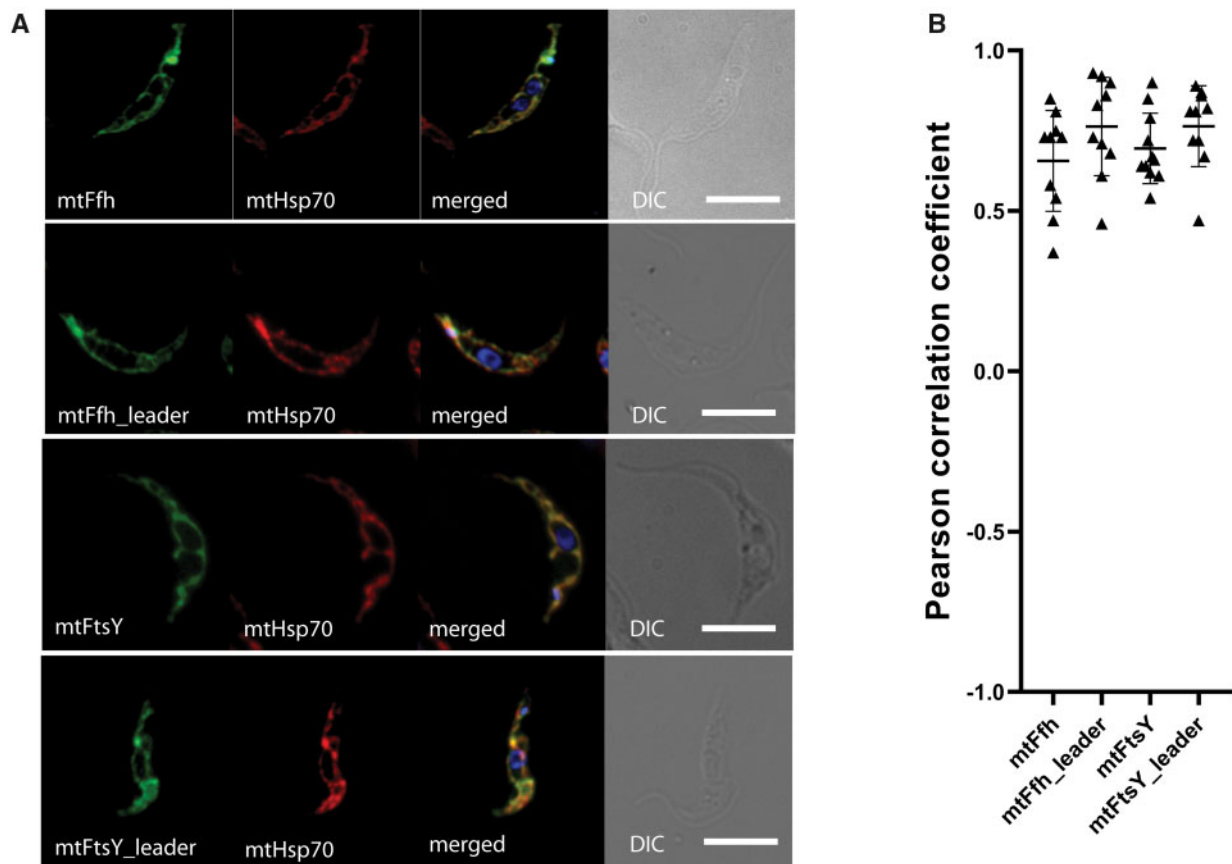
FIG. 2. Mitochondrial localization of heterologously expressed mtFfh and mtFtsY from *Naegleria gruberi*. (*A*) Full-length proteins or their N-terminal leader sequences were expressed in *Trypanosoma brucei* as translation fusions with V5-tagged mNeonGreen protein and visualized by immunofluorescence staining using an α-V5 antibody. Monoclonal α-mtHsp70 antibody served as a mitochondrial marker; DAPI (blue channel) was used to stain DNA; DIC, differential interference contrast. (*B*) PCCs of fluorescent signal colocalization for ≥10 randomly selected cells in each individual cell line. The PCC values range from 1 (i.e., 100% correlation) to −1 (i.e., 100% anticorrelation); values close to 0 mean no correlation. PCC means with standard deviations are displayed for each cell line.

(fig. 3A and C) and a lower number of informative positions (290 vs. 410), its evolutionary history could be reconstructed less robustly than that of mtFfh. Still, the mtFtsY clade is retrieved in both the FtsY/SRα-only (fig. 3C) and FtsY-only analyses (supplementary fig. S4, Supplementary Material online) and supported by 87–89% ultrafast bootstrap replicates. When the most divergent and partial mtFtsY sequence of *Percolomonas cosmopolitus* strain AE-1 was removed from the data set, the ultrafast bootstrap support for the mtFtsY clade increased to 99% in the analysis of the whole signal recognition–associated GTPase family and its internal topology became generally congruent with the known relationship among the species (fig. 3A and supplementary data set S1, Supplementary Material online). On the other hand, the phylogenetic position of the mtFtsY clade among the bacterial FtsY sequences is poorly resolved in all three analyses and the provenance of the closest relatives differs, in neither case being alphaproteobacterial. However, alternative hypotheses that mtFtsY evolved within or sister to alphaproteobacterial FtsY were not rejected by the approximately unbiased (AU) test (supplementary table S3, Supplementary Material online) when applied to the FtsY-only data set. It is worth noting that the plastidial cpSRP54 and cpFtsY are only distantly related to their respective mitochondrial homologs and that the

relation to cyanobacterial equivalents is supported only for the former protein. Similar to mtFtsY, the origin of cpFtsY remains unresolved by our analyses, but the a priori expected cyanobacterial ancestry cannot be rejected by the AU test (supplementary table S3, Supplementary Material online).

Mitochondrial Signal Recognition Particle Lacks the RNA Component

The existence of mtFfh protein in certain eukaryotes raises an obvious question whether a counterpart of the conserved SRP RNA molecule, which together with Ffh constitutes the bacterial SRP, was also retained in the mitochondrion. SRP RNA is poorly conserved in structure and sequence across distantly related taxa, ranging from the 110-nt-long 4.5S RNA of *Escherichia coli* to the 7S RNA of approximately 300 nt found in the archaeal and eukaryotic SRP (Regalia et al. 2002). Hence, we employed a sensitive search strategy using covariance models built based on two different variants of bacterial SRP RNA defined by the Rfam database. As a control, the covariance model representing protistan 7SL RNA (i.e., component of the cytoplasmic SRP) was used. Although the later model identified clear homologs in the nuclear genomes of the mtFfh-carrying species, no candidates for a bacteria-like *ffs* gene were detected.
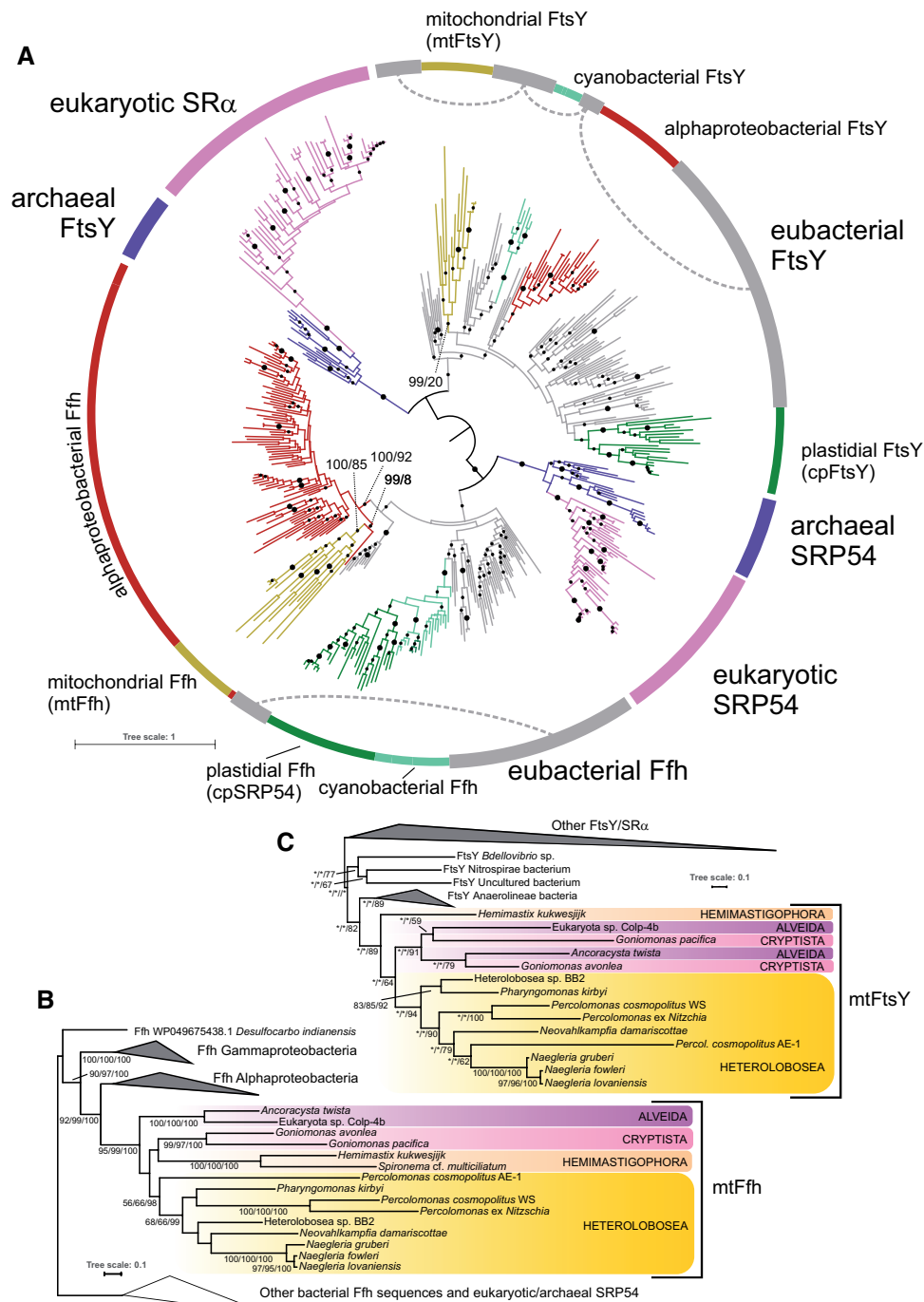
FIG. 3. Phylogenetic analysis of the signal recognition–associated GTPase family gene family showing the position of the mitochondrial Ffh and FtsY homologs. The trees were inferred with the ML method in IQ-TREE (using the LG4X substitution model). Branch support was assessed by nonparametric bootstrapping ($N = 500$, IQ-TREE; only the trees in parts B and C), rapid bootstrapping ($N = 500$, RAxML), and ultrafast bootstrapping ($N = 1,000$, IQ-TREE) using the same substitution model. (A) Phylogenetic position of newly identified mitochondrial Ffh and FtsY proteins within the family (452 sequences, 287 amino acid positions). Branch support $\geq 95\%$ in both methods or only one of them is indicated by a larger or a smaller black dot, respectively. The tree was arbitrarily rooted between the FtsY/SRα and Ffh/SRP groups. Ffh and FtsY homologs from *Paulinella* spp. chromatophore (not specifically highlighted) branch within Cyanobacteria. (B) Detailed phylogenetic analysis of Ffh/SRP54 proteins (295 sequences, 422 amino acid positions). (C) Detailed phylogenetic analysis of FtsY/SRα proteins (217 sequences, 367 amino acid positions). Full trees are provided (in the Newick format) in supplementary data set S1, Supplementary Material online.

The failure to detect the RNA component of the putative mitochondrial SRP may be formally explained by its divergence beyond recognition by bacterial covariance models. We reasoned that analogously to the plastid SRP system, the possible mitochondrial SRP RNA—if present at all—would most likely be produced by transcription of a gene residing in the mitogenome. Furthermore, such a gene would possibly be sufficiently conserved between the closely related species to allow its detection by sequence comparison. We thus systematically compared the predicted intergenic regions of the complete
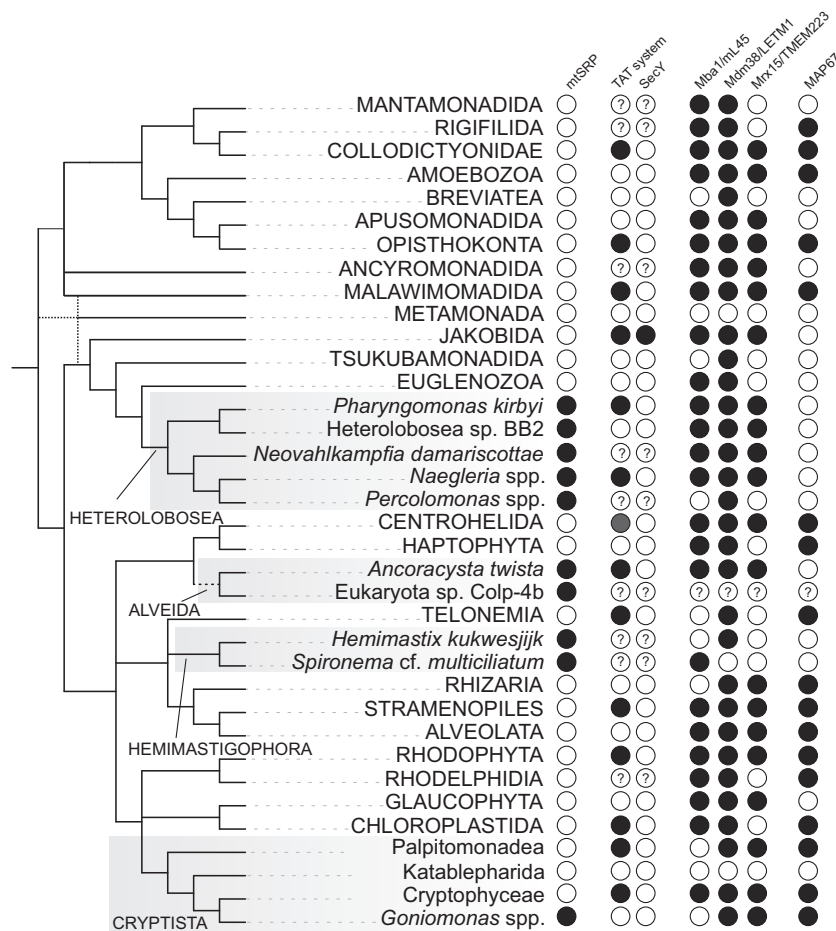
Fig. 4. Phylogenetic distribution of proteins and systems involved in membrane protein targeting and translocation in mitochondria. The schematic phylogeny of eukaryotes is plotted as a consensus of recent phylogenomic analyses of the eukaryote phylogeny (reviewed in Burki et al. 2020) The presence of the mitochondrial Ffh and FtsY (mtSRP), Twin-arginine translocation (TAT) complex, SecY, (putative) mitoribosome membrane receptors (Mba1/mL45, Mdm38/LETM1, and Mrx15/TMEM223), and the novel mitochondrial Ffh-related protein MAP67 is indicated for the main eukaryotic lineages if documented from at least one representative (filled circle). Empty circles indicate that the components are absent from the taxon or has not been identified (which may not necessarily mean true absence, given poor conservation of some of the proteins and limited sampling for many of the main eukaryotic lineages). Question marks indicate that relevant data are missing (mitochondrial genome sequences in case of TAT and SecY) or were not available for analysis (transcriptome assembly from Eukaryota sp. Colp-4b). Data on the occurrence of the TAT complex and SecY were adopted from Petrů et al. (2018) and Tong et al. (2011), respectively, with further updates based on Nishimura et al. (2019) and Wideman et al. (2020).

mitogenomes of *N. gruberi* and *N. fowleri* (the only pair of closely related mtFfh-carrying species with the mitogenome sequences available) to see if any of them exhibits conservation suggestive of a functional constraint. This allowed the identification of an unannotated homolog of Rpl19 noticed previously (Janouškovec et al. 2017) and a short open reading frame of unknown function conserved in multiple heterolobosean species (supplementary fig. S5A, Supplementary Material online), but no candidate RNA gene was found.

These results suggest that no gene for SRP RNA exists in the mitogenomes of eukaryotes harboring mtFfh, with a theoretical exception of Hemimastigophora, for which genome sequences are not yet available. In this regard, it is instructive to consider the situation with the plastidial (chloroplast) SRP (cpSRP). The *ffs* gene is present in plastid genomes of various algae and plants, but many lineages have independently lost it and there is a direct biochemical evidence for the absence of the RNA component in cpSRP of seed plants (Träger et al.

2012). In addition, the absence of *ffs* perfectly correlates with mutations in two specific motifs of cpSRP54 that are critical for its interaction with 4.5S RNA, suggesting that an alternative nuclear gene does not seem to exist in these taxa and the RNA component has indeed been lost (Träger et al. 2012; Ševčíková et al. 2019). Therefore, we checked the corresponding motifs in the mtFfh proteins and found that they are similarly mutated (supplementary fig. S5B, Supplementary Material online). This finding further supports the hypothesis that the mitochondrial SRP is devoid of an RNA component similar to its plastidial counterpart.

N-Termini of Some *N. gruberi* Mitochondrial Proteins Function as Signal Peptides

Given the evolutionary derivation of the mtFfh/mtFtsY system from the bacterial Ffh/FtsY system and considering the precedent of the analogous plastidial cpSRP54/cpFtsY system,

it is reasonable to assume that it mediates targeting of specific protein substrates into the MIM and that this targeting depends on the interaction of mtFfh with the N-terminal signal peptides of the client proteins. Consistently with this hypothesis, 20 out of 46 proteins encoded by the *N. gruberi* mitogenome carry signal peptides predicted by dedicated bioinformatics tools (supplementary table S4, Supplementary Material online). Furthermore, we have identified a strong correlation (*P* value <0.0001) between the presence of transmembrane (TM) domains and a predicted signal peptide in the mitochondrial-encoded proteins (fig. 5A). All proteins where a signal peptide was predicted with the probability exceeding 50% possess two or more TM domains, which is consistent with the assumption that signal peptide targets the protein into the MIM. One protein, the ribosomal protein S4, was predicted to contain a single TM located in the N-terminal region, which is likely a false-positive results due to the function of this protein. The respective region is neither a strong candidate for an SP (fig. 5A).

To test the functionality of these putative signal peptides in vivo, we expressed codon-optimized N-terminal regions of nine mitochondrial-encoded candidates as translational fusions with V5-tagged mNeonGreen from the pT7 vector stably integrated into the nuclear genome of *T. brucei* (for the scheme of the experiment see fig. 5B). Seven proteins were targeted into the ER, which in *T. brucei* forms a reticulated structure with a central perinuclear ring, whereas two—those with the N-terminal parts derived from *orf145* and *tatC* genes, remained in the cytoplasm, possibly in some granules (fig. 5C). Calculation of the hydrophobicity values (free insertion energy; ΔG, kcal/mol; Björkholm et al. 2015) of the putative signal peptides present in the tested N-terminal sequences revealed that the N-terminus of the *tatC*-encoded protein is by far the most hydrophobic. Thus, this protein serves as an additional specificity control excluding the possibility that the protein constructs could be dragged toward the ER simply because of their hydrophobicity. Based on these observations, we conclude that signal peptide-like N-termini of at least seven *N. gruberi* mitochondrial proteins are efficiently recognized by the cytoplasmic (i.e., eukaryotic) SRP-based targeting system in *T. brucei*. As a control, the N-terminus of NADPH-cytochrome p450 reductase from *T. brucei*, which resides in the outer mitochondrial membrane (Niemann et al. 2013) as a predicted signal-anchored protein, targeted the fused V5-tagged mNeonGreen reporter to the mitochondrion (fig. 5D). The colocalization of the tested proteins with specific ER or mitochondrial markers (TbBiP and mtHsp70, respectively) was assessed by calculating Pearson correlation coefficients (PCCs) based on ≥10 cells per each individual cell line, with the results consistent with the visual assessment of the fluorescence signals (fig. 5E). As a further control, PCC was calculated also for TbBiP and mtHsp70 antibody staining in wild-type (SMOX) cells, which indicated no correlation (fig. 5E), consistent with previous studies where those two antibodies were used (Dawoody et al. 2020). This verifies that in our overexpression system the cellular targeting mechanism

distinguishes between SRP-dependent signal peptides and similar, yet SRP-independent N-terminal targeting determinants.

Proteins Implicated in the Mitoribosome–MIM Association in Opisthokonts Are Widespread in Eukaryotes

The most common functional partner of the bacterial SRP system is the SecYEG channel residing in the plasma membrane (fig. 1A), which raises the question as to whether a similar partnership also exists in the mtFfh/mtFtsY-containing mitochondria. An SecY homolog is encoded by the mitogenomes of some jakobid flagellates (Burger et al. 2013), but the recently reported draft genome from a member of this group, *Reclinomonas americana* (Horváthová et al. 2021), indicate the absence of mtFfh and mtFtsY from these organisms. On the other hand, our reinvestigation of the mitogenomes, nuclear genomes, and/or transcriptomes of mtFfh/mtFtsY-carrying taxa did not identify any homologs of the SecYEG complex subunits, indicating a genuine absence of the SecYEG complex. Notably, cotranslational integration of a subset of bacterial proteins into a membrane does not depend on the SecYEG complex and is instead mediated by the insertase YidC as an alternative partner of the SRP system (Steinberg et al. 2018). Mitochondria of yeasts, metazoans, and plants contain two or more YidC homologs, typified by the yeast proteins Oxa1 and Cox18, which are involved in membrane integration or biogenesis of the MIM proteins in an SRP-independent fashion (Hennon et al. 2015; Kolli et al. 2018b). In opisthokonts, Oxa1 interacts directly with the mitoribosome via its C-terminal extension containing a coiled-coil motif (Jia et al. 2003). We examined the genomic and/or transcriptomic data of the mtFfh/mtFtsY-carrying species using a profile HMM specific for the YidC/Oxa1 family. One or two homologs were retrieved for each species (supplementary table S5, Supplementary Material online), and according to phylogenetic analysis, "bona fide" Oxa1 is ubiquitous among the species, with some of them additionally containing a putative Cox18 (Oxa2) ortholog (supplementary fig. S6, Supplementary Material online). Furthermore, the Oxa1 proteins from all mtFfh/mtFtsY-carrying species exhibit C-terminal extensions when compared with the bacterial YidC (supplementary fig. S7, Supplementary Material online), suggesting that the conventional mode of mitoribosome-Oxa1 interaction is preserved in these taxa and therefore was likely already present in the LECA.

The membrane association of a translating mitoribosome depends on additional proteins. Studies performed primarily in the yeast *Saccharomyces cerevisiae* identified three proteins involved in tethering the mitoribosome to the MIM. The best characterized is Mba1 (Ott and Herrmann 2010; Pfeffer et al. 2015), an ortholog of a bona fide mitoribosomal protein mL45 (Mrpl45) (Desmond et al. 2011). Recent cryo-EM studies in the mammalian systems revealed that mL45 participates in a characteristic prberance of the large mitoribosomal subunit and like Mba1 mediates the contact
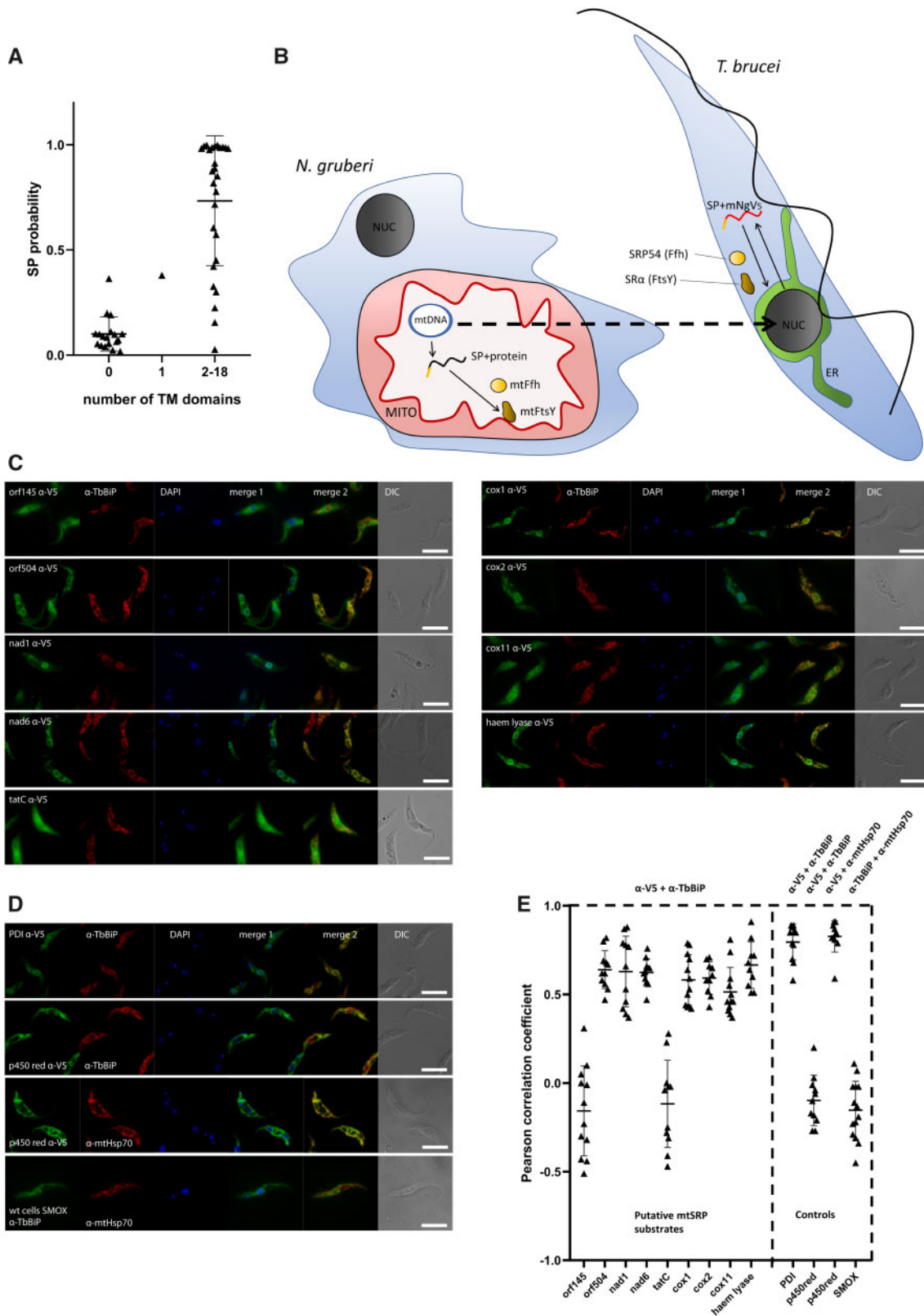
FIG. 5. N-terminal regions of selected *Naegleria gruberi* mitochondrial proteins are recognized as signal peptides by the cytosolic SRP. (*A*) Correlation between the probability of an *N. gruberi* mitochondrion-encoded protein to contain a signal peptide (as predicted with TargetP-1.1) and the number of TM domains (detected via TMHMM server). Mann–Whitney test calculated *P* value below 0.0001. The single protein with a single predicted TM is the ribosomal protein S4, representing a likely false-positive prediction. (*B*) Scheme of experiment. NUC, nucleus; MITO, mitochondrion; ER, endoplasmic reticulum. Small arrows indicate gene expression and subcellular targeting, the dashed arrow indicates nuclear transfection of *Trypanosoma brucei* with DNA constructs encoding fusion proteins comprised putative SPs of *N. gruberi* mitochondrion-encoded proteins and V5-tagged mNeonGreen (mNg). (*C*) Codon-optimized 5′ segments of *N. gruberi* mitochondrial genes (encoding N-terminal regions of

of the mitoribosome with the MIM (Greber et al. 2014; Pfeffer et al. 2015; Englmeier et al. 2017), raising the possibility that this function is more broadly conserved, if not ancestral, in eukaryotes as a whole. Mdm38 is another yeast mitoribosomal membrane receptor, with orthologs in other eukaryotes generally called LETM1 (Hashimi et al. 2013; Austin and Nowikovsky 2019). Mdm38/LETM1 are MIM-localized ion transporters, and whether they function as mitoribosome receptors in eukaryotes other than fungi is not clear. Using a specific profile HMM, we have identified orthologs of both Mba1/mL45 and Mdm38/LETM1 in most major eukaryotic lineages, including those with the mitochondrial SRP system (fig. 4 and supplementary table S6, Supplementary Material online).

Finally, Mrx15 is a newly described yeast mitoribosomal receptor organizing, jointly with Mba1, cotranslational membrane protein insertion (Möller-Hergt et al. 2018). Although proposed to be confined to fungi, our PSI-Blast search (Altschul et al. 1997) with the yeast Mrx15 as a query detected significant similarity to proteins in other eukaryotes including humans, where the homolog is called TMEM223 and besides its mitochondrial localization (Mallmann et al. 2019; Sánchez-Caballero et al. 2020), nothing is known about its function. Further analyses using a profile HMM corroborated the existence of a family of Mrx15-/TMEM223-related proteins, which is widely distributed in eukaryotes including most of the mitochondrial SRP-containing protists (supplementary table S6, Supplementary Material online). The unity of the proposed Mrx15/TMEM223 family is supported by the shared presence of two predicted transmembrane domains (supplementary fig. S8, Supplementary Material online), which were experimentally confirmed for the yeast Mrx15 (Möller-Hergt et al. 2018). Our results indicate that an ancestor of the Mrx15/TMEM223 family was likely already present in the LECA (fig. 4). This has been independently proposed in a recent study (Sánchez-Caballero et al. 2020) based on a much more restricted taxon sampling than employed here.

A Novel Mitochondrial Ffh-Related Protein Occurs in a Broad Range of Eukaryotes

While searching for mtFfh candidates in genome or transcriptome assemblies of diverse eukaryotes, we noticed in some of them weak hits different from the genuine mtFfh or other known proteins. Closer investigation of the corresponding sequences revealed that they constitute a novel protein family related to Ffh/Srp54. These proteins are generally predicted to be targeted to the mitochondrion (supplementary table S7, Supplementary Material online) and the respective representatives were found by mass spectrometry in the

mitochondrion of *Toxoplasma gondii* (TGME49_254230; Seidi et al. 2018), *Arabidopsis thaliana* (AT3G04950; Fuchs et al. 2020), and *Chlamydomonas reinhardtii* (v3 annotation ID 184930; Atteia et al. 2009). In the later species, the protein was listed among mitochondrial proteins of unknown function with the label MAP67, which we adopt here for the whole new protein family. The MAP67 family is broadly distributed in eukaryotes, being present in most major lineages, in some taxa even in more than one version (fig. 4 and supplementary table S7, Supplementary Material online). Notable exceptions are Metazoa, Fungi, Discoba, and Metamonada. Furthermore, we found MAP67 in one of the four mitochondrial SRP-bearing lineages, namely in the genus *Goniomonas*.

Based on sensitive homology searches, MAP67 proteins are along most of their length homologous to the signal peptide-binding M domain of Ffh/SRP54 proteins (fig. 6). Specifically, HHpred found a match to this domain (Pfam PF02978) with a probability of 99.66% and an E-value of 1.4e−15. In addition, the fold prediction server Phyre2 modeled 72% of the length of a reference MAP67 query (from the malawimonad *Gefionella okellyi*) with 100% confidence based on SRP54 from the archaebacterium *Methanocaldococcus jannaschii* as the best template. Interestingly, MAP67 proteins of two different eukaryote groups exhibit short conserved C-terminal extensions. In Chloroplastida, it includes a region matching the SEC-C domain (Pfam PF02810), which is also called the metal-binding domain (MBD) and occurs primarily at the C-terminus of bacterial SecA proteins (Jamshad et al. 2019). The MBD in SecA includes four positions occupied by metal ion-binding cysteine or histidine residues. Its variant in MAP67 proteins from Chloroplastida also includes four cysteine residues, although their positioning is not necessarily the same as in SecA (fig. 6). The second eukaryotic group with C-terminally extended MAP67 is Centrohelida, where the extension consists of a poorly conserved low complexity linker followed by a short highly conserved region of approximately 35 residues homologous to the C-terminus of a subset of bacterial SecA proteins (fig. 6). However, it lacks the characteristic cysteine residues, and HHpred did not detect even a remote similarity to the canonical MBD. Hence, MAP67 independently recruited two different versions of the C-terminus of bacterial SecA proteins in two different eukaryote lineages.

Discussion

Here we show that at least four distantly related eukaryotic lineages (Heterolobosea, Hemimastigophora, Alveida, and *Goniomonas* spp.) harbor homologs of the bacterial Ffh and FtsY proteins that are unrelated to the previously known

the respective proteins) were fused with the reporter V5-tagged mNeonGreen gene and integrated into the *T. brucei* nuclear genome. Except for orf145 and tatC cell lines, the fusion proteins (α-V5 antibody signal) colocalized with the signal of an α-TbBiP antibody, which served as an ER marker. (*D*) Control experiments. Top: the N-terminal region of the *T. brucei* Protein disulfide isomerase (PDI) protein targets V5-tagged mNeonGreen into the ER (positive control). Middle: the N-terminal region of the *T. brucei* NADPH-cytochrome p450 reductase targets V5-tagged mNeonGreen to the mitochondrion (specificity control). Bottom: No colocalization between ER and mitochondrial marker was observed in wild-type (SMOX) cell line. The mitochondrion was specifically labeled by the α-mtHsp70 antibody. DAPI (blue channel) represents DNA; merge 1—an overlay of α-V5 and DAPI signals; merge 2—an overlay of α-V5 and α-TbBiP signals; DIC, differential interference contrast. (*E*) PCCs of fluorescence signal colocalization for ≥10 randomly selected cells in each individual cell line (see the legend to fig. 2*B* for further details).
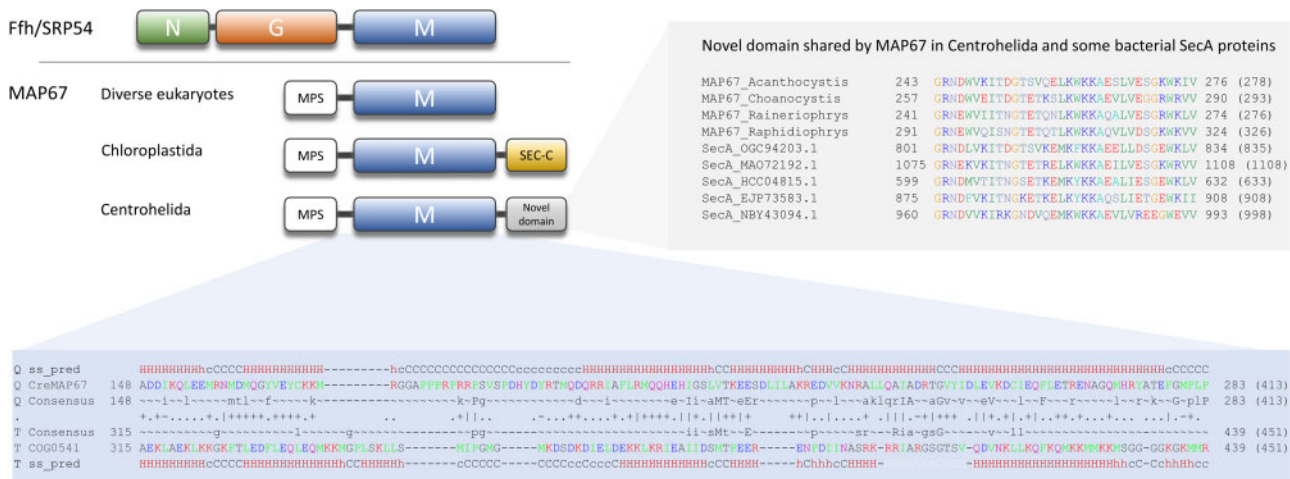
Fig. 6. MAP67, a novel mitochondrial Ffh-related protein. Top left: schematic comparison of the domain architecture of Ffh/SRP54 and different variants of MAP67. MPS—mitochondrial presequence. Bottom: alignment of profile HMMs of MAP67 (with the *Chlamydomonas reinhardtii* MAP67 sequence shown as a reference) and the family COG0541 (Ffh represented in the Clusters of Orthologous Groups database) as retrieved by HHpred. The numbers on the left and right indicate coordinates along the length of the profile HMMs (the numbers in brackets correspond to the total length of the profile HMMs). The alignment is sandwiched by predictions of the secondary structure elements provided by HHpred. Top right: multiple sequence alignment of the novel conserve domain found at the C-terminus of MAP67 proteins from Centrohelida (included are all four sequences available) and a subset of bacterial SecA proteins (five sequences selected as a reference). IDs for the centrohelid sequences are provided in supplementary table S7, Supplementary Material online. The bacterial sequences (defined by GenBank accession numbers) come from the following bacterial taxa: OGC94203.1—Candidatus Adlerbacteria bacterium RIFOXYB1_FULL_48_1; MAO72192.1—Flavobacteriales bacterium; HCC04815.1—Patescibacteria group bacterium; EJP73583.1—SAR86 cluster bacterium SAR86B; NBY43094.1—Verrucomicrobia bacterium.

cyanobacteria-derived cpSRP54 and cpFtsY functioning in the plastids. Two lines of evidence—proteomic data and expression in a heterologous system—conclusively demonstrate that the respective proteins from *N. gruberi* function within the mitochondrion. Considering additional bioinformatic evidence for the mitochondrial localization of their homologs in other eukaryotes, we labeled these proteins as mtFfh and mtFtsY. Phylogenetic analyses indicate their common origin and are consistent with the vertical inheritance of the gene pair from a common ancestor of the respective eukaryotic lineages. The four mtFfh/mtFtsY-harboring groups represent diverse lineages of the proposed "megagroup" Diphoda (Derelle et al. 2015; Lax et al. 2018), which implies that both proteins appeared no later than in the last common ancestor of this clade. However, the alphaproteobacterial origin evident for mtFfh (and not excluded for mtFtsY) suggests an even more ancient origin, specifically from genes of the proto-mitochondrion. This would by inference mean that mtFfh and mtFtsY were possibly present in the LECA and were lost multiple times in a coordinated manner, supporting their functional interdependence.

Meanwhile, it is noteworthy to compare the evolutionary patterns of the SRP system in the plastids and mitochondria. Except for euglenophytes, the plastidial system is ubiquitous (Záhonová et al. 2018), attesting to its tight integration into the molecular fabric of this cyanobacterium-derived organelle. In contrast, the mitochondrial version has been dispensed with on multiple occasions. Moreover, in some taxa, the plastidial system retains its RNA component (Träger et al. 2012), whereas the available evidence suggests that the corresponding SRP RNA had most likely been present in the alphaproteobacterial ancestor of the mitochondrion, yet

was lost prior to the LECA. Another difference rests in the fact that protein targeting mediated by the plastidial SRP system depends on an equivalent of the SecYEG translocation channel (Ziehe et al. 2017), which is missing from eukaryotes bearing the mitochondrial SRP system (at least from those where relevant data are available). Interestingly, the plastidial SRP system has become engaged in posttranslational insertion into the thylakoid membrane of the nucleus-encoded antenna proteins, which (at least in the land plants) depends on a novel protein factor called cpSRP43 interacting with cpSRP (Ziehe et al. 2017, 2018). We wondered whether analogously to cpSRP, mtFfh is accompanied by another novel factor. Following a phylogenetic profiling approach, previously successful in illuminating another patchily distributed mitochondrial system (Horváthová et al. 2021), we looked for proteins with the same or similar phylogenetic profile as mtFfh/mtFtsY but did not find any cooccurring candidates. Nevertheless, the existence of a eukaryote-specific component of the mitochondrial SRP system remains an open possibility that needs to be addressed by more direct approaches.

Since none of the eukaryotes carrying the mitochondrial SRP system is presently amenable to genetic manipulations, it is difficult to address its composition and function by experimental approaches. Assuming functional conservation dating back to bacterial ancestors of the mitochondrion as the most parsimonious alternative, the dissected system is involved in cotranslational membrane protein targeting. Hence, we evaluated the ability of the N-terminal sequences of the mitochondrial-encoded *N. gruberi* proteins that bear characteristics of a signal peptide to navigate a fused reporter fluorescent protein into the ER of genetically tractable *T. brucei*. Previous reports demonstrated that most proteins encoded by

85

the human mitogenome are mistargeted to the ER when expressed from engineered nuclear copies of the respective genes, even when provided with a strong mitochondrial presequence (Björkholm et al. 2015, 2017). This suggested capturing of transmembrane domains in these proteins by the cytosolic SRP analogously to the recognition of signal-anchor sequences in proteins normally targeted to the ER membrane. Our experiments extend these observations by showing that the N-terminal regions of some nonhuman mitochondrial proteins are interpreted by the cytosolic SRP as bona fide signal peptides. Given the fact that the eukaryotic cytosolic SRP is related, however distantly, to the eubacterial SRP, this suggests that the N-termini of these proteins are likewise recognized by mtFfh when they emerge as nascent peptides from a translating mitoribosome. Following the functional paradigm established for both the eubacterial and eukaryotic SRPs, this leads to relocation of the ribosome-nascent chain-mtFfh complex to the MIM mediated by interaction with the membrane-associated mtFtsY receptor. Since the SecYEG complex is absent from eukaryotes known to have the mtFfh/mtFtsY system, the ubiquitous YidC homolog Oxa1 is an obvious candidate for mediating cotranslational membrane insertion of the nascent protein.

Although the cotranslational function of the mitochondrial SRP system is the default hypothesis to test, the ability of the plastidial SRP system to function in a posttranslational mode suggests that such a possibility cannot be dismissed either. Indeed, numerous nucleus-encoded proteins are translocated into the matrix and then inserted into the MIM by Oxa1 working in a posttranslational mode (Stiller et al. 2016; Kolli et al. 2018a). The assistance of the mitochondrial SRP system in such a delivery route would be analogous to the role of the plastidial SRP system in the integration of light-harvesting chlorophyll *a/b*-binding proteins into the thylakoid membrane mediated by the YidC homolog Alb3 (Ziehe et al. 2018). We also considered a possibility that the mitochondrial SRP system interacts with another protein translocase of bacterial origin, the TAT complex located in the MIM of some mitochondria, although this would represent a setting that is unprecedented in bacteria. The mitochondrial TAT components are indeed present in heteroloboseans and the alveid *A. twista* (Petrů et al. 2018). The lack of mitochondrial genome sequences from Hemimastigophora precludes determining if the TAT complex is present in this group. In any case, the recently sequenced mitogenome of *G. avonlea* lacks genes for any TAT subunit (Cenci et al. 2018), making the hypothetical functional link between the mitochondrial SRP system and the TAT translocase unlikely.

Preservation of the mitochondrial SRP system in just a handful of eukaryotic lineages raises the question as to whether they share another feature that predetermined them to keep mtFfh and mtFtsY. Since the morphology and lifestyle of the mtFfh-/mtFtsY-containing eukaryotes vary widely, these provide no clue. It is, however, noticeable that they have gene-rich mitogenomes (yet to be confirmed for Hemimastogophora) with a significant fraction of genes encoding soluble proteins, such as mitoribosomal subunits

(fig. 7A). It is therefore tempting to speculate that in these organisms, stable tethering of mitoribosomes to the MIM, a situation described in yeast, human, and other eukaryotes with few if any soluble mitochondrial-encoded proteins, would not provide sufficient flexibility to the translation apparatus. Although the proportion of membrane-associated and free mitoribosomes in eukaryotes with gene-rich mitogenomes remains unknown, it is plausible that the ratio will be shifted toward the later state. In such a case, the mitochondrial SRP system would provide a means to flexibly regulate the submitochondrial localization of the translating mitoribosome, depending on the nature of the nascent protein. However, this interpretation does not explain the absence of the mitochondrial SRP system from other protists
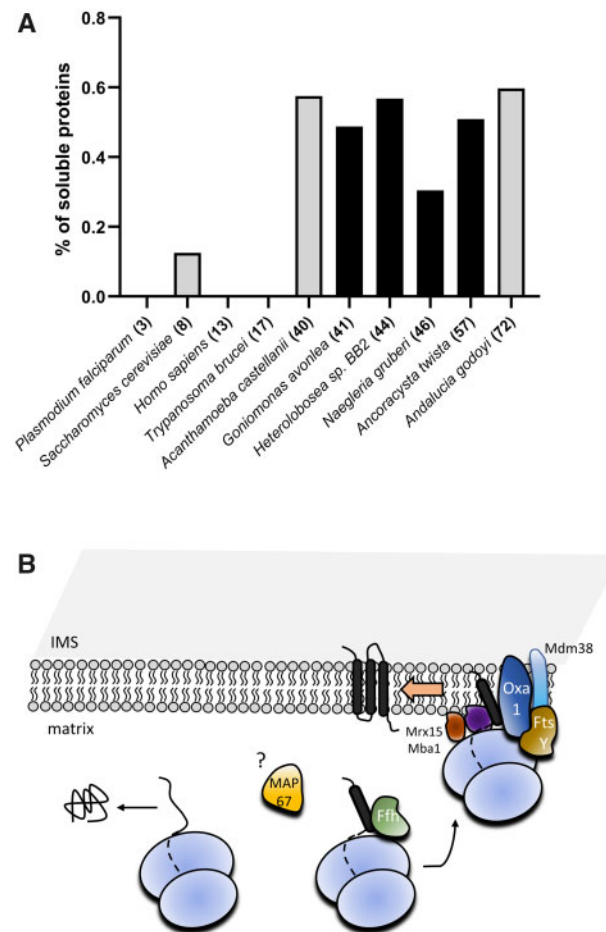
FIG. 7. Persistence of the SRP pathway in mitochondria. (A) The mtSRP pathway is present only in organisms with a high proportion of soluble proteins encoded in the mitogenome. The Y-axis represents the percentage of proteins without any predicted transmembrane domain. Numbers in brackets show the total number of protein-coding genes in the mitogenome. Black color highlights organisms where the mtSRP pathway is present. (B) Cartoon representation of the putative mtSRP pathway as deduced from the results presented in this study (compare with fig. 1D). Names of the proteins displayed in white represent proteins present in all mtSRP system-possessing taxa, whereas proteins in black are present only in some of them.

with gene-rich mitogenomes, including the jakobid *Andalucia godoyi* (Burger et al. 2013), which has retained in its organelle a range of other ancestral bacteria-like traits (Gray et al. 2020).

Interestingly, it has been previously shown that the yeast mitoribosome interacts with a bacterial Ffh when this is expressed in yeast cells and engineered to be targeted to the mitochondrion (Funes et al. 2013), suggesting that the structural prerequisites for the function of the mitochondrial SRP system have been preserved even in lineages that lost it a long time ago. In addition, expression of a mitochondrion-targeted bacterial Ffh in the yeast Δ*mba1 oxa1*Δ*C* strain (i.e., a mutant lacking the *mba1* gene and expressing a truncated version of Oxa1 without the C-terminal ribosome-binding tail) partially rescued the growth defects conferred by the mutations (Funes et al. 2013), which would be compatible with the idea that the ancestral mitochondrial SRP system and the Mba1-driven mitochondrion-specific mechanisms of ribosome membrane association are at least partially functionally redundant. However, the rescue effect of the bacterial Ffh did not depend on the presence of a bacterial FtsY in the yeast mutant (Funes et al. 2013), raising the question about the actual biochemical mechanisms of the Ffh action in the yeast mitochondrion. In this context, it is notable that our comparative genomic and phylogenetic analyses revealed broad conservation of proteins involved in the mitoribosome–MIM association in eukaryotes, including the (putative) mitoribosome receptors Mba1/mL45, Mdm38/LETM1, and Mrx15/TMEM223, as well as the C-terminal extension of Oxa1. Their distribution is consistent with the notion that they could have mediated the mitoribosome–MIM association already in the LECA. Meanwhile, the presence of these proteins in the mtFfh-/mtFtsY-carrying protists suggests that in these eukaryotic lineages, the original SRP-dependent mechanism of protein targeting has coexisted with the newly evolved mechanisms of mitoribosome–MIM association for approximately 1.5 billion years (Betts et al. 2018) and hence is unlikely to be functionally redundant with them.

Furthermore, we identified MAP67, a novel mitochondrial protein that also occurs broadly in eukaryotes and was most likely already present in the LECA. Its obvious evolutionary relationship to Ffh raises the possibility that it is a highly modified ortholog of mtFfh. However, the four mtFfh-bearing lineages are interspersed among taxa with MAP67 and at least one of them, the genus *Goniomonas*, harbors both genes. Therefore, we propose that MAP67 and mtFfh coexisted in early eukaryotes and their current distribution reflects extensive differential loss. Unfortunately, MAP67 is not sufficiently similar to the M domains of Ffh/SRP54 to make a conventional phylogenetic analysis meaningful, but the most parsimonious explanation of its origin is that it emerged from a duplicated copy of mtFfh by an internal deletion that removed its N and G domains. Presently, we can only speculate about the function of MAP67, but it has already been shown to be essential in two model apicomplexans, *T. gondii* (TGGT1_254230; Sidik et al. 2016) and *Plasmodium falciparum* (PF3D7_1004900; https://plasmodb.

org/plasmo/, last accessed March 31, 2021). As it represents a divergent version of the Ffh/SRP54 M domain responsible for binding the signal peptide (Janda et al. 2010), it may still bind the N-terminal regions of mitochondrial proteins with characteristics of a signal peptide and mediate their membrane targeting. Such a role of MAP67 is further supported by the accretion, in Chloroplastida and Centrohelida, of two alternative C-terminal domains of SecA, a bacterial protein unknown from mitochondria that is involved in posttranslational membrane protein targeting (Steinberg et al. 2018). One more piece of evidence for our hypothesis was provided by a recent cryo-EM study of the structure of the mitoribosome from the ciliate *Tetrahymena thermophila* that detected a novel protein, denoted mL105, associated with the mitoribosome tunnel (Tobiasson and Amunts 2020). The authors noticed homology of mL105 to the M domain of Ffh and proposed that it may be involved in protein targeting in the mitochondrion. Unsurprisingly, our inspection of the *T. thermophila* mL105 protein (TTHERM_000931898) identified it as an MAP67 ortholog.

In conclusion, with the identification of mtFfh and mtFtsY, we have unveiled a novel mitochondrial attribute that joins the growing list of components present in the proto-mitochondrial endosymbiont but is retained only by marginal extant eukaryotic groups. We predict that with further exploration of the protist diversity, the reconstructed complexity of the mitochondrial cenancestor and its bacterial character will further increase. Somewhat surprisingly, the mitochondrial SRP system seems to be absent from a group where its presence was suspected based on the previous knowledge, namely the mitochondrial SecY-containing jakobids. The apparently nonoverlapping distribution of the mitochondrial SecY (or possibly a full SecYEG translocon) and the mtSRP system is puzzling and cannot be readily explained without functional characterization of both elements. These uncertainties notwithstanding, we hypothesize that protein targeting in certain extant mitochondria relies on a modified SRP-dependent pathway (fig. 7B) and may represent an impediment for translocation of the corresponding mitochondrial genes into the nuclear genome in the respective eukaryote lineages. Furthermore, our discovery of the broadly occurring MAP67 family that likely evolved from an mtFfh paralog suggests that vestiges of the SRP pathway in mitochondria may not be restricted to the mtFfh-/mtFtsY-carrying taxa. Direct experimental studies of MAP67 in appropriate model systems are necessary to establish its exact role in mitochondrial biology and to understand why MAP67 was lost from many eukaryotes, including metazoans and fungi.

Materials and Methods

Identification of mtFfh and mtFtsY Sequences
In order to identify homologs of mitochondrial SRP pathway in other eukaryotes outside the genus *Naegleria*, we performed a phylogeny-directed search for close homologs of

its two protein components (mtFfh and mtFtsY) in publicly available databases. Using *Naegleria* sequences as a query, we collected 500 best tblastn hits from NCBI Transcriptome Shotgun Assembly (TSA) database, 2000 best tblastn hits from The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP), and approximately 400 sequences from the NCBI nonredundant protein database (100 best BlastP hits from each of the following: unclustered Archaea, Eubacteria, Eukaryota, and from clustered database "nr70"). Transcripts obtained from the MMETSP and NCBI TSA databases were translated into proteins using the TransDecoder utility (Haas et al. 2013). We additionally searched with tblastn transcriptome assemblies from various poorly studied protist lineages that were reported in the literature but are not included in the NCBI database; these were downloaded from the specific public repositories or were obtained upon request from the authors. In several cases, the sequences of special interest that were found truncated were extended by iterative manual blastn searches and recruitment of raw unassembled RNAseq reads available in the SRA database at NCBI. A partial FtsY transcript from *A. twista* was assembled similarly, starting from a seed read identified by an iterative tblastn search against the respective database of RNAseq reads. Some current gene models in genome annotations proved inaccurate and were manually corrected using evidence from transcriptome data and/or comparison with conserved regions in homologs. Ffh and FtsY homologs were also identified in our unpublished genome sequence assembly from the heterolobosean *Neovahlkampfia damariscottae* and manually annotated to define the exon–intron structure of the respective genes. All protein sequences were aligned with MAFFT version 7, using the auto mode (Katoh and Standley 2016) and trimmed manually. A preliminary phylogenetic analysis was performed in RAxML version 8.2.11 (Stamatakis 2014) under the simple PROTCATLG model with 100 rapid bootstraps. Sequences branching in the vicinity mtFtsY and mtFfh were retained for further analyses together with representative sequences from the archaeal, eubacterial, eukaryotic, and plastidial SRP54, Ffh, FtsY, and SRα proteins. Subsequently, several rounds of reciprocal Blastp and phylogenetic analyses were performed to remove contaminants and to add homologs from undersampled lineages. All mtFtsY and mtFfh sequences are listed in supplementary table S1, Supplementary Material online.

Analyses of Protein Sequences

Subcellular targeting of candidate proteins was predicted by using TargetP-1.1 (Emanuelsson et al. 2007; http://www.cbs.dtu.dk/services/TargetP-1.1/index.php, last accessed March 31, 2021), MitoFates (Fukasawa et al. 2015; http://mitf.cbrc.jp/MitoFates/cgi-bin/top.cgi; prediction model: metazoa, last accessed March 31, 2021), MitoProt (Claros and Vincens 1996; https://ihg.gsf.de/ihg/mitoprot.html, last accessed March 31, 2021), and Predotar (Small et al. 2004; https://urgi.versailles.inra.fr/predotar/; animal or fungal sequences, last accessed March 31, 2021). TargetP was also used for the prediction of signal peptides. TMHMM tool (Krogh et al. 2001; http://www.cbs.dtu.dk/services/TMHMM/, last

accessed March 31, 2021) served for the detection of transmembrane domains. Sensitive homology detection tools were employed to search for homologs of proteins of interest that evolve too rapidly to be always detectable across distant relationships by using Blastp (Oxa1, Mba1/mL45, Mrx15/TMEM223, MAP67). The HMMER3 package (Eddy 2011) was used to search a locally maintained protein sequence database (combining data protein sequences downloaded from public resources or inferred from nucleotide sequence data) in parallel to the recently reported EukProt database (Richter et al. 2020). The searches employed as queries profile HMMs built based on seed multiple protein sequence alignments downloaded from the Pfam database (El-GebAli et al. 2019) or custom alignments of previously identified reference sequences prepared with MAFFT. Where appropriate or needed, profile HMMs were iteratively updated by expanding the template alignments with new homologs recognized in the previous search. HMMER searches of the NCBI nr database were carried out using a public server (https://toolkit.tuebingen.mpg.de/tools/hmmer, last accessed March 31, 2021). The identity of the hits was assessed by backward Blastp searches against the NCBI nr database, conserved domain (CD) searches against the NCBI Conserved Domain Database (Yang et al. 2020), and by HHpred searches (Zimmermann et al. 2018; https://toolkit.tuebingen.mpg.de/tools/hhpred, last accessed March 31, 2021). The later searches were initiated either with individual reference query sequences with the default maximal three Multiple sequence alignment (MSA) generation steps utilizing HHblits, or multiple prealigned sequences with no extra MSA generation step. Four databases of profile HMMs—PDB_mmCIF30, COG_KOG, Pfam-A, and NCBI_CDs—were searched at once. Homology of MAP67 (using the sequence from the presumably slowly evolving malawimonad *Gefionella okellyi*) was also investigated by using the fold recognition server Phyre2 (Kelley et al. 2015).

SRP RNA Analysis Using Covariance Models

Alignments of small bacterial SRP RNA (RF00169), large bacterial SRP RNA (RF01854), and protozoan signal recognition particle RNA (RF01856) were downloaded from the Rfam database (Kalvari et al. 2018; http://rfam.xfam.org/clan/CL00003, last accessed March 31, 2021) and processed using tools of the Infernal package version 1.1.2 (Nawrocki and Eddy 2013). Particularly, cmbuild was used to build a covariance model; *E*-value parameters for covariance models were calibrated by cmcalibrate, and cmsearch was used in combination with a particular calibrated model to screen available mtDNAs from Heterolobosea (*Pharyngomonas kirbyi*, Heterolobosea sp. BB2, *Stachyamoeba lipophora*, *Naegleria* spp., *Acrasis kona*, *N. damariscottae*), *A. twista*, and *G. avonlea* as well as nuclear genome assemblies from three *Naegleria* spp., *G. avonlea*, and our unpublished genomic data from *N. damariscottae*.

Phylogenetic Analyses

In an attempt to evaluate the phylogenetic position and robustness of the phylogenetic placement of identified

mitochondrial FtsY and Ffh proteins, we performed a set of phylogenetic analyses using sequences of signal recognition–associated GTPase family identified and collected by phylogeny-directed search (see above). We prepared four taxonomically balanced data sets. Specifically, the broad data set representing the diversity of the whole signal recognition–associated GTPase family (452 taxa) and three more focused data sets: FtsY-only data set containing eubacterial and organellar FtsY sequences (154 operational taxonomic units, or OTUs); FtsY/SRα data set (217 OTUs), and Ffh/SRP54 data set (295 OTUs). Protein sequences were aligned with MAFFT version 7 (Katoh and Standley 2016), using the G-INS-i method with BLOSUM30 scoring matrix and unalignlevel 0.8 (the broad data set) or unalignlevel 0.0 (FtsY-only data set) or the L-INS-i method with BLOSUM30 (Ffh/SRP54 and FtsY/SR data set). Alignments were trimmed manually (FtsY-only data set) or automatically (other data sets) using BMGE version 1.12 (Criscuolo and Gribaldo 2010) with adjusted parameters: BLOSUM30 matrix to estimate entropy-like value for each position; length of selected blocks at least two; maximum gap rate per position 0.6 or 0.8. Divvier 1.0 (Ali et al. 2019) under standard divvying setting was used to remove low confidence homologies from the broad data set before trimming.

Maximum likelihood (ML) phylogenetic analyses were carried out with IQ-TREE multicore version 1.6.10 (Hoang et al. 2018) and RAxML version 8.2.11 (Stamatakis 2014) under the LG4X substitution model suggested by ModelFinder (Kalyaanamoorthy et al. 2017). Branch supports were estimated by using three approaches: ultrafast bootstrapping with an activated "bnni" option to reduce the risk of overestimating branch supports (IQ-TREE, $N = 1{,}000$), rapid bootstrapping ($N = 500$, RAxML), and in case of Ffh/SRP54 also with nonparametric bootstrapping ($N = 400$, IQ-TREE). All bootstrap replicates were mapped on the best IQ-TREE topology using the "sup" option; final trees were visualized with CorelDRAW Home & Student Suite X8. For the ML phylogenetic analysis of the FtsY-only data set (supplementary fig. S4, Supplementary Material online), the AU test (Shimodaira 2002) was performed as implemented in the IQ-TREE multicore version 1.6.10 to evaluate two hypotheses for the phylogenetic origin of the mitochondrial and plastidial FtsY: mtFtsY branching with sequences from alphaproteobacteria and cpFtsY branching with homologs from cyanobacteria, respectively. The AU tests were conducted with hypothetical groupings (loosely constrained) under the LG4X model. The optimized trees were compared with 10,000 resamplings using the RELL method. Each hypothesis was tested in triplicate to show the consistency of the results. Maximum log-likelihoods (logL) of each constraint and replicate, as well as their differences from the unconstrained ML tree (deltaL) are listed in supplementary table S3, Supplementary Material online. The hypotheses within the 95% confidence interval that could not be rejected are those with $P\text{-AU} \geq 0.05$.

Cell Cultivation, Cloning, and Expression

T. brucei procyclic cell line SMOX 927 (Poon et al. 2012) was grown at 27 °C in SDM79 medium (Schönenberger 1979),

whereas *N. gruberi* strain NEG-M (ATCC 30224) was grown axenically at 27 °C in M7 medium (Fulton 1974). Both media were supplemented with 10% fetal bovine serum. The N-terminal region of the mitochondrial-encoded genes from *N. gruberi* were codon-optimized for the expression in *T. brucei* (https://eu.idtdna.com/CodonOpt, last accessed March 31, 2021) and designed as a partially overlapping opposing long primers, which served both as a template and as primers in one cycle PCR. Analogously, N-terminal regions of PDI and NADPH-cytochrome p450 reductase from *T. brucei* were used as a positive and specificity control, respectively. This led to the synthesis of inserts up to 180 bp in length, which were along with the full-length CDS or N-terminal regions corresponding to the predicted mitochondrial signals of the NgFfh and NgFtsY individually subcloned in pT7 plasmid (Shaner et al. 2013) modified by insertion of the mNg gene in front of the V5 tag. The plasmid was linearized with *Not*I restriction enzyme and nucleofected into the *T. brucei* procyclic stage as described earlier (Kaurov et al. 2018). Expression of the proteins was induced with doxycycline overnight or for just a few hours, as was the case of full-length NgFfh and NgFtsY.

Immunofluorescence Microscopy

Trypanosoma brucei procyclic cells were harvested, washed twice (900 g, 5 min at room temperature [RT]) with Voorheis's-modified phosphate-buffered saline (vPBS; PBS supplemented with 10 mM glucose and 46 mM sucrose, pH 7.6) and the cell suspension was transferred on a microscopic slide covered with poly-L-lysine. Attached cells were fixed for 15 min with 4% paraformaldehyde at RT. Afterward, the cells were permeabilized with 0.1% Triton X-100 in PBS for 15 min. Blocking was performed for 1 h in 1% bovine serum albumine (BSA) in PBS supplemented with 0.033% Triton-X-100 and the same buffer (but without BSA) was also used for all washing steps. The expressed proteins were visualized using rabbit α-V5 antibody (Sigma–Aldrich), with α-mtHsp70 and α-TbBiP antibodies (Bangs et al. 1993; Panigrahi et al. 2008) used as mitochondrial and ER markers, respectively. Goat α-rabbit Alexa Fluor 488 and goat α-mouse Alexa Fluor 555 (both Life Technologies) were used as secondary antibodies. DNA was stained with ProLong1 Gold antifade reagent with 4′,6-diamidine-2′-phenylindole dihydrochloride (DAPI) (Molecular Probes), and stained cells were observed with Zeiss microscope Axioplan 2 equipped with an Olympus DP73 digital camera. Images were processed using the Fiji software and Pearson correlation coefficient for signals from different channels was calculated using the Coloc 2 plugin with default settings (Schindelin et al. 2012).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Data Availability

The data underlying this article are available in the article and in its Supplementary Material online.

References

Akopian D, Shen K, Zhang X, Shan S. 2013. Signal recognition particle: an essential protein-targeting machine. *Annu Rev Biochem*. 82:693–721.

Ali RH, Bogusz M, Whelan S, Tamura K. 2019. Identifying clusters of high confidence homologies in multiple sequence alignments. *Mol Biol Evol*. 36(10):2340–2351.

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(17):3389–3402.

Aphasizheva I, Alfonzo J, Carnes J, Cestari I, Cruz-Reyes J, Göringer HU, Hajduk S, Lukeš J, Madison-Antenucci S, Maslov DA, et al. 2020. Lexis and grammar of mitochondrial RNA processing in trypanosomes. *Trends Parasitol*. 36(4):337–355.

Atteia A, Adrait A, Brugière S, Tardif M, van Lis R, Deusch O, Dagan T, Kuhn L, Gontero B, Martin W, et al. 2009. A proteomic survey of *Chlamydomonas reinhardtii* mitochondria sheds new light on the metabolic plasticity of the organelle and on the nature of the α-proteobacterial mitochondrial ancestor. *Mol Biol Evol*. 26(7):1533–1548.

Austin S, Nowikovsky K. 2019. LETM1: essential for mitochondrial biology and cation homeostasis? *Trends Biochem Sci*. 44(8):648–658.

Bangs JD, Uyetake L, Brickman MJ, Balber AE, Boothroyd JC. 1993. Molecular cloning and cellular localization of a BiP homologue in *Trypanosoma brucei*. Divergent ER retention signals in a lower eukaryote. *J Cell Sci*. 105:1101–1113.

Beech PL, Nheu T, Schultz T, Herbert S, Lithgow T, Gilson PR, McFadden GI. 2000. Mitochondrial FtsZ in a chromophyte alga. *Science* 287(5456):1276–1279.

Behrens M, Michaelis G, Pratje E. 1991. Mitochondrial inner membrane protease 1 of *Saccharomyces cerevisiae* shows sequence similarity to the *Escherichia coli* leader peptidase. *Mol Gen Genet*. 228(1–2):167–176.

Betts HC, Puttick MN, Clark JW, Williams TA, Donoghue PCJ, Pisani D. 2018. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat Ecol Evol*. 2(10):1556–1562.

Björkholm P, Harish A, Hagström E, Ernst AM, Andersson SG. 2015. Mitochondrial genomes are retained by selective constraints on protein targeting. *Proc Natl Acad Sci U S A*. 112(33):10154–10161.

Björkholm P, Ernst AM, Hagström E, Andersson SG. 2017. Why mitochondria need a genome revisited. *FEBS Lett*. 591(1):65–75.

Burger G, Gray MW, Forget L, Lang BF. 2013. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. *Genome Biol Evol*. 5(2):418–438.

Burki F, Roger AJ, Brown MW, Simpson AGB. 2020. The new tree of eukaryotes. *Trends Ecol Evol*. 35(1):43–55.

Cavalier-Smith T, Chao EE, Lewis R. 2018. Multigene phylogeny and cell evolution of chromist infrakingdom Rhizaria: contrasting cell organisation of sister phyla Cercozoa and Retaria. *Protoplasma* 255(5):1517–1574.

Cenci U, Sibbald SJ, Curtis BA, Kamikawa R, Eme L, Moog D, Henrissat B, Maréchal E, Chabi M, Djemiel C, et al. 2018. Nuclear genome sequence of the plastid-lacking cryptomonad *Goniomonas avonlea* provides insights into the evolution of secondary plastids. *BMC Biol*. 16(1):137.

Christian BE, Spremulli LL. 2012. Mechanism of protein biosynthesis in mammalian mitochondria. *Biochim Biophys Acta*. 1819(9–10):1035–1054.

Claros MG, Vincens P. 1996. Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem*. 241(3):779–786.

Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol*. 10:210.

Dawoody NL, Stumpe M, Rauch M, Hemphill A, Schneiter R, Bütikofer P, Serricchio M. 2020. Mitochondrial sphingosine-1-phosphate lyase is essential for phosphatidylethanolamine synthesis and survival of *Trypanosoma brucei*. *Sci Rep*. 10(1):8268.

Derelle R, Torruella G, Klimeš V, Brinkmann H, Kim E, Vlček Č, Eliáš M. 2015. Bacterial proteins pinpoint a single eukaryotic root. *Proc Natl Acad Sci U S A*. 112(7):E693–E699.

Desmond E, Brochier-Armanet C, Forterre P, Gribaldo S. 2011. On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes. *Res Microbiol*. 162(1):53–70.

Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol*. 7(10):e1002195.

El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, et al. 2019. The Pfam protein families database in 2019. *Nucleic Acids Res*. 47(D1):D427–D432.

Emanuelsson O, Brunak S, von Heijne G, Nielsen H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*. 2(4):953–971.

Englmeier R, Pfeffer S, Förster F. 2017. Structure of the human mitochondrial ribosome studied in situ by cryoelectron tomography. *Structure* 25(10):1574–1581.

Fritz-Laylin LK, Prochnik SE, Ginger ML, Dacks JB, Carpenter ML, Field MC, Kuo A, Paredez A, Chapman J, Pham J, et al. 2010. The genome of *Naegleria gruberi* illuminates early eukaryotic versatility. *Cell* 140(5):631–642.

Fuchs P, Rugen N, Carrie C, Elsässer M, Finkemeier I, Giese J, Hildebrandt TM, Kühn K, Maurino VG, Ruberti C, et al. 2020. Single organelle function and organization as estimated from Arabidopsis mitochondrial proteomics. *Plant J*. 101(2):420–441.

Fukasawa Y, Tsuji J, Fu S-C, Tomii K, Horton P, Imai K. 2015. MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. *Mol Cell Proteomics*. 14(4):1113–1126.

Fulton C. 1974. Axenic cultivation of *Naegleria gruberi*. Requirement for methionine. *Exp Cell Res*. 88(2):365–370.

Funes S, Westerburg H, Jaimes-Miranda F, Woellhaf MW, Aguilar-Lopez JL, Janßen L, Bonnefoy N, Kauff F, Herrmann JM. 2013. Partial suppression of Oxa1 mutants by mitochondria-targeted signal recognition particle provides insights into the evolution of the cotranslational insertion systems. *FEBS J*. 280(3):904–915.

Gakh O, Cavadini P, Isaya G. 2002. Mitochondrial processing peptidases. *Biochim Biophys Acta: Mol Cell Res*. 1592(1):63–77.

Glick BS, Von Heijne G. 1996. *Saccharomyces cerevisiae* mitochondria lack a bacterial-type sec machinery. *Protein Sci*. 5(12):2651–2652.

Gray MW, Burger G, Derelle R, Klimeš V, Leger MM, Sarrasin M, Vlček Č, Roger AJ, Eliáš M, Lang BF. 2020. The draft nuclear genome sequence and predicted mitochondrial proteome of *Andalucia godoyi*, a protist with the most gene-rich and bacteria-like mitochondrial genome. *BMC Biol*. 18(1):22.

Greber BJ, Boehringer D, Leitner A, Bieri P, Voigts-Hoffmann F, Erzberger JP, Leibundgut M, Aebersold R, Ban N. 2014. Architecture of the large subunit of the mammalian mitochondrial ribosome. *Nature* 505(7484):515–519.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc.* 8(8):1494–1512.

Hashimi H, McDonald L, Stříbrná E, Lukeš J. 2013. Trypanosome Letm1 protein is essential for mitochondrial potassium homeostasis. *J Biol Chem.* 288(37):26914–26925.

Hennon SW, Soman R, Zhu L, Dalbey RE. 2015. YidC/Alb3/Oxa1 family of insertases. *J Biol Chem.* 290(24):14866–14874.

Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol.* 35(2):518–522.

Horváthová L, Žárský V, Pánek T, Derelle R, Pyrih J, Motyčková A, Klápšťová V, Klimeš V, Petrů M, Vaitová Z., et al. Forthcoming 2021. Analysis of diverse eukaryotes suggests the existence of an ancestral mitochondrial apparatus derived from the bacterial type II secretion system. *Nat Commun.* bioRixv. 10.1101/790865.

Jamshad M, Knowles TJ, White SA, Ward DG, Mohammed F, Rahman KF, Wynne M, Hughes GW, Kramer G, Bukau B, et al. 2019. The C-terminal tail of the bacterial translocation ATPase SecA modulates its activity. *eLife* 8:e48385.

Janda CY, Li J, Oubridge C, Hernández H, Robinson CV, Nagai K. 2010. Recognition of a signal peptide by the signal recognition particle. *Nature* 465(7297):507–510.

Janouškovec J, Tikhonenkov DV, Burki F, Howe AT, Rohwer FL, Mylnikov AP, Keeling PJ. 2017. A new lineage of eukaryotes illuminates early mitochondrial genome reduction. *Curr Biol.* 27(23):3717–3724.

Jia L, Dienhart M, Schramp M, McCauley M, Hell K, Stuart RA. 2003. Yeast Oxa1 interacts with mitochondrial ribosomes: the importance of the C-terminal region of Oxa1. *EMBO J.* 22(24):6438–6447.

Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, Bateman A, Finn RD, Petrov AI. 2018. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* 46(D1):D335–D342.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 14(6):587–589.

Kamikawa R, Shiratori T, Ishida KI, Miyashita H, Roger AJ. 2016. Group II intron-mediated trans-splicing in the gene-rich mitochondrial genome of an enigmatic eukaryote, *Diphylleia rotans. Genome Biol Evol.* 8(2):458–466.

Karnkowska A, Treitli SC, Brzoň O, Novák L, Vacek V, Soukal P, Barlow LD, Herman EK, Pipaliya SV, Pánek T, et al. 2019. The oxymonad genome displays canonical eukaryotic complexity in the absence of a mitochondrion. *Mol Biol Evol.* 36(10):2292–2312.

Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petrželková R, Eme L, Novák L, Žárský V, Barlow LD, Herman EK, et al. 2016. A eukaryote without a mitochondrial organelle. *Curr Biol.* 26(10):1274–1284.

Katoh K, Standley DM. 2016. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* 32(13):1933–1942.

Kaur B, Záhonová K, Valach M, Faktorová D, Prokopchuk G, Burger G, Lukeš J. 2020. Gene fragmentation and RNA editing without borders: eccentric mitochondrial genomes of diplonemids. *Nucleic Acids Res.* 48(5):2694–2708.

Kaurov I, Vancová M, Schimanski B, Cadena LR, Heller J, Bílý T, Potěšil D, Eichenberger C, Bruce H, Oeljeklaus S, et al. 2018. The diverged Trypanosome MICOS complex as a hub for mitochondrial cristae shaping and protein import. *Curr Biol.* 28(21):3393–3407.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 10(6):845–858.

Kiefel BR, Gilson PR, Beech PL. 2004. Diverse eukaryotes have retained mitochondrial homologues of the bacterial division protein FtsZ. *Protist* 155(1):105–115.

Kolli R, Soll J, Carrie C. 2018a. Plant mitochondrial inner membrane protein insertion. *Int J Mol Sci.* 19:1188–1196.

Kolli R, Soll J, Carrie C. 2018b. OXA2b is crucial for proper membrane insertion of COX2 during biogenesis of complex IV in plant mitochondria. *Plant Physiol.* 179(2):601–615.

Krogh A, Larsson B, Von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 305(3):567–580.

Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387(6632):493–497.

Lax G, Eglit Y, Eme L, Bertrand EM, Roger AJ, Simpson AGB. 2018. Hemimastigophora is a novel supra-kingdom-level lineage of eukaryotes. *Nature* 564(7736):410–414.

Leger MM, Kolisko M, Kamikawa R, Stairs CW, Kume K, Čepička I, Silberman JD, Andersson JO, Xu F, Yabuki A, et al. 2017. Organelles that illuminate the origins of *Trichomonas* hydrogenosomes and *Giardia* mitosomes. *Nat Ecol Evol.* 1(4):0092.

Leger MM, Kolísko M, Stairs CW, Simpson AGB. 2019. Mitochondrion-related organelles in feee-living protists. In: Tachezy J, editor. Hydrogenosomes and mitosomes: mitochondria of anaerobic eukaryotes. Cham: Springer Nature Switzerland AG. p. 287–308.

Leger MM, Petrů M, Žárský V, Eme L, Vlček Č, Harding T, Lang BF, Eliáš M, Doležal P, Roger AJ. 2015. An ancestral bacterial division system is widespread in eukaryotic mitochondria. *Proc Natl Acad Sci U S A.* 112(33):10239–10246.

Lukeš J, Wheeler R, Jirsová D, David V, Archibald JM. 2018. Massive mitochondrial DNA content in diplonemid and kinetoplastid protists. *IUBMB Life.* 70(12):1267–1274.

Mallmann R, Ondacova K, Moravcikova L, Jurkovicova-Tarabova B, Pavlovicova M, Moravcik R, Lichvarova L, Kominkova V, Klugbauer N, Lacinova L. 2019. Four novel interaction partners demonstrate diverse modulatory effects on voltage-gated CaV2.2 Ca2+ channels. *Pflugers Arch .* 471(6):861–874.

Martijn J, Vosseberg J, Guy L, Offre P, Ettema TJG. 2018. Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature* 557(7703):101–105.

Möller-Hergt BV, Carlström A, Stephan K, Imhof A, Ott M. 2018. The ribosome receptors Mrx15 and Mba1 jointly organize cotranslational insertion and protein biogenesis in mitochondria. *Mol Biol Cell.* 29(20):2386–2396.

Nawrocki EP, Eddy SR. 2013. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29(22):2933–2935.

Niemann M, Wiese S, Mani J, Chanfon A, Jackson C, Meisinger C, Warscheid B, Schneider A. 2013. Mitochondrial outer membrane proteome of *Trypanosoma brucei* reveals novel factors required to maintain mitochondrial morphology. *Mol Cell Proteomics.* 12(2):515–528.

Nishimura Y, Shiratori T, Ishida KI, Hashimoto T, Ohkuma M, Inagaki Y. 2019. Horizontally-acquired genetic elements in the mitochondrial genome of a centrohelid *Marophrys* sp. SRT127. *Sci Rep.* 9(1):4850.

Ott M, Herrmann JM. 2010. Co-translational membrane insertion of mitochondrially encoded proteins. *Biochim Biophys Acta.* 1803(6):767–775.

Panigrahi AK, Zíková A, Dalley RA, Acestor N, Ogata Y, Anupama A, Myler PJ, Stuart KD. 2008. Mitochondrial complexes in *Trypanosoma brucei. Mol Cell Proteomics* 7(3):534–545.

Petrů M, Wideman J, Moore K, Alcock F, Palmer T, Doležal P. 2018. Evolution of mitochondrial TAT translocases illustrates the loss of bacterial protein transport machines in mitochondria. *BMC Biol.* 16(1):141.

Pfeffer S, Woellhaf MW, Herrmann JM, Förster F. 2015. Organization of the mitochondrial translation machinery studied in situ by cryoelectron tomography. *Nat Commun.* 6:6019.

Ponce-Toledo RI, Deschamps P, López-García P, Zivanovic Y, Benzerara K, Moreira D. 2017. An early-branching freshwater cyanobacterium at the origin of plastids. *Curr Biol.* 27(3):386–391.

Poon SK, Peacock L, Gibson W, Gull K, Kelly S. 2012. A modular and optimized single marker system for generating *Trypanosoma brucei*

cell lines expressing T7 RNA polymerase and the tetracycline repressor. *Open Biol.* 2(2):110037.

Regalia M, Rosenblad MA, Samuelsson T. 2002. Prediction of signal recognition particle RNA genes. *Nucleic Acids Res.* 30(15):3368–3377.

Richter DJ, Berney C, Strassert JFH, Burki F, de Vargas C. 2020. EukProt: a database of genome-scale predicted proteins across the diversity of eukaryotic life. bioRxiv. doi:10/1101/2020.06.30.180687

Roger AJ, Muñoz-Gómez SA, Kamikawa R. 2017. The origin and diversification of mitochondria. *Curr Biol.* 27(21):R1177–R1192.

Sánchez-Caballero L, Elurbe DM, Baertling F, Guerrero-Castillo S, van den Brand M, van Strien J, van Dam TJP, Rodenburg R, Brandt U, Huynen MA, et al. 2020. TMEM70 functions in the assembly of complexes I and V. *Biochim Biophys Acta Bioenerg.* 1861(8):148202.

Santos HJ, Makiuchi T, Nozaki T. 2018. Reinventing an organelle: the reduced mitochondrion in parasitic protists. *Trends Parasitol.* 34(12):1038–1055.

Saraogi I, Shan SO. 2014. Co-translational protein targeting to the bacterial membrane. *Biochim Biophys Acta.* 1843(8):1433–1441.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods.* 9(7):676–682.

Schönenberger BR. 1979. Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. *Acta Trop.* 36:289–292.

Seidi A, Muellner-Wong LS, Rajendran E, Tjhin ET, Dagley LF, Aw VYT, Faou P, Webb AI, Tonkin CJ, van Dooren GG. 2018. Elucidating the mitochondrial proteome of *Toxoplasma gondii* reveals the presence of a divergent cytochrome c oxidase. *Elife* 7:e38131.

Ševčíková T, Yurchenko T, Fawley KP, Amaral R, Strnad H, Santos LMA, Fawley MW, Eliáš M. 2019. Plastid genomes and proteins illuminate the evolution of eustigmatophyte algae and their bacterial endosymbionts. *Genome Biol Evol.* 11(2):362–379.

Shaner NC, Lambert GG, Chammas A, Ni Y, Cranfill PJ, Baird MA, Sell BR, Allen JR, Day RN, Israelsson M, et al. 2013. A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum. Nat Methods.* 10(5):407–409.

Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst Biol.* 51(3):492–508.

Sidik SM, Huet D, Ganesan SM, Huynh MH, Wang T, Nasamu AS, Thiru P, Saeij JPJ, Carruthers VB, Niles JC, et al. 2016. A genome-wide CRISPR screen in *Toxoplasma* identifies essential apicomplexan genes. *Cell* 166(6):1423–1435.

Small I, Peeters N, Legeai F, Lurin C. 2004. Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4(6):1581–1590.

Smith DR, Keeling PJ. 2015. Mitochondrial and plastid genome architecture: reoccurring themes, but significant differences at the extremes. *Proc Natl Acad Sci U S A.* 112(33):10177–10184.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.

Steinberg R, Knüpffer L, Origi A, Asti R, Koch HG. 2018. Co-translational protein targeting in bacteria. *FEMS Microbiol Lett.* 365:fny095.

Stiller SB, Höpker J, Oeljeklaus S, Schütze C, Schrempp SG, Vent-Schmidt J, Horvath SE, Frazier AE, Gebert N, Van Der Laan M, et al. 2016. Mitochondrial OXA translocase plays a major role in biogenesis of inner-membrane proteins. *Cell Metab.* 23(5):901–908.

Tobiasson V, Amunts A. 2020. Ciliate mitoribosome illuminates evolutionary steps of mitochondrial translation. *eLife* 9:e59264.

Tong J, Dolezal P, Selkrig J, Crawford S, Simpson AGB, Noinaj N, Buchanan SK, Gabriel K, Lithgow T. 2011. Ancestral and derived protein import pathways in the mitochondrion of *Reclinomonas americana. Mol Biol Evol.* 28(5):1581–1591.

Träger C, Rosenblad MA, Ziehe D, Garcia-Petit C, Schrader L, Kock K, Vera Richter C, Klinkert B, Narberhaus F, Herrmann C, et al. 2012. Evolution from the prokaryotic to the higher plant chloroplast signal recognition particle: the signal recognition particle RNA is conserved in plastids of a wide range of photosynthetic organisms. *Plant Cell* 24(12):4819–4836.

Wideman JG, Monier A, Rodríguez-Martínez R, Leonard G, Cook E, Poirier C, Maguire F, Milner DS, Irwin N, Moore K, et al. 2020. Unexpected mitochondrial genome diversity revealed by targeted single-cell genomics of heterotrophic flagellated protists. *Nat Microbiol.* 5(1):154–165.

Yabuki A, Gyaltshen Y, Heiss AA, Fujikura K, Kim E. 2018. *Ophirina amphinema* n. gen., n. sp., a new deeply branching discobid with phylogenetic affinity to jakobids. *Sci Rep.* 8(1):16219.

Yang M, Derbyshire MK, Yamashita RA, Marchler-Bauer A. 2020. NCBI's conserved domain database and tools for protein domain analysis. *Curr Prot Bioinformatics.* 69:e90.

Záhonová K, Füssy Z, Birčák E, Novák Vanclová AMG, Klimeš V, Vesteg M, Krajčovič J, Oborník M, Eliáš M. 2018. Peculiar features of the plastids of the colourless alga *Euglena longa* and photosynthetic euglenophytes unveiled by transcriptome analyses. *Sci Rep.* 8(1):17012.

Ziehe D, Dünschede B, Schünemann D. 2017. From bacteria to chloroplasts: evolution of the chloroplast SRP system. *Biol Chem.* 398(5–6):653–661.

Ziehe D, Dünschede B, Schünemann D. 2018. Molecular mechanism of SRP-dependent light-harvesting protein transport to the thylakoid membrane in plants. *Photosynth Res.* 138(3):303–313.

Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol.* 430(15):2237–2243.

Zwieb C, Bhuiyan S. 2010. Archaea signal recognition particle shows the way. *Archaea* 2010:485051.

(A)

Mba1
mtFfh
Mdm38
Oxa1
Mrx15

All proteins identified
Mito markers
Mito proteins
Perox markers
Perox proteins
Proteins of interest
ER markers

SRa
Srp54
mtFtsY

ER markers:
Calreticulin, PDI,
Calnexin, Sec61

(B)
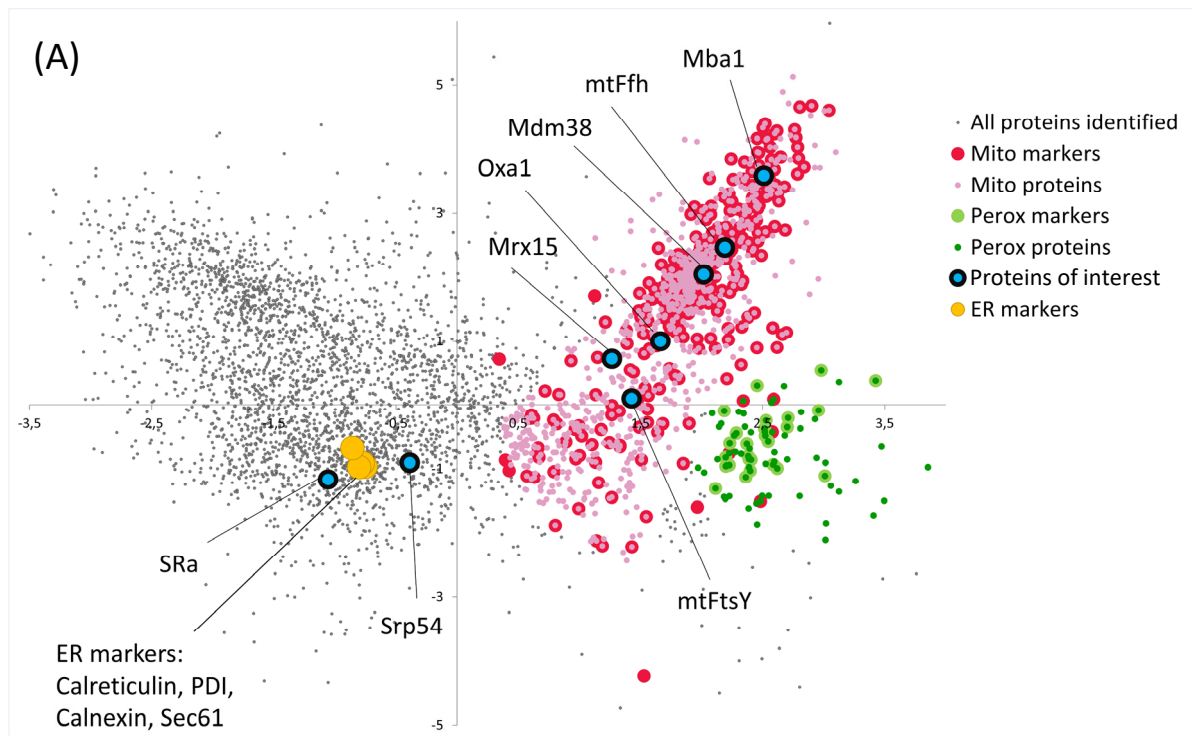
mtFfh

EDVEEAMQK
MFKSNSRSLLSRFLKTANTHHHCKIVFSSSEVFSPALCHHTSLFSSNNKNVGLQHVLQQYNQQQVRFFHSSPKQYGLFSNLTSNMSDAFTNLFKKKTLTKEDVEEAMQK

VALLDADVAESVVSQFVK   SVIMVTGIQGSGK
              EATEDAIGIAAVK   LSELLGGGVAPLELIPSTEGK
VRVALLDADVAESVVSQFVKEATEDAIGIAAVKQVESQENKGILSRIAKWVGPNQKAEELPKSATVYLMVMDRLSELLGGGVAPLELIPSTEGKSVIMVTGIQGSGKTT

                          AMEYVGPNGEQFDTVIFDTAGR   AIVQPNETLLVAD
                                         MHIDEQLMTELEELR
SSAKLALQLKRKENRRVLLVSLDTYRPAAQMQLQTLAQQIQVESLPIIPEQNPIEIAKRAMEYVGPNGEQFDTVIFDTAGRMHIDEQLMTELEELRAIVQPNETLLVAD

SMLGNDAVNIATQFHDR   MDDLETFDPQSLAK   AKDAMNLDAAQAQQK   GAYT
                                              RILGGGDIMTLAQK
SMLGNDAVNIATQFHDRVGLSGIVLTRMDGTSSGGCAISMKQVVGLSVKYIGIGERMDDLETFDPQSLAKRILGGGDIMTLAQKAKDAMNLDAAQAQQKLVQFSKGAYT

FKDYQDQIQMLK   MLKMYEK
  DYQDQIQMLK   DKLDNMDLSFIDNHEHIISCMSEQEK   AKVEVTDVNK
FKDYQDQIQMLKKMGSLKNMASYLPEQFIGKFRDKLDNMDLSFIDNHEHIISCMSEQEKLQPLLVESSSARRLDLAKRAKVEVTDVNKMLKMYEKMKSFISKAGSAAMK

DPKKMSEQMMKDPMFFANMFPMKVKKQLIRPPKR

mtFtsY

KIFNEEDFKLLEK   QIENNPTAEENVKPMK   RL
          ALLSSDVGNNTTQLLLTR   QIENNPTAEENVK   L
MFNKILGGLSKTSIFSKIGSILGGGATSRKIFNEEDFKLLEKALLSSDVGNNTTQLLLTRMKTQVSEIEKQIENNPTAEENVKPMKNILREEMLKLFQYPMQQQIVKRL

QQNQGTPEGGVSIPLIPLNPK   TSGCTLVDLTESEK
QQNQGTPEGGVSIPLIPLNPK
QQNQGTPEGGVSIPLIPLNPKSRPTVVQICGVNGSGKTTTIGKLLHKYRESGTVRNIVVAAADTVRAAAPDQLRTWVERTSGCTLVDLTESEKLKKQHTTTPQQVKVKN

VPAESVVYEAIQQALRKEDVDLVFVDTAGR
VPAESVVYEAIQQALR
              KEDVDLVFVDTAGR   ELSLITEMCSK
IRQVSYKVPAESVVYEAIQQALRKEDVDLVFVDTAGRLQNQEASMKELSLITEMCSKSRKGAPDHTWLILDGTIGQNSIQQAKLFQKYVRISGIIVTKLDGSAKGGVIL

AIANELKIPVLYIGLGESVSDLKPFYPEQFVDSILSVASAEEKQTKSEEDDE

Supplementary fig. S1. Proteomic identification of mtFfh, mtFtsY, and other mitochondrial proteins of interest in cell fractions of *N. gruberi*. (**A**) PCA analysis of 4,198 proteins proteomically measured in cell fractions differentially enriched in mitochondria. The cluster of mitochondrial proteins was defined based on 376 mitochondrial markers. The boundaries of the cluster of co-purified peroxisomal proteins were established by 26 peroxisomal markers, the position of four selected endoplasmic reticulum (ER) markers is also shown. All *N. gruberi* proteins specifically discussed in the main text as supposedly mitochondrial, including mtFfh and mtFtsY, clearly co-purify with the mitochondrial markers, in contrast to the components of the cytoplasmic/ER SRP system (SRP54 and SRa) following the ER markers. (**B**) Peptides (in red) from *N. gruberi* mtFfh and mtFtsY identified by tandem mass spectrometry (nLC-MS2). The N-terminal regions of the two proteins highlighted in yellow correspond to the leader sequences tested (as fusions with mNeonGreen) in subcellular targeting experiments in the heterologous system of *Trypanosoma brucei*. Further technical details on the proteomic analysis are provided in Horváthová et al. (2021).

(A) *N. gruberi* mtFfh gene: revised sequence and model

(B) Alignment of mtFfh protein sequences from *Naegleria* spp.

(C) *N. gruberi* mtFtsY gene: revised model

(D) Alignment of mtFtsY protein sequences from *Naegleria* spp.

```
Ngr_mtFtsY    MFNKILGGLSKTS--IFSKIGSILGGGATS-RKIFNEEDFKLLEKALLSSDVGNNTTQLLLTRMKTQVSEIEKQIENNPTAEENVKPMKNILREEMLKLFQYPMQQQIVKRLQQNQGTPEGGVSIPLIPLNPKSRPTVVQICGVNGSGKT
Nfo_mtFtsY    MFSRILGGLSKTSSTIFSKIGSILGGSANAQRKVFNEEDFKLLESALLSSDVGNNTTQLLLQRMKSQVTNIEKELENNPNMTD-VKPMKSILREEMLKLFQNPMQQQIVKRLQQNQGTPEEEISVSLTPLSMKSRPTVVQICGVNGSGKT
Nlo_mtFtsY    MFNRILGGLSKTSSTIFSKIGSILGGSSTGPRKVFNEEDFKLLESALLSSDVGNNTTQLLLQRMKTQVTNIEKELELNPNMTD-IKPMKSILREEMLKLFQNPMQQQIVKRLQQNQGTAEQEISVSLIPISVKSRPTVVQICGVNGSGKT
              **.:*********  **********.:.. **:*********.****************.***:**::***::* **.  : :****.*********** ***************.*  :*:.* *:. ******************

Ngr_mtFtsY    TTIGKLLHKYRESGTVRNIVVAAADTVRAAAPDQLRTWVERTSGCTLVDLTESEKLKKQ-HTTTPQQVKVKNIRQVSYKVPAESVVYEAIQQALRKEDVDLVFVDTAGRLQNQEASMKELSLITEMCSKSRKGAPDHTWLILDGTIGQNS
Nfo_mtFtsY    TTIGKLLHKYRQSGTVRHMIVAAADTVRAAAPDQLRSWVERTPDCSIVDLTESEKLKKQSE--THQPVKTKNIRQVSYKVPSESVVYEAIQQGLRKEDVDVVFVDTAGRLQNQEASMKELALINEMCSRSRKGAPDHTWLILDGTIGQNS
Nlo_mtFtsY    TSIGKLLHKYRQSGTVRHMIVAAADTVRAAAPDQLRSWVERTPDCSIVDLTESEKLKKQSESATHQPVKTKNIRQVSYKVPAESVVYEAIQQGLRKEDVDVVFVDTAGRLQNQEASMKELALINEMCSRSRKGAPDHTWLVLDGTIGQNS
              *:*********:*****::::****************:*****..*::************   .  * * **.**********.*********.*******:***********************:**.****:**********:*********

Ngr_mtFtsY    IQQAKLFQKYVRISGIIVTKLDGSAKGGVILAIANELKIPVLYIGLGESVSDLKPFYPEQFVDSILSVASAEEKQTKSEEDDE
Nfo_mtFtsY    IQQAKLFQKYVRISGIIVTKLDGSAKGGVILAIANELKIPVLYIGLGESVQDLRPFYPEQFVDSILSVTEQNNEKQDDDEEE-
Nlo_mtFtsY    IQQAKLFQKYVRISGIIVTKLDGSAKGGVILAIANELKIPVLYIGLGESVQDLRPFYPEQFVDSILSVTEQTKEKPQDEEEE-
              **************************************************.**:************:.  :::  ..:*::
```

Supplementary fig. S2. Revisions of existing models of *Naegleria gruberi* mtFfh and mtFtsY genes. (**A**) and (**C**) Nucleotide sequences of the respective genes extracted from the current genome assembly; mtFfh gene – GenBank accession number ACER01000115.1 (positions 64001 to 66000, reverse complement sequence), mtFtsY – GenBank accession number ACER01000034.1 (positions 330136 to 331575). Coding sequences as delimited in the existing gene models are highlighted in green, extensions of the coding sequences resulting from model revisions are highlighted in turquoise. Conceptual translation in all three sense frames is shown beneath the sequences; amino acids constituting the currently predicted protein (mtFfh – XP_002680081.1, mtFtsY – XP_002682642.1) are in red, N-terminal extensions resulting from revised gene models are in blue. The sequence of the mtFfh gene is impacted by a one-nucleotide deletion (highlighted in yellow), as is apparent from a comparison with raw Sanger sequencing reads (segments of four exemplar reads aligning to the region containing the assembly error are shown above the gene sequence). (**B**) and (**D**) Multiple alignments of revised (*N. gruberi*) or *de novo* predicted (*Naegleria fowleri* and *Naegleria lovanensis*) sequences of mtFfh and mtFtsY proteins, respectively. The alignments are shown to document that homologous sequences from other *Naegleria* species support the validity of the revisions of the existing *N. gruberi* gene modes. Note that the N-terminus of mtFfh in *Naegleria* spp. is poorly conserved and it is possible that in case of *N. fowleri* and *N. lovanensis* the actual start of the protein is at a methionine residue downstream of that used in the scheme. Still, the alignment clearly supports the extension of the N-terminal region of the *N. gruberi* beyond the existing model.

Supplementary fig. S3. Multiple sequence alignment of bacterial and mitochondrial Ffh proteins. All detected mitochondrial homologs possess N-terminal extensions compared to bacterial proteins (positions 1–171). The alignment was constructed in Geneious Prime 2019.2.3 using the Geneious alignment tool with default settings. The full alignment contains 750 positions, only positions 1–270 are shown. Note that the sequences from *Percolomonas* sp. ex *Nitzschia* and *Percolomonas cosmpolitus* WS are truncated and the N-terminal extensions of the respective proteins are presumably longer than shown in the figure.
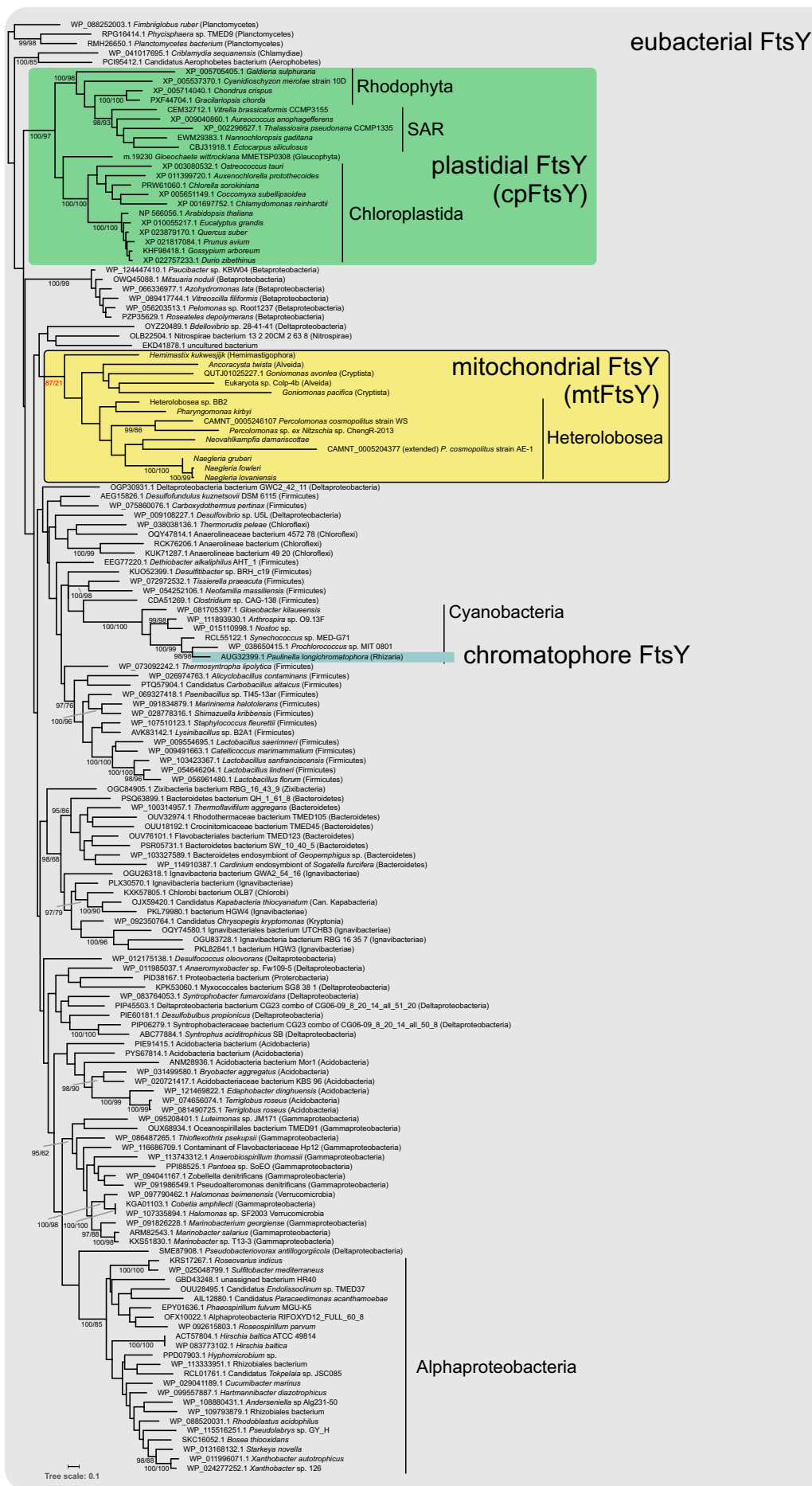
Supplementary fig. S4. Phylogenetic analysis of eubacterial and organellar FtsY proteins. The tree topologies shown correspond to maximum likelihood tree (LG4X substitution model) computed with IQ-TREE based on a multiple alignment of 154 sequences and 287 amino acid positions. Branch support was assessed by ultrafast bootstrapping (N=1000, IQ-TREE) and rapid bootstrapping (N=500, RAxML) using the same model. The full trees from both methods are also provided in the Newick format in supplementary dataset S1, Supplementary Material online.

(A)

```
                          10        20        30        40
                 ....|....|....|....|....|....|....|....|....|
Naegleria gruberi      MKLFKFYVMNEVILIDADGSSKCNFLAIDFSFCNNYYNLEHYLI*   AF288092.1:8630..8764
Naegleria fowleri      MKLFKFYNLSETILIGHDGSSEYLFLTLDFNFCNNFYNIEIYLI*   JX174181.1:8283..8417
Naegleria lovanensis   MKLFKFYNLNETIIINSDGSCEYLFLTLDFNFCNNFHNIEIYLI*   PYSW01000110.1:8249..8383
Willaertia magna       MKLLRLNSFTEVTLIHFDGASKFHFLALDFSFCNNYYNLEQYVI*   Scaffold_1522:4839..4973
```

(B)

ffs

```
mtFfh Naegleria gruberi            FIDNHEHIISCMSEQEKLQPLLV---ESSSARRLDLAKRAKVEVTDVNKMLKM  528
mtFfh Naegleria fowleri            FFDAHEGIISHMSEQEKLQPLLL---QSSSARRLDLAKKAKVDITEINKMLKM  666
mtFfh Naegleria lovaniensis        FFDAHEGIINHMSEQEKLQPLLI---QSSSARRLDLAKKAKVDITEINKMLKM  570
mtFfh Percolomonas cosmopolitus WS KLEDFRIIYAEMQPDERRNPMLI---KVSSARRMRVIKSSGVTLKRIQECLQM  504
mtFfh Percolomonas sp. ex Nitzschia SSEAAARAIYQSMENDERMNPLLI---KMSSIRKMRISKATGLTIKNVLEEIVH  460
mtFfh Percolomonas cosmopolitus AE-1 QHSTAIRVVDFMTQEEKENPILI---RRMASRRKTLASRAGVAITDVSAMIKD  499
mtFfh Neovahlkampfia damariscottae FIDVHLRFINLMTEKEKRNPNMF---KQMASCRIGLAKRAKVPVVEVNKFLKM  480
mtFfh Pharygomonas kirbyi          DREQQLAIIKQLSDQDKSNPKLI---SRSGAKKKAVAQAGVQIVAVNRMLKQ  547
mtFfh Heterolobosea sp. BB2        LIATQIAILEAMTEEERRNPEIT---KSSSTRKTKLAAAKVQMIEVNRLFKM  559
mtFfh Ancoracysta twista           AFQKEEALLNAMTDEERSIPNSI-----KGPRKKALAKKCGQSIHDVNRLIKK  466
mtFfh Eukaryota sp. Colp-4b        GLALEESLLMAMTDEERANPSTI-----RGPRKKALAKSCSQSIHEVHKLLKK  474
mtFfh Goniomonas avonlea           QFDVEAKVVSHMTPEERARPKII-----KADRKLALATAAGTDSAAINHMIKR  477
mtFfh Goniomonas pacifica          DMATLERITGMMTEQEKQQPDRL-----PAQRKKQLATELNMELSDLNVIFRR  463
mtFfh Hemimastix kukwesjijk        TLGINDMLLKAMTDQERKNPALI---ASSAPLRRLAEATGQSVTDVNRLIMH  483
mtFfh Spironema cf. multiciliatum  NFYIHEKLIRAMTPEQRANPSFI---NNSVARQ-KLADAAGQPLSAVLNLVRE  559
cpSRP54 Arabidopsis thaliana       NLLVMEAMIEVMTPEERERPELL---AESPERRKRIAKDSGKTEQQVSALVAQ  494
cpSRP54 Physcomitrella patens      SLKIMESMINSMTPKERADPDLL---AKSPSRRRRVANGSGRSQEQVSALVAQ  543
cpSRP54 Chlamydomonas reinhardtii  QFGVYEAIIGAMDEEERSNPEVI---IKNLARRRRVAQDSGKSEAEVTKLMAA  482
cpSRP54 Coccomyxa subellipsoidea   QFKMFEALINSMTPEERTNPDLL---AKTASRRKRLARGAGRIEVDVTNMIGT  418
cpSRP54 Fragilariopsis cylindrus   RLKKNEAMIAVMTEEERSNPDLLIKDSKALERLIRLAKFSDMPLSDVKQFMSE  441
cpSRP54 Phaeodactylum tricornutum  RLKKSKSMISSMTKKERANPELLIKDSSARSRLIRITKGSGCGLDEGGQFMSE  414
cpSRP54 Cyanidioschyzon merolae    RLRLAESVINSMTPRERANPDLLTVDKTAASRMRRIARGSGRSLEQVESLMRD  531
Ffh Wolbachia endosymbiont         RVKKYIAIINSMTKKEKRDPDIL-----NCKRRLRIAKGSGTNVMDINLLIKQ  416
Ffh Magnetospirillum magneticum    MVARQEAIITSMTKAERRNPDLI-----KASRRILARGAGVEVQDVNKLLKQ   418
Ffh Escherichia coli               VLVRMEAIINSMTMKERAKPEII-----KCSRKRRIAAGCGMQVQDVNRLLKQ  419
Ffh Bacillus subtilis              QINHVEAIIKSMTVLEKEQPDLL-----NASRRKRIAKGSGTSVQEVNRLLKQ  419
                                           SM                   GXG
```

Supplementary fig. S5. No evidence for mitochondrial 4.5S RNA. (**A**) A novel short conserved unannotated ORF found in mitogenomes of certain Heterolobosea when searching for potential missed *ssrA* gene in intergenic regions. Accession numbers of the respective nucleotide sequences and coordinates defining the ORF are given to the right. The mitogenome sequence from *Willaertia magna* was detected in the whole genome assembly reported by Hasni et al. (*Scientific Reports* 2019, 9:18318, doi:10.1038/s41598-019-54580-6). (**B**) Mitochondrial Ffh has substituted the amino acid residues critical for binding 4.5S RNA (the RNA component of the signal recognition particle). Displayed is a segment of a multiple sequence alignment of mtFfh (the protein component of the mitochondrial SRP) and their homologs, including plastidial cpSRP45 and bacterial Ffh. The region shown includes two motifs involved in binding of the 4.5S RNA molecule (the four residues critical for the binding are highlighted in red). The red rectangles on the left indicate the presence of a discernible 4.5S RNA-specifying *ffs* gene in the respective organellar (mtFfh, cpSRP54) or bacterial (Ffh) genome. Note the perfect correlation between the conservation of the 4.5S RNA-binding residues in the protein and the presence of the *ffs* gene.

Tree scale: 1

Oxa1

Tetr. thermophila
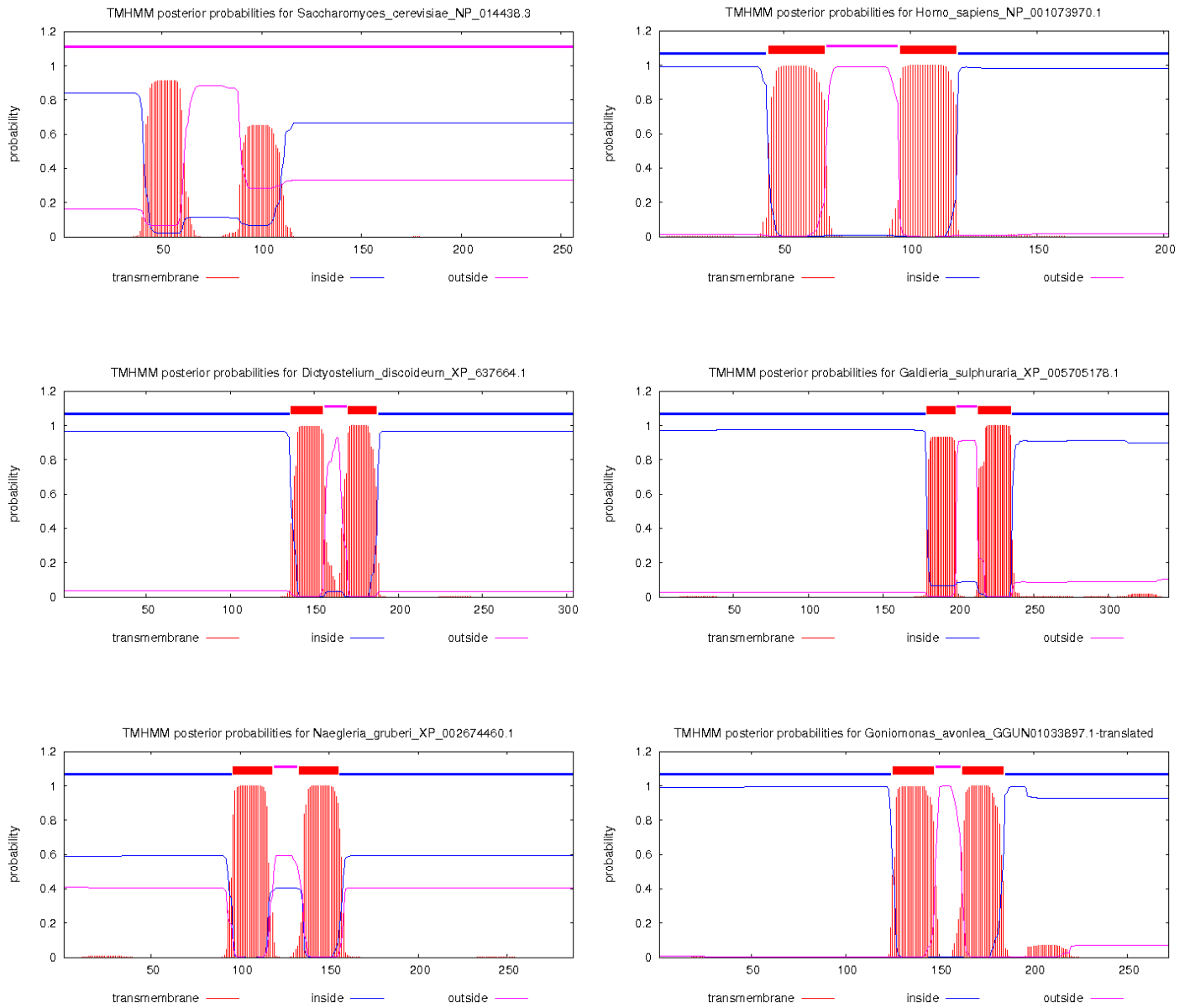XP001014647.2

Cox18/Oxa2

99

Supplementary fig. S6. Phylogenetic analysis of the YidC/Alb3/Oxa1 family. A selection of bacterial YidC sequences and their eukaryotic homologs, i.e. the plastidial Alb3 and the mitochondrial Oxa1 and Cox18, was aligned using PASTA (Mirab et al. 2015 - doi.org/10.1089/cmb.2014.0156) and trimmed manually to remove poorly conserved regions. The final alignment (222 amino acid positions, 251 sequences) was subjected to tree inference using IQ-TREE multicore version 1.6.12, with the LG+C20+F+G4 substitution model and 1000 ultrafast bootstraps (bnni). Bootstrap values are shown when ≥70%. The tree was arbitrarily rooted between YidC/Alb3 and Oxa1/Cox18 groups and separated into two parts, each displayed on a separate page. The resolution of the tree is inherently limited by the low sequence conservation and the small length of the alignment. As a result, the Oxa1 and Cox18, presumably two separate deeply diverged eukaryotic paralogs, are not clearly separated in the tree (Oxa1 is paraphyletic). The sequences from mtFfh/mtFtsY-carrying species are highlighted in boldface; their sequence IDs are provided in supplementary table S5. Note that the Oxa1-2 sequence from *Hemimastix kukwesjijk* (Hemi2|18767_TR6208_c0_g1_i1) is very short and hence was not included in the analysis.

Supplementary fig. S7. Putative ribosome-binding C-terminal extensions in Oxa1 proteins of diverse eukaryotes. The figure shows multiple sequence alignment of the C-terminal segment, including the last transmembrane helix (highlighted in a gray box) and the regions downstream, of selected representatives of the YidC/Alb3/Oxa1 family. Note the presence of C-terminal extensions (poorly conserved in sequence and length) in the eukaryotic Oxa1 proteins, including those from mtFfh/mtFtsY-carrying taxa (highlighted in red) compared to the reference bacterial YidC sequences. The extensions of some sequences (marked with an asterisk) are in fact longer than apparent from the figure, as the available sequences are truncated. Sequence IDs of proteins from the mtFfh/mtFtsY-carrying taxa are provided in supplementary table S5. The classification of the sequences into the Oxa1 and Cox18 subfamilies is based on the phylogenetic analyses presented in supplementary fig. S6.

Supplementary fig. S8. Two transmembrane domains are a conserved feature of proteins of the Mrx15/TMEM223 family. Selected representatives of the family (identified with PSI-BLAST of HMMER searches) were evaluated by the TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/). Note that despite the below-threshold probability of transmembrane domains in the *S. cerevisiae* Mrx15 protein (top left), their presence has been confirmed experimentally (Möller-Hergt et al. 2018).

Chapter Summary

Here we report on the presence of homologs of the bacterial Ffh and FtsY proteins in various unrelated plastid-lacking unicellular eukaryotes, with the following outcomes:

1. *Naegleria gruberi* possesses mitochondrial homologs of Ffh and FtsY.

2. Mitochondrial signal recognition particle lacks the RNA component.

3. N-termini of some *N. gruberi* mitochondrial proteins function as signal peptides.

4. Proteins implicated in the mitoribosome association in opisthokonts are widespread in eukaryotes.

5. Identification a novel mitochondrial protein (MAP67) present in diverse eukaryotes and related to the signal peptide-binding domain of Ffh.

Chapter III.

Orthologues of presequence translocase-associated motor subunits are essential for kinetoplast DNA replication in procyclic form of *Trypanosoma brucei*

Manuscript in preparation

Orthologues of presequence translocase-associated motor subunits are essential for kinetoplast DNA replication in procyclic form *Trypanosoma brucei*

Corinne von Känel[1], Bettina Warscheid[2], Vendula Rašková[3,4], Ignacio Durante[3], Silke Oeljeklaus[2], Andre Schneider[1]

[1]Department of Chemistry and Biochemistry, University of Bern, Bern, Switzerland

[2]Biochemistry and Functional Proteomics, Institute of Biology II, Faculty of Biology and Signalling Research Centres BIOSS and CIBSS, University of Freiburg, Freiburg, Germany

[3]Institute of Parasitology, Biology Centre, České Budějovice, Czech Republic

[4]Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

Abstract

The majority of mitochondrial proteins is encoded in the nuclear genome, synthesized in the cytosol, and then imported into mitochondria. Mitochondrial protein import has been most studied in the yeast *Saccharomyces cerevisiae* and much of this work can be generalized to many eukaryotic lineages. However, analysis of the protein import machinery in the parasitic protist *Trypanosoma brucei* revealed that its import complexes are highly diverged from the ones found in other eukaryotes. In yeast, the driving force to import proteins into the mitochondrial matrix is mostly provided by the presequence translocase-associated motor (PAM). Several homologues of PAM components were found in the *T. brucei* mitochondrial proteome. Among these are orthologues of the yeast PAM subunits Pam18 and Pam16, which were termed TbPam18 and TbPam16. Surprisingly, TbPam18 and TbPam16 are not associated with the trypanosomal PAM and not involved in mitochondrial protein import. Instead, we show that TbPam18 and TbPam16 play lifecycle-specific roles in the replication of the mitochondrial DNA. The single-unit mitochondrial genome of *T. brucei*, termed kinetoplast DNA, consists of a complex network of two types of interlocked DNA rings, the mini- and maxicircles. Intriguingly, TbPam18 and TbPam16 are specifically involved in maxicircle replication, in a not yet determined, possibly regulatory, process linked to the inner mitochondrial membrane.

Introduction

The parasitic protist *Trypanosoma brucei* exhibits several unusual biological properties. One of the most striking features is the single, reticulated mitochondrion with its particular single-unit mitochondrial genome, termed kinetoplast DNA (kDNA). The kDNA consists of two types of DNA rings, called mini- and maxicircles, which are arranged into a complex intercatenated network. There are approximately 25 maxicircles, each of which is 23 kilobases (kb) in size[1,2]. They encode for 16 genes of the oxidative phosphorylation (OXPHOS) pathway, two mitochondrial ribosomal proteins (MRPs), as well as the 9S and 12S mitoribosomal RNAs[1–3]. Twelve of these maxicircle transcripts need to be post-transcriptionally modified in a process termed RNA editing[4–6]. RNA editing depends on small guide RNAs, almost all of which are encoded on the minicircles[6]. The heterogenous minicircles are approximately one kb in size, but with roughly 5000 copies, they are much more abundant than the maxicircles and make up the bulk of the kDNA network[1,7]. Each minicircle is linked to three other minicircles and maxicircles are interwoven into the minicircle network[8]. Additionally, all maxicircles seem to be interlocked with each other[9]. These networks are condensed into a disk-like structure. The kDNA disk is positioned in a specific region of the mitochondrial matrix, near the basal body of the flagellum that is located outside of the mitochondrion[2,10]. It is not surprising that the replication of this complex kDNA network is a complicated process and estimates predict that up to 150 proteins must be involved[1,2].

The replication of minicircles starts with their individual release into the region be- tween the kDNA disk and the inner mitochondrial membrane (IM), termed the kinetoflagellar zone (KFZ)[11]. In the KFZ, replication is initiated and proceeds unidirectionally via theta structures[2,10,12]. Replicated minicircles then migrate to the antipodal sites, which are protein complexes at opposing sites of the kDNA disc. At these antipodal sites, most gaps between Okazaki fragments are repaired and minicircles are attached to the periphery of the kDNA network[2,13,14]. Maxicircles are also replicated unidirectionally *via* theta structures. However, in contrast to minicircles, maxicircles are never released from the kDNA network and are replicated while being interlocked with minicircles and other maxicircles. This makes it considerably harder to study the exact mechanism of maxicircle replication and, therefore, the process is less well understood[2,10,15]. The cycle of minicircle release from the network and reattachment to the poles passively concentrates the still catenated maxicircles in the central zone[9]. The replicated maxicircles eventually serve as the connection between the two daughter minicircle networks[2,10,16]. The two replicated kinetoplasts are segregated by the separating basal

bodies that are connected to the kDNA through a protein complex termed tripartite attachment complex (TAC)[17,18]. The last stepin kDNA separation is the cleavage of the maxicircles that are still connecting the segregating networks, to unlink the daughter kDNAs[2,10,19].

This complicated replication mechanism, combined with the equally complex RNA editing of most kDNA transcripts, their further post-transcriptional processing and mitochondrial translation processes, ensure the expression of the small number of mitochondrially encoded genes. However, the few proteins encoded in the kDNA, make up only a fraction of the proteins needed by the mitochondrion to fulfil the multitude of its functions. The total mitochondrial proteome of *T. brucei* comprises roughly 1200 proteins[20]. Like in other eukaryotes, most mitochondrial proteins are encoded in the nuclear genome, translated by cytosolic ribosomes and then imported into the organelle[21]. This process known as mitochondrial protein import, has been best studied in theyeast *Saccharomyces cerevisiae*. Yeast, as well as most other studied eukaryotes, possesses two translocases of the IM (TIM), the TIM22 and TIM23 complexes[22,23]. The TIM22complex mediates the insertion of proteins with multi-spanning membrane domains, suchas mitochondrial carrier proteins (MCPs), into the IM[24,25]. The TIM23 complex imports presequence-containing precursor proteins, which represent 60 to 70% of all mitochondrial proteins, across or into the IM[26–28]. To import presequence-containing proteins into the mitochondrial matrix of yeast, TIM23 associates with the matrix-exposed presequence translocase-associated motor (PAM). The yeast PAM consists of five essential components[27,29] that are well conserved across eukaryotes[22]. This includes the mitochondrial heat shock protein 70 (mHsp70)[30,31], its co-factors Pam18[32,33] and Pam16[34], Tim44, which tethers mHsp70 to TIM23[35], as well as the nucleotide exchange factor Mge1[36–38].

Surprisingly, in *T. brucei* only a single TIM complex can be found. This trypanosomal TIM complex, with small compositional variations, can import presequence-containing proteins as well as MCPs[39]. To import presequence-containing proteins into the matrix, the trypanosomal TIM complex also associates with a PAM module[40]. This PAM contains an orthologue of mHsp70 (TbmHsp70) that was shown to be essential for mitochondrial presequence protein import[40,41]. Furthermore, *T. brucei* contains *bona fide* orthologues of Pam18 and Pam16, termed TbPam18 and TbPam16, that are essential for normal growth. Surprisingly, they are not involved in mitochondrial protein import[40]. The function of Pam18, and possibly Pam16, in the PAM of trypanosomes is carried out by the non- orthologous and euglenozoan-specific J-domain protein TbPam27[40]. Thus, during kineto- plastid evolution Pam18 and Pam16 must have been replaced in the import motor by TbPam27. The finding that TbPam18 and TbPam16 are retained

in trypanosomes and other kinetoplastids, even though they are no longer required to import proteins, is surprising.

Here, we present data suggesting that TbPam18 and TbPam16 are involved in the maintenance or the replication of the kDNA maxicircles in *T. brucei*. Interestingly, this function seems to be specific for procyclic form *T. brucei* and it is linked to the IM.

Results

TbPam18 and TbPam16 are interaction partners and present in a high molecular weight complex

Even though TbPam18 and TbPam16 are orthologues of the PAM subunits in other eukaryotes, identified in previous studies[40,39], they were not found to be associated with the single trypanosomal TIM complex. Furthermore, we have reported earlier that RNAi-mediated knockdowns of TbPam18 and TbPam16 lead to growth retardations after four days of RNAi induction. However, no evidence for an involvement in general mitochondrial protein import has been found. Therefore, TbPam18 and TbPam16 are no components of the trypanosomal PAM[40]. We thus wondered whether TbPam18 and TbPam16 have interaction partners outside of the import machinery.

First, we wanted to find out, if C-terminally HA-tagged versions of TbPam18 and TbPam16 (TbPam18-HA and TbPam16-HA) are functional proteins. This was done in order to establish whether the tag interferes with their function(s) and whether previous experiments that have been done with TbPam18-HA or TbPam16-HA should be re-evaluated. To that end, we generated RNAi-resistant (RNAi-res.) TbPam18-HA and TbPam16- HA variants. The codons in regions of the open reading frame (ORF) that are targeted by RNAi were changed such that their transcripts are resistant to RNAi, but still translate into the same amino acid sequence as in the endogenous protein. We then established cell lines allowing the inducible ectopic expression of these RNAi-res. TbPam18-HA or TbPam16-HA versions, in the background of RNAi against the respective endogenous proteins (exclusive expressor cell lines). A growth curve of the respective cell line showed that RNAi-res. TbPam18-HA cannot complement the growth retardation caused by the ablation of the endogenous TbPam18 (*Fig. 1A*). This result demonstrates that TbPam18-HA isnot a functional protein. Findings that have been, or will be, attained using TbPam18-HA, thus, should be interpreted carefully. *Fig. 1B* shows that TbPam16-HA, on the other hand,can fully complement the RNAi-induced growth phenotype and, therefore,

is a functional protein. Thus, findings obtained with TbPam16-HA are representative of the wildtype version of TbPam16.
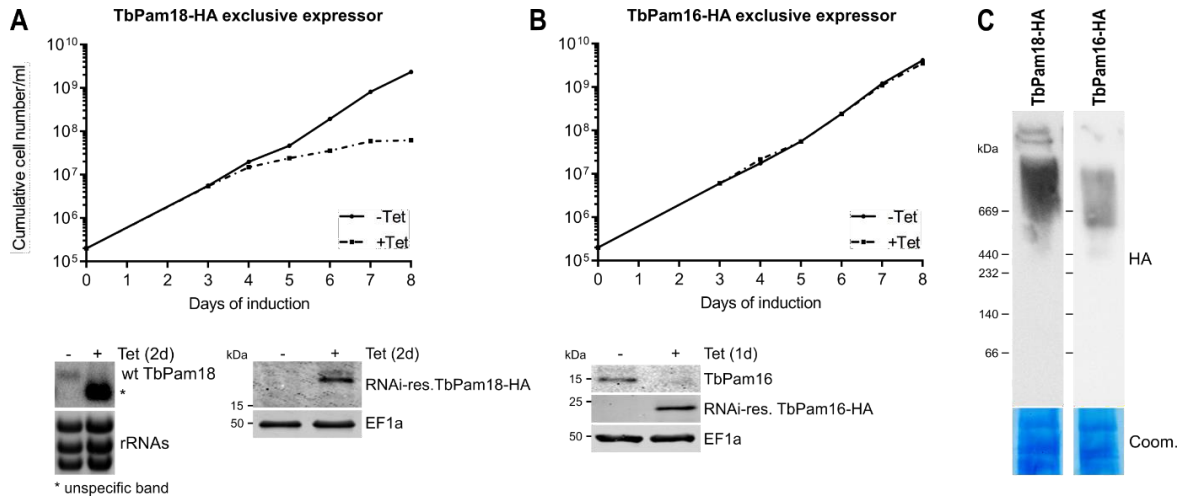
Figure 1 - TbPam18 and TbPam16 are present in a HMW complex: (A) Upper panel: Growth curve of uninduced (-Tet) and RNAi-induced cells (+Tet) expressing RNAi-resistant (RNAi-res.) TbPam18-HA inthe background of RNAi targeting the wildtype (wt) TbPam18 (TbPam18-HA exclusive expressor).Lower panel, left: Northern blot of total RNA extracted from uninduced (-) and two days induced cells (+), probed for the mRNA of wt TbPam18. Asterisk (*) indicates a prominent RNA product that results from the RNAi against TbPam18. Ethidium bromide-stained ribosomal RNAs (rRNAs) serve as loading control. Lower panel, right: Immunoblot analysis of whole cell extracts of uninduced (-) and two days RNAi-induced (+) cells probed for RNAi-res. TbPam18-HA and elongation factor 1 alpha (EF1a) as loading controls. **(B)** Upper panel: Growth curve of uninduced (-Tet) and induced(+Tet) cells expressing RNAi-res. TbPam16-HA in the background of RNAi targeting wt TbPam16 (TbPam16 exclusive expressor). Lower panel: Immunoblot analysis of whole-cell extracts of uninduced (-) and one day RNAi-induced (+) cells probed for TbPam16, RNAi-res. TbPam16-HA and EF1a serve as loading controls. **(C)** Blue native (BN)-PAGE analysis of digitonin-extracted, mitochondria-enriched fractions of cells expressing TbPam18-HA or TbPam16-HA. Coomassie-stained gel sections (Coom.) serve as a loading control.

TbPam16-HA has previously been localized to the IM[40]. Furthermore, TbPam16-HA was utilized as the bait in a quantitative co-immunoprecipitation (CoIP) experiment that was done under stable isotope labelling by amino acids in cell culture (SILAC) conditions and then analysed by mass spectrometry (MS)[40]. In this experiment, the only protein enriched to similar levels as the

bait was TbPam18 suggesting that the two proteins are interaction partners. Given the interaction of TbPam18 and TbPam16, we are confident that our previously reported finding that TbPam18, just as TbPam16, is localized to the IM[40] is also reliable, even though this result was obtained with the non-functional TbPam18-HA. However, a TbPam18-HA SILAC CoIP experiment done in our previous study, had resulted in the enrichment of numerous non-mitochondrial proteins and the interaction with TbPam16 was not detected[40]. Due to the non-functionality of TbPam18-HA the results of this SILAC CoIP can be considered as an artefact.

To learn more about the TbPam18-TbPam16 interaction, digitonin-extracted mitochondria-enriched fractions of cell lines expressing TbPam18-HA or TbPam16-HA, respectively, were subjected to a blue native (BN)-PAGE analysis. The results in *Fig. 1C* show that both, TbPam18-HA as well as TbPam16-HA, are present in high molecular weight complexes above 669 kDa. Since the result of TbPam18-HA phenocopies the result obtained with its interaction partner, the functional TbPam16-HA, we consider it reliable.

Taken together, the observations of TbPam18 and TbPam16 being present in a HMW complex (*Fig. 1C*) and of TbPam18 being TbPam16's only interaction partner[40], suggest that the two proteins exist in a hetero-oligomeric complex with each protein being present in multiple copies.

Ablation of TbPam18 or TbPam16 predominantly affects MRPs and OXPHOS components

To find out more about the function of TbPam18 and TbPam16, we analysed global changes in the mitochondrial proteome upon RNAi-mediated TbPam18 and TbPam16 depletion. To that end, we executed a quantitative proteomic analysis of the steady-state levels of mitochondrial proteins in the TbPam18 and TbPam16 RNAi cell lines by using SILAC combined with MS. Uninduced and induced TbPam18 or TbPam16 RNAi cells were grown in a medium containing stable isotope-labelled forms of arginine and lysine. After four days of RNAi induction, which corresponds to the onset of the growth retardations[40], equal numbers of uninduced and induced cells were mixed, and mitochondria-enriched fractions generated by digitonin extraction were analysed by quantitative MS.

Surprisingly, neither TbPam18 nor TbPam16 were detected in either of the two SILAC RNAi experiments. To elaborate on this observation, we have obtained information on the steady-state levels of TbPam16 in the RNAi cell lines, using a newly generated TbPam16 antibody (see Materials and Methods). After only one day of TbPam18 or TbPam16 RNAi induction, respectively, TbPam16 levels are strongly reduced (*Fig. 2A*). These results show that TbPam16

stability is highly dependent on the presence of TbPam18 and that RNAi mediated TbPam16 turnover is very efficient and fast. Attempts to generate an antibody against TbPam18 were not successful.
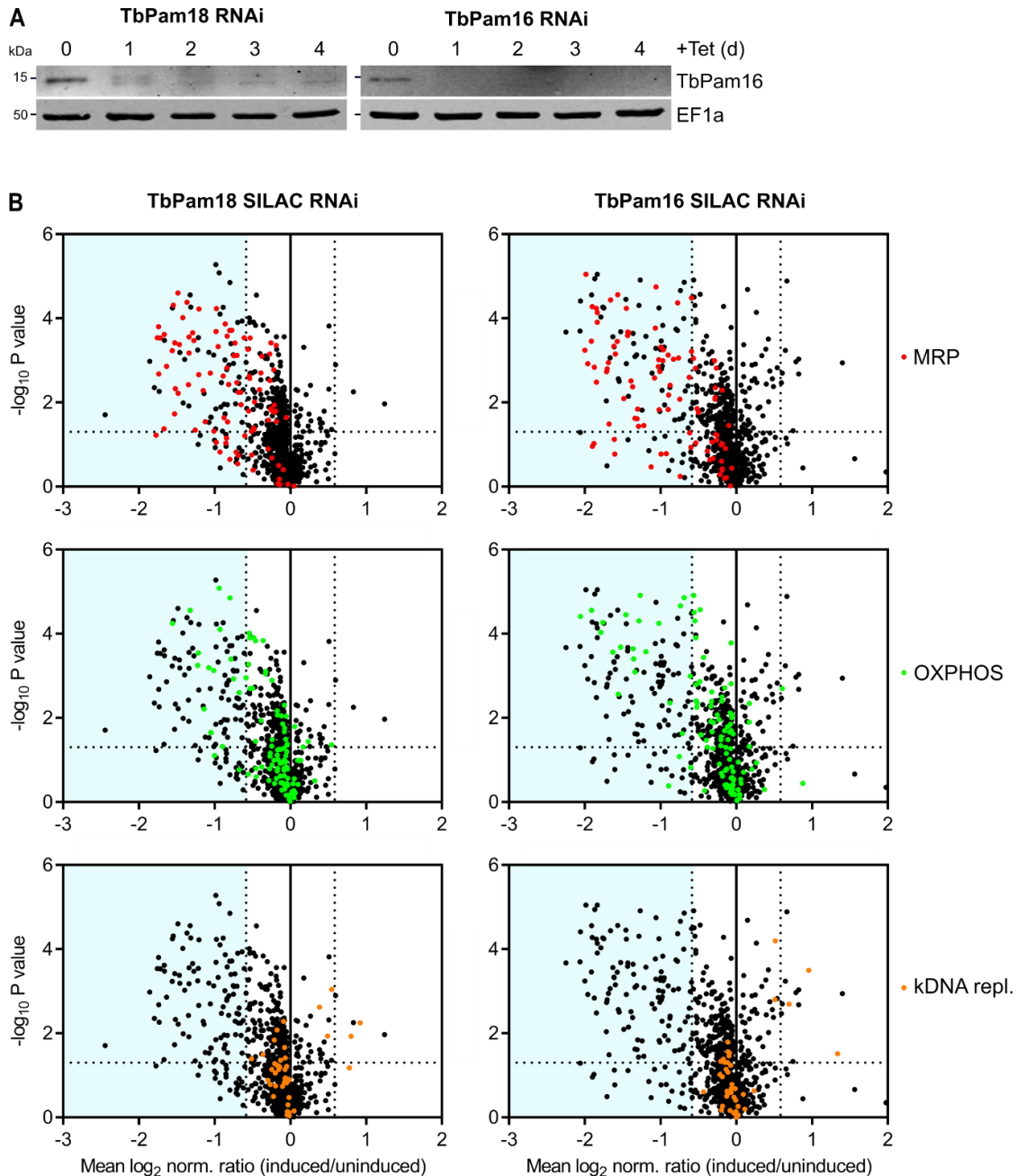
Figure 2 - TbPam18 and TbPam16 RNAi predominantly affects MRPs and OXPHOS components. Immunoblot analysis of steady-state protein levels of TbPam16 in whole-cell extracts of TbPam16 or TbPam18 RNAi cell lines over four days of induction. EF1a serves as a loading control.

Global mitochondrial proteome changes upon ablation of TbPam18 (left panels) or TbPam16 (right panels). Mitochondria-enriched fractions of uninduced and four days RNAi-induced TbPam18 and TbPam16 RNAi cells were analysed by SILAC-based quantitative MS. Datasets were filtered for mitochondrial proteins[20] and the mean log2 of normalized ratios (induced/uninduced) was plottedagainst

the corresponding log10 P value (two-sided t-test). Datasets were then filtered for mitochondrial ribosomal proteins (MtRPs, top panels, red)[3], components of the oxidative phosphorylation pathway (OXPHOS, middle panels, green)[42] and proteins known to be involved in kDNA replication, maintenance and segregation (kDNA repl., bottom panels, orange)[1,2,18,42–50]. The horizontal dotted line in each volcano plot marks a t-test significance level of 0.05. The vertical dotted lines indicate a fold-change in protein abundance of ±1.5. The blue background marks a downregulation of 1.5-fold.

The proteins detected in the MS analysis of the TbPam18 and TbPam16 SILAC RNAi experiments, were filtered for mitochondrial proteins by using a *T. brucei* mitochondrial proteome reference dataset[20]. 943 and 899 mitochondrial proteins were detected, respectively, which corresponds to 84.3% and 80.3% of the mitochondrial proteome (*Fig. 2B*)[20]. In the TbPam18 SILAC RNAi experiment, 13.5% of the detected mitochondrial proteins are downregulated more than 1.5-fold. In the TbPam16 SILAC RNAi experiment this number constitutes 15.6%. The 50 most downregulated proteins in the respective datasets are listed in *Tables S1 and S2*. To determine which proteins are predominantly affected by the TbPam18 and TbPam16 depletions, we filtered the datasets for different protein groups. Interestingly, in the TbPam18 dataset, 61.6% of all detected mitochondrial ribosomal proteins (MtRPs)[3] showed a 1.5-fold decreased steady-state level (*Fig. 2B, top panel*). Upon TbPam16 RNAi, this number amounts to almost equal fraction, namely 65.6%. In both experiments two thirds of the more than 1.5-fold downregulated MtRPs are components of the large subunit (LSU) and one third are proteins of the small subunit (SSU) of the mitochondrial ribosome (mitoribosome)[3]. Furthermore, we found that in the TbPam18 and TbPam16 SILAC RNAi experiments, 17.3% and 20.9% of all detected components of the OXPHOS pathway[42] are reduced more than 1.5-fold, respectively (*Fig. 2B, middle panels*). In both experiments, respiratory complex IV is affected the most, followed by complexes I, III and V. It is worth noting that while detected and quantified in the experiments, subunits of complex II were not found to be downregulated more than 1.5-fold.

A common feature of the LSU and SSU of the mitoribosome, as well as the OXPHOS complexes I, III, IV and V is that they contain components encoded in the kDNA. Coding regions of not only the 12S and 9S mitoribosomal RNAs, but also two MtRPs can be found in the kDNA[3]. Furthermore, six subunits of complex I, one complex III subunit, three subunits of complex IV and one complex V subunit are encoded in the kDNA[1,3,51]. In contrast, all complex II subunits are encoded in the nuclear genome[51].

In the next step, we filtered the datasets for proteins that are associated with kDNA replication, maintenance, and segregation (*Fig. 2B, bottom panels*)[1,2,18,42–50]. 44 and 42 such

factors were detected in the TbPam18 or TbPam16 SILAC RNAi datasets, respectively, none of which was downregulated more than 1.5-fold. Interestingly, two proteins, the small and large subunits of DNA topoisomerase IB, were found to be significantly upregulated more than 1.5-fold in the TbPam18 and the TbPam16 SILAC RNAi datasets. DNA topoisomerase IB was described to be involved in the replication of the nuclear as well as the mitochondrial genome[2,52,53]. The third protein significantly upregulated more than 1.5-fold upon TbPam16 RNAi, is the universal minicircle sequence binding protein 1 (UMSBP1). UMSBP1 is not only required for the initiation of minicircle replication, but is also involved in the regulation of kDNA segregation and nuclear division[2,54]. A subset of the kDNA replication factors we screened our datasets for, including the three proteins that have been significantly upregulated, are not part of the mitochondrial reference proteome[20]. These proteins, therefore, appear as additional datapoints in the volcano plots in *Fig. 2B*.

Taken together, the findings that upon TbPam18 and TbPam16 RNAi, numerous MtRPs and OXPHOS components are downregulated more than 1.5-fold, and certain kDNA replication factors are significantly upregulated, suggested that TbPam18 and TbPam16 ablation could affect the kDNA.

Depletion of TbPam18 and TbPam16 causes loss of maxicircles

To analyse the fate of the kDNA upon TbPam18 or TbPam16 depletion, we dissected DAPI-stained RNAi cells by fluorescence microscopy. Interestingly, we found that after four days of RNAi induction, at onset of the growth retardations[40], many TbPam18 and TbPam16 RNAi cells had visibly smaller kDNAs when compared to uninduced cells (*Fig. 3A, upper panels*). For each timepoint, the kDNA size in 86 to 140 cells was measured and the mean of the kDNA sizes in uninduced RNAi cells was set to 100% (*Fig. 3A, lower panels*). After three days of induction the mean of the kDNA size, when compared to uninduced cells, is significantly reduced to roughly 90% in TbPam18 and TbPam16 RNAi cells. This percentage decreases to 74% and 60% for TbPam18 and 64% and 55% for TbPam16 in four and five days induced cells, respectively. Shrinkage of the kDNA disc is a well-known phenomenon observed previously, when proteins involved in replication of mini- as well as maxicircles were ablated[49,50].

A

TbPam18 RNAi

TbPam16 RNAi

B

TbPam18 RNAi
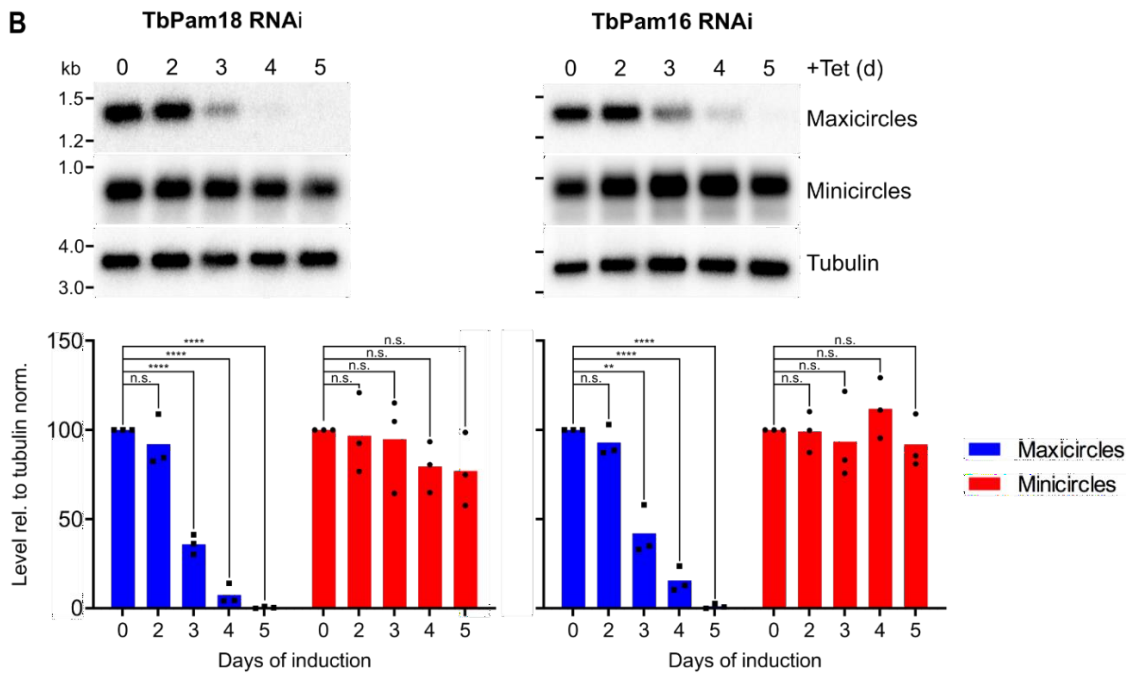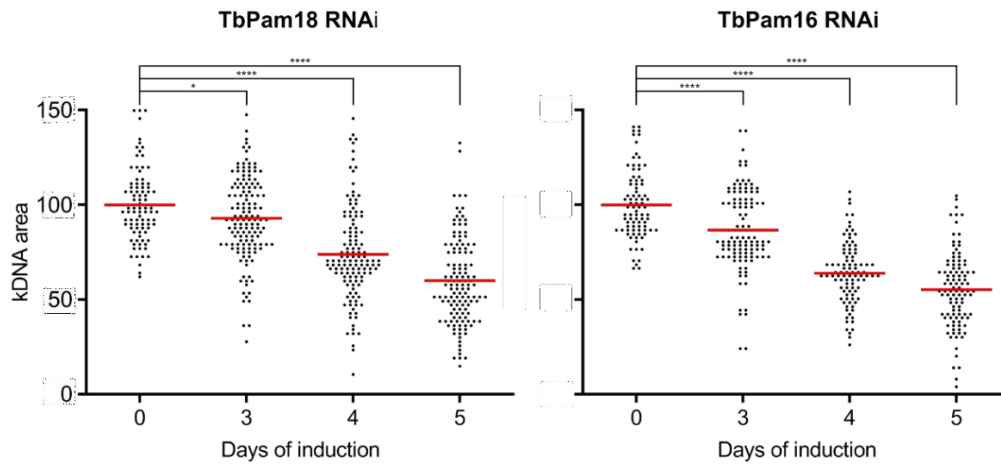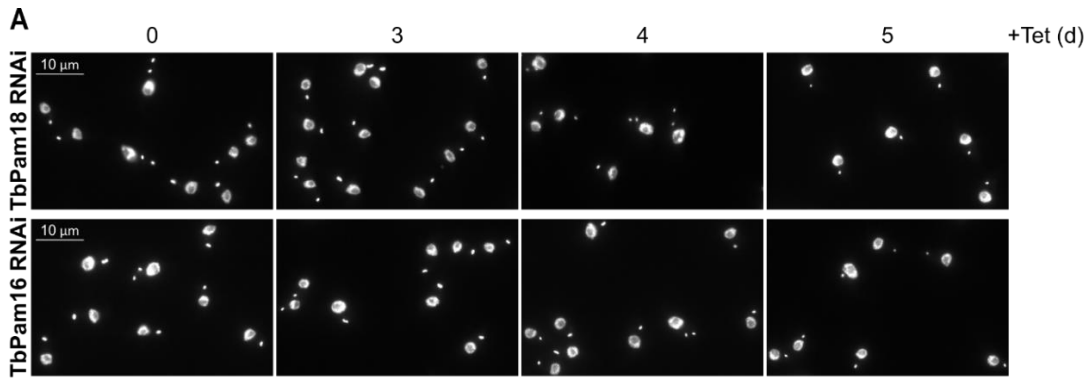
TbPam16 RNAi

Figure 3 - TbPam18 and TbPam16 ablation causes loss of maxicircles: (A) Upper panels: Fluorescence microscopy analysis of DAPI-stained uninduced and three to five days RNAi-induced TbPam18 and TbPam16 RNAi cells as indicated. Lower panels: Quantification of kDNA areas in 86 to 140 DAPI-stained RNAi cells induced for the indicated amount of time. Red line indicates the mean of the kDNA sizes at each timepoint. The mean of the uninduced cells was set to 100%. *: p- value<0.05, ****: p-value<0.0001, as calculated by an unpaired two-tailed t-test. **(B)** Southern blot analysis of steady-state levels of mini- and maxicircles in the TbPam18 and TbPam16 RNAi cell lines. Upper panels: Total DNA from uninduced or three to five days induced cells was isolated and digested with *Hind*III and *Xba*I. Probes specifically recognizing mini- or maxicircles were used. A probe detecting a 3.6-kb fragment of the tubulin intergenic region serves as a loading control. Lower panels: Densitometric quantification of mini- and maxicircle abundance on Southern blots during TbPam18 or TbPam16 depletion. The ratio of the mini- or maxicircle abundance and the respective loading control (tubulin) was normalized (norm.) to the ratios of uninduced cells. Blue (maxicircles) and red (minicircles) bars represent the mean of three independent biological replicates. n.s.: not significant, **: p-value<0.01, ****: p-value<0.0001, as calculated by an unpaired two-tailed t-test.

To study the effects on the kDNA in more detail, we performed Southern blot analysis of mini- and maxicircles in TbPam18 and TbPam16 RNAi cells (*Fig. 3B*). Total DNA was extracted from uninduced cells and two to five days after RNAi induction. To analyse total amounts of mini- and maxicircles, we utilized specific mini- and maxicircle probes. As a loading control, we used a probe targeting the tubulin intergenic region. Already after three days of induction, when RNAi-induced cells still grow at the same rate as uninduced trypanosomes[40], maxicircle levels are significantly reduced to 36% upon TbPam18 and 42% upon TbPam16 depletion, when compared to levels in uninduced cells. In cells induced for five days, maxicircles become almost undetectable. The levels of minicircles, on the other hand, are not significantly reduced over five days of TbPam18 and TbPam16 RNAi induction. Based on these results, we suggest that TbPam18 and TbPam16 are involved in maxicircle replication and/or maintenance. The fact that the impact on the maxicircles is the earliest observed phenotype and occurs even before the onset of the growth retardation, points towards a direct effect.

Ablation of TbPam18 decrease the activity of respiratory complexes and membrane potential

Next, to determine whether the downregulation of respiratory complexes observed during SILAC experiments would affect their function, the enzymatic activity of complex II (cII), complex III (cIII) and complex IV (cIV) in isolated mitochondria was measured.

As shown in *Fig. 4 A*, a statistically significant downregulation of both cIII and cIV activity was observed for TbPam18 RNAi at day 4 post-induction. A slight upregulation in the activity of cII upon RNAi induction remained below statistical significance, in agreement with the previous results obtained for this element during SILAC-MS analysis. On the contrary, no statistically significant changes in cII, cIII and cIV activities were observed for TbPam16, as shown in *Fig. 4 B*. As it could be concomitantly expected, a pronounced drop in the membrane potential of live cells at day 6 after Tet induction was observed, depicted in *Fig. S2 A,B*.

These results correlate and recapitulate the observations regarding OXPHOS relative abundance during TbPam18 and TbPam16 SILAC-MS showing that the enzymatic activity of respiratory complexes, crucial in the procyclic stage of *T. brucei*, is affected by TbPam18 depletion and this produces a drop in membrane potential as a late effect.
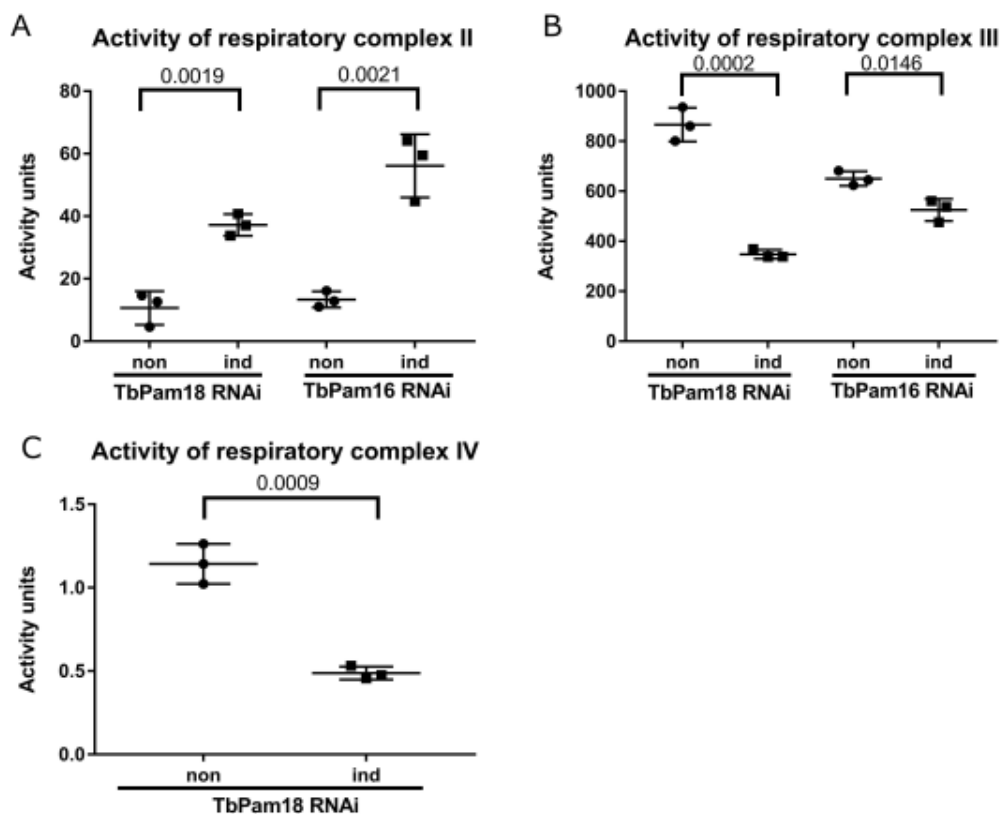
Fig. 4 - Ablation of TbPam18, TbPam16 affects the activities of respiratory complexes. Activities of three respiratory complexes are elevated in induced cells 4 days postinduction. The p-value of unpaired t-test is shown where a statistically significant difference was detected. Means from three independent technical replicates are displayed. Error bars represent standard deviation. One unit (U) of appropriate activity catalyses the reduction/oxidation per minute of 1 nmol of 2,6-dichlorophenolindophen for succinate dehydrogenase (complex II), and 1 µmol of cytochrome c for both cytochrome c reductase (complex III) and cytochrome c oxidase (complex IV). Specific activity is calculated as U per mg of mitochondrial proteins.
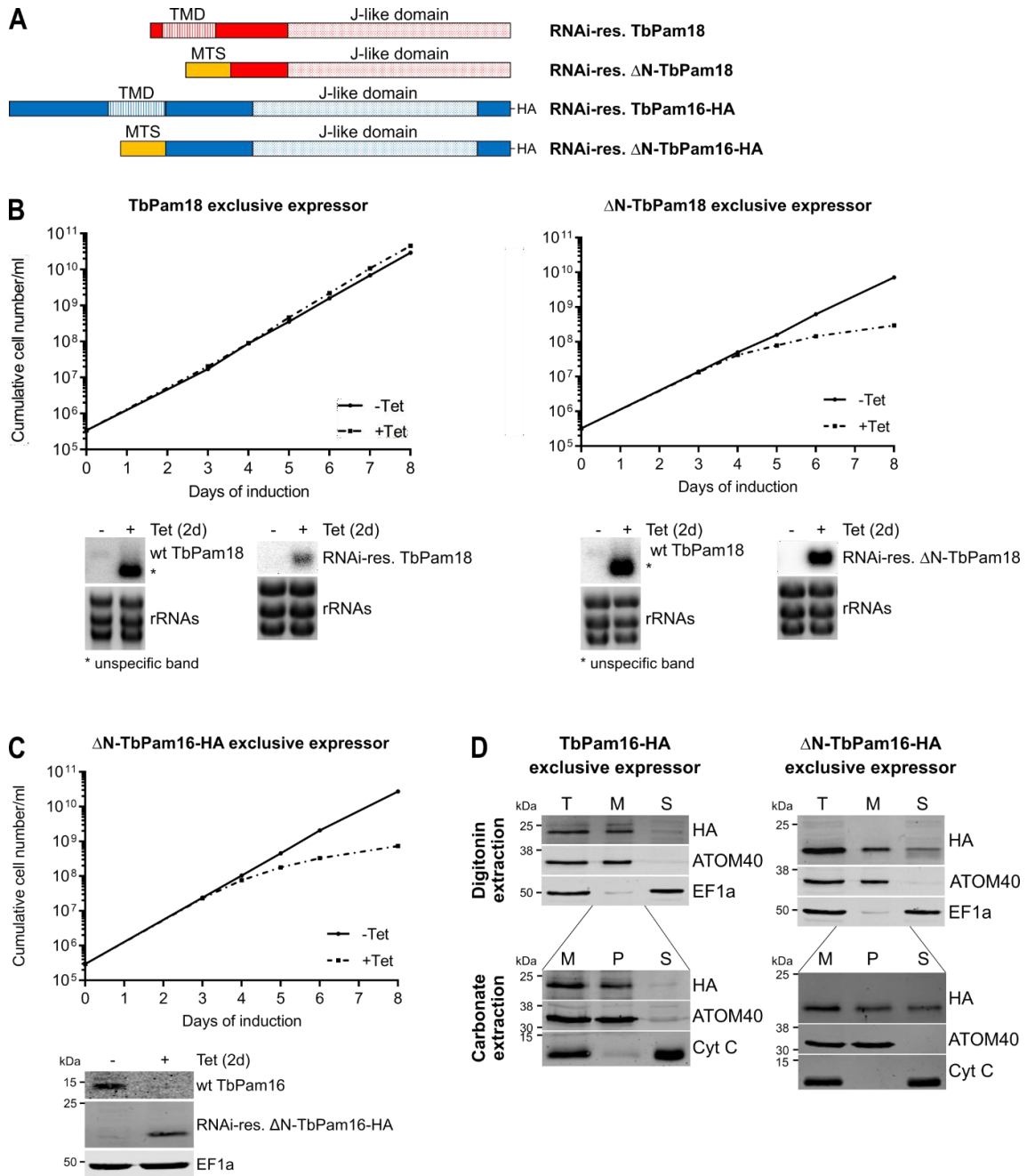
Integral membrane localization is crucial for the function of TbPam18 and TbPam16

In our previous publication, we have demonstrated that TbPam18 and TbPam16 are integral IM proteins[40]. All attempts to further pinpoint this localization within the mitochondrion by fluorescence microscopy, for example to the proximity of the kDNA, were not successful. To the best of our knowledge, there is no protein described yet that is localized in the IM and directly involved in kDNA replication. To answer the question, whether the integral mem-brane localization is essential for their function, we established full-length and truncated TbPam18 and TbPam16 variants lacking their predicted transmembrane domains (TMDs). These variants were all cloned from the synthetic, RNAi-res. TbPam18 or TbPam16 genes introduced above (*Fig. 5A*). The truncated variant of TbPam18 (ΔN-TbPam18) lacks its predicted N-terminal TMD (amino acids (aa) 5-22) and five aa upstream of it, which were replaced by the mitochondrial targeting sequence (MTS) of the trypanosomal mitochondrial heat shock protein 60 (TbmHsp60). This ensures its import into the mitochondrion. The full-length and the truncated RNAi-res. variants of TbPam18 were not tagged since we found that C- as well as N-terminal tags render the protein unfunctional (*Figs. 1C, 5A and S1*). The predicted TMD of TbPam16 (aa 34-52) was replaced by the MTS of TbmHsp60 in the RNAi-res. ΔN-TbPam16 variant. Additionally, the protein was C-terminally HA-tagged, since we showed earlier that the tagged full-length TbPam16 remained functional (*Figs. 1B and 5A*). The described constructs were expressed in the background of RNAi targeting wt TbPam18 and TbPam16, respectively, resulting in the variant being the only version of the respective protein expressed (exclusive expressor cell lines).

The full-length TbPam18 version fully complements the growth phenotype, while expression of ΔN-TbPam18 does not rescue the growth retardation caused by TbPam18 RNAi (*Fig. 5B*). Expression of ΔN-TbPam16-HA, as well, is not capable of rescuing the loss of wt TbPam16 (*Fig. 5C*). To monitor if ΔN-TbPam16-HA is imported into the mitochondrion

as expected, the ΔN-TbPam16-HA exclusive expressor cell line was subjected to a cell fractionation using low concentrations of digitonin. As shown in *Fig. 5D*, full-length TbPam16-HA, as well as ΔN-TbPam16-HA co-fractionate with the mitochondrial marker ATOM40. Furthermore, we used an alkaline carbonate extraction to generate a pellet containing integral membrane proteins and a supernatant fraction containing soluble proteins. Whilethe full-length TbPam16-HA is only recovered in the pellet fraction, ΔN-TbPam16-HA is found in the pellet, as well as in the supernatant fractions at about equal amounts. The finding that not all ΔN-TbPam16-HA is found in the soluble fraction, even though its TMD was deleted,might be due to a close interaction with another membrane protein, likely TbPam18.

Based on these results, we suggest that the membrane integral-localization of TbPam18 and TbPam16 is essential for their function. This finding, for the first time, linksthe process of maxicircle replication to the IM. Interestingly, the domain structures of yeast Pam18 and Pam16 were studied by implementing similar and even more extensive truncations[57]. It was shown that for both proteins, the only essential part is their J or J- like domain, respectively. The findings that TbPam18 and TbPam16 functions depend on the presence of their TMDs, therefore, sets them even further apart from the functionalityof yeast Pam18 and Pam16.

A

TMD | J-like domain — **RNAi-res. TbPam18**

MTS | J-like domain — **RNAi-res. ΔN-TbPam18**

TMD | J-like domain —HA | **RNAi-res. TbPam16-HA**

MTS | J-like domain —HA | **RNAi-res. ΔN-TbPam16-HA**

B

TbPam18 exclusive expressor

ΔN-TbPam18 exclusive expressor

Cumulative cell number/ml vs Days of induction

-Tet / +Tet

Tet (2d): wt TbPam18, * ; RNAi-res. TbPam18; rRNAs
* unspecific band

Tet (2d): wt TbPam18, * ; RNAi-res. ΔN-TbPam18; rRNAs
* unspecific band

C

ΔN-TbPam16-HA exclusive expressor

Cumulative cell number/ml vs Days of induction

-Tet / +Tet

Tet (2d): wt TbPam16; RNAi-res. ΔN-TbPam16-HA; EF1a

D

TbPam16-HA exclusive expressor

ΔN-TbPam16-HA exclusive expressor

Digitonin extraction: T M S — HA, ATOM40, EF1a

Carbonate extraction: M P S — HA, ATOM40, Cyt C

119

Figure 5 - Integral membrane localization is crucial for TbPam18's and TbPam16's functions: (A) Schematic representation of RNAi-resistant (RNAi-res.) full-length (TbPam18 and TbPam16-HA) and N-terminally truncated variants of TbPam18 (ΔN-TbPam18) and TbPam16 (ΔN-TbPam16-HA). TbPam18 constructs are untagged, while TbPam16 constructs carry a C-terminal HA-tag. Predicted transmembrane domains (TMD) and J-like domains are indicated. To ensure a mitochondrial localization, the N-terminally truncated variants were expressed with the mitochondrial targeting sequence (MTS) of the trypanosomal mitochondrial heat shock protein 60 (mHsp60). TMDs were predicted by TMHMM-2.0[58], The J-like domain of TbPam18 was predicted by InterPro[59] and the one of TbPam16 was predicted by alignment with the yeast Pam16. **(B)** Upper panels: Growth curves of uninduced (-Tet) and RNAi-induced (+Tet) cell lines ectopically expressing RNAi-resistant (RNAi-res.) full-length TbPam18 (left) or ΔN-TbPam18 (right) in the background of RNAi targeting the wildtype (wt) TbPam18 (TbPam18 and ΔN-TbPam18 exclusive expressors). Lower panels: Northern blots of total RNA isolated from uninduced (-) and two days induced (+) cells probed for either the mRNA of wt TbPam18 or RNAi-res. ΔN-TbPam18. This was done to confirm efficient RNAi of wt TbPam18 and simultaneous inducible ectopic expression of RNAi-res. TbPam18 or ΔN-TbPam18. Asterisk (*) indicates a prominent RNA product that results from the RNAi against TbPam18. Ethidium bromide-stained rRNAs serve as loading control. **(C)** Upper panel: Growth curve of uninduced (-Tet) and RNAi-induced (+Tet) cells ectopically expressing RNAi-res. ΔN-TbPam16- HA in the background of RNAi targeting the wt TbPam16 (ΔN-TbPam16-HA exclusive expressor). Lower panel: Immunoblot analysis of whole-cell extracts of uninduced (-) and two days induced (+) cells probed for wt TbPam16, RNAi-res. ΔN-TbPam16-HA and EF1a serve as loading controls. **(D)** Upper panels: Immunoblot analysis of total cells (T), digitonin-extracted, mitochondria-enriched (M) and soluble cytosolic (S) fractions of TbPam16-HA and ΔN-TbPam16-HA exclusive expressor cell lines. Immunoblots were probed with anti-HA antibodies and antisera against ATOM40 and EF1a, which serve as mitochondrial and cytosolic markers, respectively. Lower panels: Digitonin-extracted crude mitochondrial fractions (M) were subjected to an alkaline carbonate extraction resulting in a pellet enriched in integral membrane proteins (P) and a soluble supernatant fraction (S). Immunoblots were probed with anti-HA and antisera against ATOM40 and cytochrome C (Cyt C), which serve as marker for integral membrane and soluble proteins, respectively.

The J-domain of yeast Pam18 cannot complement the loss of the TbPam18 J-like domain.

Our findings that TbPam18's TMD is essential sets it apart from the yeast Pam18 (ScPam18), whose only essential domain is its J-domain[57]. Similar to ScPam18, the J-domain of TbPam18 is located at the C-terminus of the protein (aa 47-120) (*Fig. 6A*). However, within the J-domains of ScPam18 and TbPam18, a striking difference can be found. A defining feature of J-domains is the conserved tripeptide His-Pro-Asp (HPD), which is essential for the stimulation of the ATPase activity of their Hsp70 partners[60,61]. While the J-domain of ScPam18 contains

an intact HPD motif, in the TbPam18 J-domain, the motif has changed to His-Ser-Asp (HSD), making it to a J-like domain[61]. The HPD motif is well conserved in Pam18 homologues across eukaryotes (*Fig. S2A*). Among the 13 Pam18 homologues from representative eukaryotes we analysed, TbPam18 is the only one with an aberrant HPD motif. In other trypanosomatid species, however, the HSD motif, as found in TbPam18, is a conserved feature (*Fig. S2B*).
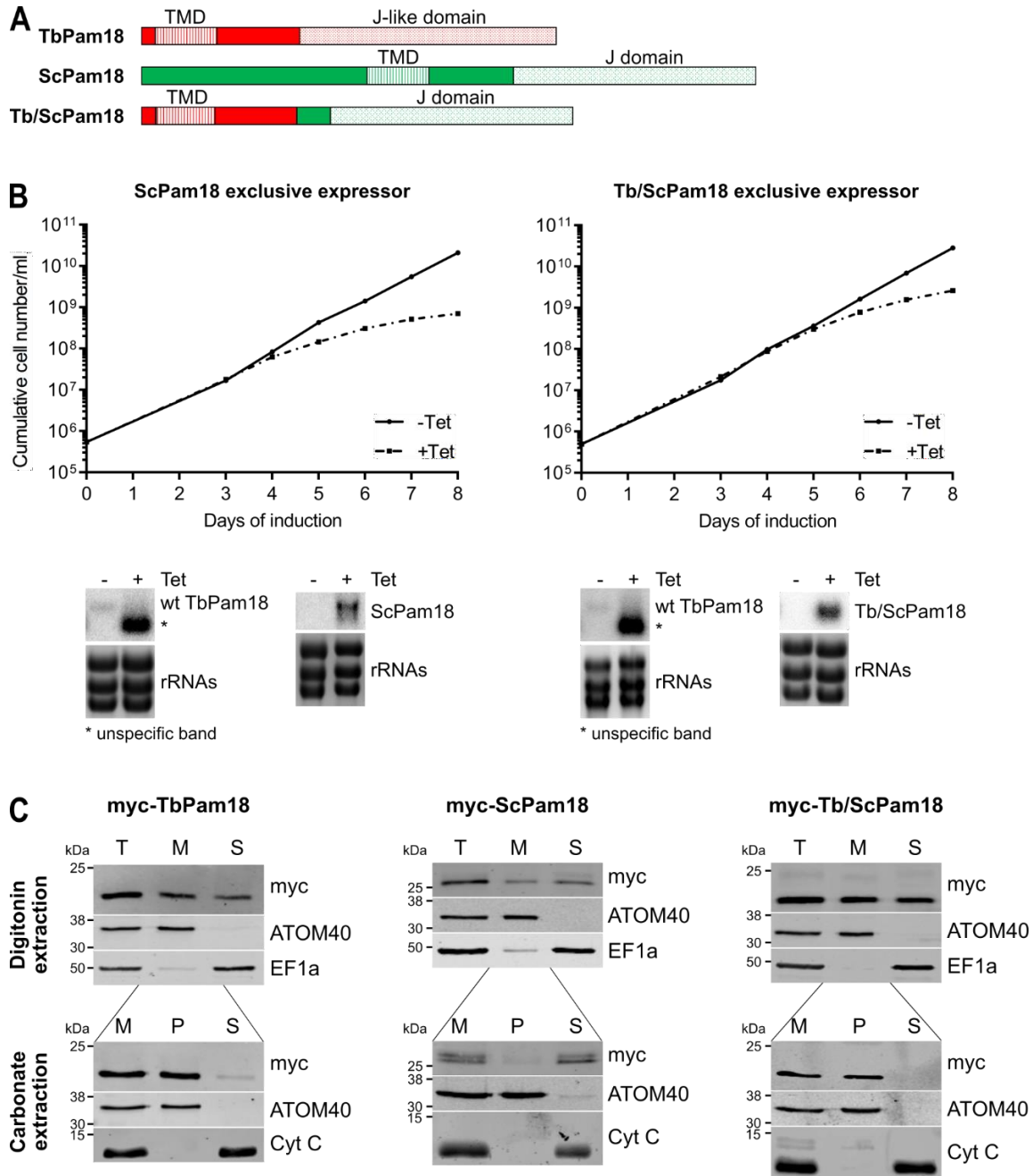
Figure 6 - The J-domain of ScPam18 cannot complement the loss of the J-like domain of TbPam18: **(A)** Schematic representation of TbPam18, yeast (Sc) Pam18 and the Tb/Sc fusion Pam18 (Tb/ScPam18), in which the J-like domain of TbPam18 was replaced by the J-domain of ScPam18. The J-domain of TbPam18 was predicted by InterPro[59]. The J-domain of ScPam18 was previously described[57]. **(B)** Upper panels: Growth curves of uninduced (-Tet) and induced (+Tet) cells ectopically expressing

ScPam18 (left) or Tb/ScPam18 (right) in the background of RNAi targeting the endogenous wildtype (wt) TbPam18 (ScPam18 and Tb/ScPam18 exclusive expressor cell lines). Lower panels: Northern blots of total RNA extracted from uninduced (-) and two days induced (+) cells, probed for wildtype (wt) TbPam18, to confirm efficient RNAi, or ScPam18 as well as Tb/ScPam18, to confirm inducible ectopic expression. Asterisk (*) indicate a prominent RNA product that results from the RNAi against TbPam18. Ethidium bromide-stained rRNAs serve as loading control. **(C)** Upper panels: Immunoblot analysis of total cells (T), digitonin-extracted mitochondria-enriched (M), and soluble cytosolic (S) fractions of cell lines expressing N-terminally myc-tagged TbPam18, ScPam18 or Tb/ScTbPam18. Blots were probed with anti-myc antibodies and antisera against ATOM40 and EF1a, which serve as mitochondrial and cytosolic markers, respectively. Lower panels: Digitonin-extracted crude mitochondrial fractions (M) were subjected to an alkaline carbonate extraction resulting in a pellet enriched in integral membrane proteins (P) and a soluble supernatant fraction (S). Immunoblots were probed with anti-myc and antisera against ATOM40 and Cyt C, which serve as maker for integral membrane and soluble proteins, respectively.

We asked the question if the domain-specific changes in TbPam18 are prerequisites for its function in maxicircle replication, or whether this function could be taken over by a more conventional Pam18 homologue. To pursue this question, we generated a cell line allowing the inducible ectopic expression of ScPam18 in TbPam18 RNAi background (ScPam18 exclusive expressor). As shown in *Fig. 6B*, expression of ScPam18 cannot complement the growth phenotype caused by TbPam18 ablation. To find out more about the fate of ScPam18, we generated an inducible cell line expressing a N-terminally myc-tagged ScPam18 version and subjected these cells to digitonin and alkaline carbonate extractions. *Fig. 6C* shows that myc-ScPam18 does only partially localize to the mitochondrial fraction in a digitonin extraction. However, the same pattern can be observed for a N-terminally myc-tagged version of TbPam18. This observation can possibly be explained by the overexpression of the proteins, or it is also conceivable that the myc-tag at the N-terminus partially impedes recognition of the mitochondrial import signals. However, while the mitochondrially located myc-TbPam18 completely localizes to the integral membrane-fraction in an alkaline extraction, as expected, myc-ScPam18 remains completely soluble. In future experiments, usage of untagged ScPam18 in a combination with a ScPam18 antibody could help to clarify if this mislocalization is due to the tag or if ScPam18 indeed cannot be incorporated into the trypanosomal IM.

To increase the likelihood for correct localization and to pursue the question, whether an intact J-domain can take over the function of TbPam18's J-like domain, we replaced the J-like domain of TbPam18 by the J-domain of ScPam18 (*Fig. 6A*, Tb/ScPam18). This Tb/ScPam18 fusion protein was then expressed in TbPam18 RNAi background (Tb/ScPam18 exclusive

expressor). Expression of Tb/ScPam18 delayed the onset of the growth phenotype by one day, but could not rescue the growth retardation at later time points (*Fig. 6B*). Digitonin extractions of a cell line expressing C-terminally myc-tagged Tb/ScPam18 revealed that myc-Tb/ScPam18 as well is only partially found in the mitochondria-enriched fraction (*Fig. 6C*). However, in an alkaline carbonate extraction it behaves the same as myc-TbPam18 and completely localizes to the integral membrane fraction.

The findings that neither the wildtype ScPam18, nor its J-domain linked to TbPam18's TMD, can rescue the loss of TbPam18 or its J-like domain, respectively, suggest that TbPam18, whilst still sharing homology with ScPam18[40], has evolved different features that make TbPam18 and ScPam18 not interchangeable. The seemingly subtle change from a HPD to an HSD motif in TbPam18's J-domain suggests that it can no longer directly stimulate Hsp70's ATPase activity and, thus, most likely adopted a substantially different functionality.

TbPam18 and TbPam16 have a procyclic form-specific function.

Trypanosomes have a complex life cycle alternating between an insect vector, the tsetse fly (*Glossina* spp.), and a mammalian host. In the insect vector, trypanosomes occur in the procyclic form (PCF). PCF parasites contain an extensively reticulated cristae-rich mitochondrion that is capable of oxidative phosphorylation. The replicative form of trypanosomes in the mammalian host is termed bloodstream form (BSF). BSF trypanosomes have a tube-shaped, cristae-poor mitochondrion with a highly reduced mitochondrial metabolism[1,42]. The mitochondrion in BSF parasites lacks the respiratory complexes, except for the ATP-synthase. The ATP-synthase, however, functions in reverse and, at the expense of ATP, pumps protons from the matrix into the intermembrane space, in order to maintain the proton gradient over the IM[62,63]. Subunit a of the membrane integral F_o subunit, is critical for this proton translocation. The fact that subunit a is encoded in the kDNA, and that its pre-mRNA requires substantial RNA editing, explains why BSF, just as PCF trypanosomes, depend on an intact kDNA network[64,65]. In our previous studies, we have demonstrated that TbPam18 and TbPam16 are not essential for normal growth in the BSF strain γL262P[40]. γL262P BSF cells harbor a single point mutation in subunit γ of the F_1 ATP-synthase, which can compensate for the loss of subunit a and, therefore, for the total loss of kDNA[65]. Since we have now demonstrated that TbPam18 and TbPam16 are involved in kDNA replication (*Fig. 3*), it does no longer seem surprising that their knockdown does not affect the growth of γL262P cells.
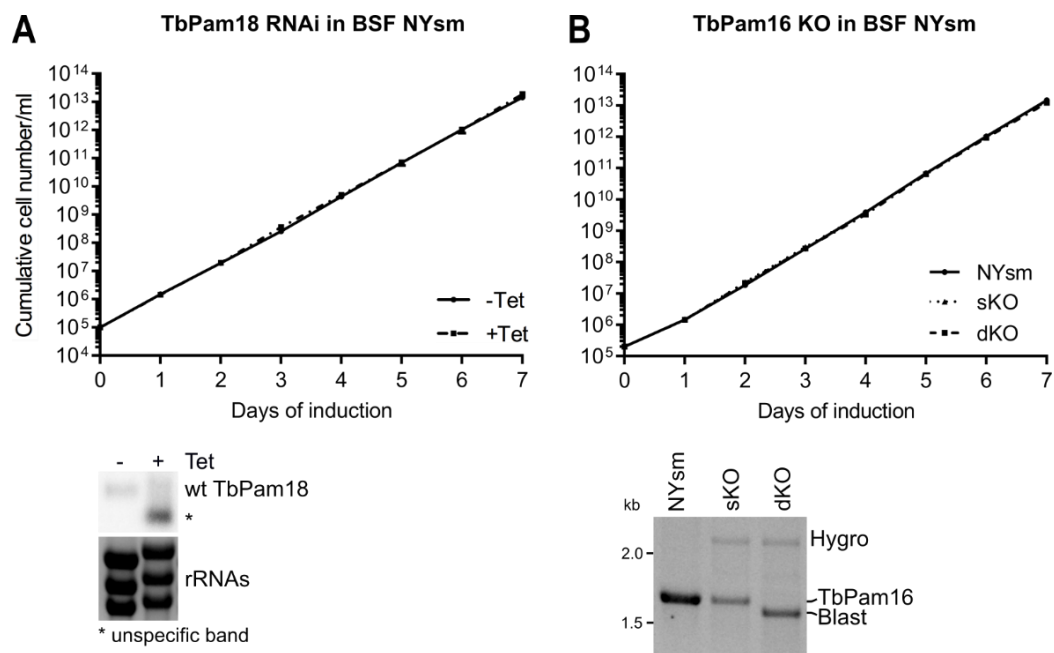
Figure 7 - TbPam18 and TbPam16 are not essential in BSF trypanosomes: (A) Upper panel: Growth curve of uninduced (-Tet) and RNAi-induced (+Tet) bloodstream form (BSF) New York single marker (NYsm) RNAi cell line ablating TbPam18. Error bars correspond to the standard deviation (n=3). Lower panel: Northern blot of total RNA extracted from uninduced (-) and two days induced(+) cells, probed for the TbPam18 ORF. Asterisk (*) indicates a prominent RNA product that resultsfrom the RNAi against TbPam18. Ethidium bromide-stained rRNAs serve as loading control. **(B)** Upper panel: Growth curve of NYsm, TbPam16 single knockout (sKO) and double knockout (dKO)BSF cell lines. Lower panel: Verification of sKO and dKO by PCR using one primer pair to amplify the TbPam16 ORF (~1.7 kilobases (kb)), the hygromycin (Hygro, ~2.3 kb) or blasticidin (Blast, ~1.6kb) resistance cassettes at the same time. Hygro was used to replace the first allele and Blast wasused to replace the second allele.

To analyze the effect of their ablation in the BSF flagellates that depend on an intact kDNA network, we depleted TbPam18 and TbPam16 in the BSF New York single marker (NYsm) strain[66]. As RNAi never eliminates all mRNAs, we aimed to establish TbPam18 and TbPam16 double knockout (dKO) cell lines. While attempts to produce a TbPam18 dKO cell line failed, TbPam18 could be knocked down via RNAi (*Fig. 7A*). For TbPam16 a double knockout (dKO) cell line was successfully generated (*Fig. 7B*). As shownin *Fig. 6*, neither TbPam18 nor TbPam16 are essential for normal growth of NYsm BSF cells. These results demonstrate that the functions of TbPam18 and TbPam16 are stage-specific for PCF trypanosomes. This suggests that their function in maxicircle replication might be redundant in the BSF cells or, less likely, taken over by other proteins.

Discussion

TbPam18 and TbPam16 are *bona fide* orthologues of the PAM subunits Pam18 and Pam16 in other eukaryotes. Unexpectedly, in *T. brucei* they are neither involved in mitochondrial protein import, nor are they part of the PAM[40]. Here, we present data showing that TbPam18 and TbPam16 are essential for proper replication and/or maintenance of the kDNA maxicircles in the procyclic form of the parasite (*Fig. 3*). Maxicircles, similar to mitochondrial genomes in other eukaryotes, code for the two mitochondrial rRNAs, two MtRPs and several subunits of the OXPHOS complexes I, III, IV and V[1,3,51]. Thus, it is not surprising that depletion of TbPam18 and TbPam16 leads to a downregulation of steady-state levels of MRPs and components of OXPHOS complexes I, III, IV and V (*Fig. 2B*). Interestingly, ablation of TbPam18 and TbPam16 also leads to upregulated steady-state levels of DNA topoisomerase IB subunits and UMSBP1. Topoisomerase IB, which in trypanosomes consists of two subunits, is associated with the nuclear and mitochondrial genome replication, yet its exact role remains unclear[2,52,53]. UMSBP1 is required for the initiation of minicircle replication, but is also involved in the regulation of kDNA condensation and segregation, and possibly also in nuclear division[2,54,67]. It is tempting to speculate that these upregulations are the result of a stress reaction that tries to compensate for the loss of maxicircles upon TbPam18 and TbPam16 ablation (*Fig. 2B*). However, levels of other proteins involved in kDNA replication, maintenance, and segregation remain widely unchanged upon TbPam18 or TbPam16 ablation. This is in line with the naturally occurring *T. brucei* ecotypes *T. b.* f. *equiperdum* and *T. b.* f. *evansi* that have either partially [dyskinetoplastic (dk)] or completely [akinetoplastic (ak)] lost their kDNA. Nevertheless, proteins involved in kDNA replication and maintenance are still imported into their mitochondria[68]. Thus, it seems that the levels of most proteins involved in kDNA replication or maintenance are independent of the presence or absence of the kDNA.

Interestingly, minicircle levels remain stable after the loss of maxicircles upon TbPam18 and TbPam16 depletion (*Fig. 3B*). Similar observations have been reported, upon depletion of the mitochondrial DNA primase 1 (TbPRI1)[55], the mitochondrial DNA helicase TbPIF2[56] and the mitochondrial heat shock protein machinery TbmHsp70/TbmHsp40[44], all of which are linked to maxicircle replication and maintenance. Additionally, in the cases of TbPRI1[55] and TbmHsp70/TbmHsp40[44] a shrinkage of kDNA disks was observed *via* DAPI-staining, similar to what we observe after TbPam18 and TbPam16 ablation (*Fig. 3A*). A recent study investigating kDNA condensation steps, suggested that the linkage between minicircles and maxicircles stabilizes the network structure[69]. A destabilized network upon loss of maxicircles could lead to a reduced intensity of the DAPI signal. However, even though

minicircles are much smaller in size than maxicircles, they are much more abundant and represent the bulk of the kDNA network[1,7].Thus, it seems unlikely that a shrinkage of the network could be observed, if only maxicircles are lost. Indeed, electron microscopy analysis of dk trypanosomes of the *T. b.* f. *equiperdum* ecotype that lost all maxicircles, but retained minicircles, shows that the size and ultrastructure of the kDNA disk is indistinguishable from its wild-type form[68]. Therefore, the shrinkage of the kDNA observed upon ablation of maxicircle replication factors mentioned above must be mainly a consequence of the depletion of minicircles from the kDNA disk. During replication, minicircles are detached from the kDNA network[2]. Upon TbPRI1 and TbmHsp70/TbmHsp40 depletion, free minicircle species that could still be replicated, butnot reattached to the network, were shown to accumulate[44,55]. This observation could explain a progressive shrinkage of the kDNA disk as well as why the total number of minicircles, as measured in Southern blots, remains steady for several days after the loss of maxicircles. These findings suggest that there are no control mechanisms in place to stop minicircle replication if replication of maxicircle fails. Thus, it could be that such an accumulation offree minicircles also occurs upon TbPam18 and TbPam16 ablation.

In what way could TbPam18 and TbPam16 be involved in maxicircle replication or maintenance? TbPam18 and TbPam16 both contain a single TMD. Furthermore, they were previously shown to be localized in the IM[40]. Using N-terminally truncated variants of TbPam18 and TbPam16 that lack their TMDs, we present data suggesting that the membrane-integral localization of TbPam18 and TbPam16 is essential for their function (*Fig. 6*). To our knowledge, no other experimentally studied kDNA replication factor is an integral membrane protein. Factors involved in kDNA replication are typically locateddirectly at the kDNA disk, albeit in specific sub-locations[2]. Thus, our finding links maxicircle replication to the IM for the first time. The only other known IM protein that is associated with kDNA inheritance is p166, a subunit of the TAC[70,71]. The TAC is essential for kDNA segregation and knockdownof TAC subunits leads to a complete loss of the kDNA and cells with small or enlarged, over-replicated kDNAs[18]. Since knockdown of TbPam18 and TbPam16 leads to a progressive shrinkage of the kDNA, it is unlikely that they are involved in kDNA segregation and the TAC function.

The finding that the TMDs of TbPam18 and TbPam16 are essential for their functions, differs from the situation in yeast. The TMD of ScPam18 is not essential and the yeast Pam16 homologue is not an integral membrane protein[57]. Another difference can be found in the J-domains of TbPam18 and ScPam18. In TbPam18 the conserved HPD tripeptide was changed to HSD. J-domain proteins are obligate co-chaperones of their Hsp70 partners. Via their J-domain they stimulate Hsp70's ATPase activity and stabilize its inter-action with substrate polypeptides.

The HPD motif within the J-domain is essential for the stimulation of ATP hydrolysis by Hsp70[60,72]. The differences in their J-domains and TMDs underline the functional divergence of TbPam18 and TbPam16 from their yeast counterparts. This is furthermore emphasized by the finding that the J-domain of yeast Pam18, which notably contains an intact HPD motif, cannot complement for the loss of the J-like domain of TbPam18 (*Fig. 6*). The HSD in place of the HPD tripeptide is a conserved feature among trypanosomatid Pam18 homologues (*Fig. S3B*). These findings suggest that the ac-quired mutation in the HPD motif is essential for TbPam18's new function outside of the import motor. In future experiments, it will be interesting to see, how TbPam18's function is impacted, if the HSD in its J-domain is changed back to a HPD motif.

The finding that TbPam18 lacks an intact HPD tripeptide in its J-domain suggests that it does not directly interact with Hsp70. In recent years, several reports have collected evidence that there are numerous J-domain proteins, whose chaperone activity does not solely, or not at all, depend on an intact J-domain[73–75]. Many such examples were found in the plant *Arabidopsis thaliana,* which contains 21 J-like proteins involved in various pro-cesses[75]. A few of those J-like proteins in *A. thaliana* were described to indirectly regulate the activity of the Hsp70 system. They do this through interactions with specific J-domain proteins containing a conventional J-domain, much like the regulation of Pam18 by Pam16 in yeast and *A. thaliana*. However, numerous *A. thaliana* J-like proteins also seem to function independently of Hsp70 altogether[74,75]. Would it, thus, be possible that TbPam18 and TbPam16 could also indirectly regulate the Hsp70 machinery, or even function as chaperones on their own? A recent study reported that the *T. brucei* genome encodes for 67 different putative J-domain proteins[76]. Based on different proteomics studies[20,77–80], as well as on online prediction programs, the same report assigned 38 of these putative J-domain proteins to mitochondria. Three of the putative mitochondrial J-domain proteins, including TbPam18, contain abrogated HPD motifs[76]. However, TbPam16, which lacks the HPD motif altogether, was not among the identified 67 putative J-domain proteins. This suggests the presence of even more J-like proteins in *T. brucei* that are, due to their atypical J-like domains, more difficult to recognize. It is unknown why trypanosomes contain such an immense number of mitochondrial J-domain proteins, even though they have a single mitochondrial Hsp70 isoform (TbmHsp70) only[81]. Could it be that the many mitochondrial J-domain proteins are used to differentially regulate the TbmHsp70 machinery? Interestingly, TbmHsp70 and the J-domain protein TbmHsp40 are involved in the replication of maxicircles, but what their exact function in the process is, is not known[44]. The complexity of kDNA replication could suggest that different members of the J-domain protein family are required to regulate TbmHsp70 during the different stages of the process. TbPam18 and TbPam16, besides their TMDs and J-like domains, do not

contain additional domains, which would suggest that they can act as chaperones on their own. It is possible that their membrane-anchored J-like domains regulate J-domains of other proteins and retain them close to the membrane. Membrane-retention potentially could serve as a way for spatial separation until these, yet elusive, J-domain proteins are needed in the complex process of kDNA replication.

A recent survey[76] compared J-domain protein expression levels between different *T. brucei* life cycle stages using several proteomic data sets[82–84]. Interestingly, the survey found that the levels of many J-domain proteins are different between BSF and PCF trypanosomes, suggesting lifecycle specific roles[76]. We present data that TbPam18 and TbPam16 are essential for normal growth in PCF, but not BSF trypanosomes. If TbPam18 and TbPam16 interact with life cycle dependently expressed J-domain proteins, this observation can be easily explained. It will be an interesting topic for future research to identify transient interaction partners of TbPam18 and TbPam16, which are possibly other J- domain proteins, and to determine if they play a role in kDNA replication.

The function of trypanosomal Pam18 in the PAM and, thus, in protein import, which is essential for PCF and BSF trypanosomes, was replaced by TbPam27[40]. Our current hypothesis regarding this replacement implies that in ancient pro-kinetoplastids, TIM22 and TIM23 complexes including a canonical PAM were present[40,85]. The recruitment of TbPam27 to the TIM22 complex enabled TIM22 to import presequence-containing proteins and, thus, to take over the function of the TIM23 complex and its associated PAM[40,85]. Here, we present data demonstrating that instead of being lost, TbPam18 and TbPam16 underwent architectural and sequence-specific alterations. Interestingly, the two proteins are still interaction partners, and their stability is interdependent [40] (*Figs. 1 and 2A*). However, they are involved in a completely unrelated function, the maxicircle replication or maintenance. Therefore, we link this function, for the first time, to integral IM proteins. TbPam18's and TbPam16's function is PCF-specific, possibly because their J-like domains are used to regulate other J-domain proteins that they themselves are expressed in a life cycle stage-specific manner. Our findings once again underline that the presence of an orthologue does not proof that its function is also conserved. While we can only speculate at this point, what the exact function of TbPam18 and TbPam16 might be, we cannot help but being fascinated by the intriguing ways evolution takes.

Materials and methods

Transgenic cell lines

Transgenic *T. brucei* cell lines are either based on the procyclic form (PCF) strain 29-13 or the blood stream form (BSF) strain New York single marker (NYsm)[66]. PCF cells were grown in SDM-79[86] supplemented with 10% (v/v) fetal calf serum (FCS) at 27°C. BSF cells were cultivated in HMI-9[87] containing 10% (v/v) FCS at 37°C.

RNAi against TbPam18 (Tb927.8.6310) and TbPam16 (Tb927.9.13530) has been described previously[40]. For complementation experiments with TbPam18 and TbPam16, synthetic genes (Biomatik) were used. The codons in ORF regions that are targeted by RNAi were changed such that is transcripts are RNAi resistant but still translate into the same amino acid sequence as in the endogenous protein. To produce constructs allowing expression of the N-terminally truncated TbPam18 (ΔN-TbPam18) and TbPam16 (ΔN- TbPam18) variants, the corresponding DNA fragments were amplified from the synthetic genes. To ensure targeting to mitochondria, the MTS of mitochondrial Hsp60 (TbmHsp60,Tb927.10.6510) was cloned in front of the truncated constructs. For the Tb/ScPam18 fusion protein, another synthetic gene (Biomatik) was used. The first 138 nucleotides of the RNAi-resistant TbPam18 were fused in front of the last 213 nucleotides of the wildtype yeast Pam18 gene (YLR008C). Sequences of the synthetic genes are shown in *Fig S4*.

To generate plasmids for ectopic expression of untagged, N- or C-terminal triple c- myc- or HA-tagged RNAi-res. full-length or N-terminally truncated TbPam18 or TbPam16, as well as wildtype ScPam18 and Tb/ScPam18, the complete or truncated ORFs of the respective genes were amplified by PCR. The PCR products subsequently were cloned into a modified pLew100 vector[66,88], which contains a puromycin resistance gene and either no epitope tag or a triple c- myc- or HA-tag[89].

The TbPam16 double knockout (dKO) cell line was generated by fusing the 500 nucleotides up- and downstream of the TbPam16 alleles to the N- or C-terminus of the hygromycin (Hygro) or blasticidin (Blast) resistance cassette, respectively. The first TbPam16 allele in the BSF strain NYsm was replaced by Hygro resulting in the single KO (sKO). To generate the dKO, the second TbPam16 allele was replaced by Blast.

Antibodies

Polyclonal rabbit antiserum against TbPam16 was commercially produced (Eurogentec, Belgium) using aa 153-167 (VKDSHGNSRGNDAMW) as antigen. For western blots (WB) the TbPam16 antiserum was used at a 1: 500 dilution. Commercially available antibodies were: Mouse anti-c-myc (Invitrogen, dilution WB 1: 2,000) mouse anti-HA (Enzo Life Sciences AG, dilution WB 1: 5,000) and mouse anti-EF1a (Merck Millipore, dilution WB 1:10'000). Polyclonal rabbit anti-ATOM40 (dilution WB 1: 10,000) and polyclonal rabbit anti-Cyt C (dilution WB 1: 100) were previously produced in our laboratory[80,90]. Secondary antibodies used: goat anti-mouse IRDye 680LT conjugated (LI-COR Biosciences, dilution WB 1: 20,000) and goat anti-rabbit IRDye 800CW conjugated (LI-COR Biosciences, dilution WB 1: 20,000).

Digitonin extraction

Cell lines were induced with tetracycline for one day prior to the experiment to ensure expression of epitope-tagged proteins. To selectively solubilize the plasma membrane, 1×10^8 cells were incubated at 4°C for 10 min in a buffer containing 0.6 M sorbitol, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA (pH 8.0) and 0.015% (w/v) digitonin. A mitochondria-enriched pellet was separated from a supernatant that is enriched in cytosolic proteins by centrifugation (6'800 g, 5 min, 4°C). Equivalents of 2×10^6 cells of each fraction were analysed by SDS-PAGE and western blotting.

Alkaline carbonate extraction

A mitochondria-enriched pellet, generated as described above, was resuspended in 100 mM Na_2CO (pH 11.5) and incubated at 4°C for 10 min. Centrifugation (100'000 g, 10 min, 4°C) yielded in a pellet enriched in integral membrane proteins and a supernatant enriched in soluble or loosely membrane-associated proteins. Equivalents of 2×10^6 cells of each fraction were analyzed by SDS-PAGE and western blotting.

Co-immunoprecipitation (CoIP)

A mitochondria-enriched digitonin pellet from 1×10^8 cells expressing TbPam16-HA was solubilized in a buffer containing 20 mM Tris-HCl (pH 7.4), 0.1 mM EDTA, 100 mM NaCl, 10% glycerol, 1X Protease Inhibitor mix (Roche, EDTA-free) and 1% (w/v) digitonin for 15 min at

4°C. After centrifugation (20'000 g, 15 min, 4°C), the lysate was transferred to 50 μl HA bead slurry (anti-HA affinity matrix, Roche), which had been equilibrated in wash buffer (20 mM Tris-HCl (pH 7.4), 0.1 mM EDTA, 1 mM NaCl, 10% glycerol, 0.2% (w/v) digitonin). Subsequent to incubation in an end-over-end shaker for 1 hr at 4°C, the supernatant containing the unbound proteins was removed. After washing the bead slurry three times with wash buffer, the bound proteins were eluted by boiling the resin for 5 min in 2% SDS in 60 mM Tris-HCl (pH 6.8).

Blue native (BN)-PAGE

Mitochondria-enriched digitonin pellets of 1×10^8 cells expressing TbPam18-HA or TbPam16-HA were incubated in a buffer containing 20 mM Tris-HCl (pH 7.4), 50 mM NaCl, 10% glycerol, 0.1 mM EDTA, 1 mM PMSF, 1% (w/v) digitonin for 15 min at 4°C to solubilize mitochondrial membranes. After centrifugation (20'817 g, 15 min, 4°C), the resulting supernatants were separated on a 4-13% gradient gel. To facilitate protein transfer, the gel was incubated in SDS-PAGE running buffer (25 mM Tris, 1 mM EDTA, 190 mM glycine, 0.05% (w/v) SDS) prior to western blotting.

Fluorescence microscopy and kDNA size quantification

TbPam18 and TbPam16 RNAi cells were fixed with 4% paraformaldehyde in PBS, postfixed in cold methanol and mounted using VectaShield containing 4′,6-diamidino-2-phenylin-dole (DAPI) (Vector Laboratories). Images were acquired by a DMI6000B microscope and a DFC360 FX monochrome camera (both Leica Microsystems).

Images were analyzed using ImageJ[91]. The kDNA size analysis was performed on binarized 8-bit format images. The size of particles was measured in arbitrary units (a.u.) and kDNA particles >0.0 a.u. and <0.75 a.u. were included in the analysis. Boomerang shaped, dividing kDNAs and randomly picked up particles were manually removed from the analysis. Significance of these results was calculated by an unpaired two-tailed t-test.

RNA extraction and northern blotting

Acid guanidinium thiocyanate-phenol-chloroform extraction to isolate total RNA from uninduced and induced (two days) RNAi cells was done as described elsewhere[92]. Total RNA was separated on a 1% agarose gel in MOPS buffer containing 0.5% formaldehyde. Northern probes

were generated from gel-purified PCR products corresponding to the RNAi inserts (as described previously[40]) or the overexpressed proteins described above, and radiolabelled by means of the Prime-a-Gene labelling system (Promega).

DNA extraction and Southern blotting

For DNA isolation, 5 x 10^7 cells were washed once in NTE buffer (100 mM NaCl, 10 mM Tris (pH 7.5) and 5 mM EDTA) and then resuspended in NTE buffer containing 0.5% SDS for cell lysis and 0.2 mg/ml RNase A to degrade RNA. After incubation for 1 hr at 37°C, 1 mg/ml proteinase K was added, followed by 2 hr of incubation at 37°C. DNA was isolated by phenol-chloroform extraction and subsequent ethanol precipitation.

For Southern blotting, 5 µg of DNA were digested overnight at 37°C with *Hind*III and *Xba*I. Digested DNA was separated in a 1% agarose gel in 1X TAE buffer. Gel processing and blotting was done as described elsewhere[44,47]. For kDNA detection sequence-specific mini- and maxicircle probes were generated by PCR. The minicircle probe was a 0.1 kb stretch of the conserved minicircle sequence[47]. A 1.4 kb fragment served as the maxicircle probe[44,93]. For normalization, a tubulin probe binding to a 3.6 kb stretch within the intergenic region between α- and β-tubulin, was used[47]. Probes were radiolabelled by means of the Prime-a-Gene labelling system (Promega).

SILAC RNAi experiments

TbPam18 and TbPam16 RNAi cells were washed in PBS and resuspended in SDM-80[94] containing 5.55 mM glucose, 10% dialyzed FCS (BioConcept, Switzerland) and either light ($^{12}C_6$/$^{14}N_\chi$) or heavy (13C6/15Nχ) isotopes of arginine (1.1 mM) and lysine (0.4 mM) (Euroisotope). The cells were grown in SILAC medium for six to ten doubling times, to ensure a complete labelling of all proteins with heavy amino acids. Uninduced and induced (four days) TbPam18 or TbPam16 RNAi cells were mixed in an one to one ratio and digitonin- extracted mitochondria-enriched pellets were generated. TbPam18 and TbPam16 SILAC RNAi experiments were done in three biological replicates including a label-switch and analysed by liquid chromatography-mass spectrometry (LC-MS).

Respiratory complexes activity assessment

The enzymatic activities of respiratory complexes II, III, and IV were determined spectrophotometrically in lysates of mitochondrial enriched fractions as described previously[96]. Briefly, mitochondrial enriched fraction obtained from $5*10^8$ cells by hypotonic lysis were lysed in 0.5 M aminocaproic acid (SIGMA) and 2% [w/v] dodecyl maltoside. Upon incubation, the lysate was spun for 30 min at maximum g at 4 °C and supernatant was used. Protein concentration in samples was determined by Bradford method[97].

For succinate dehydrogenase activity (complex II), five μl of the mitochondrial lysate was added to the 1 ml of reaction buffer (25 mM KPi, pH 7.2; 5 mM $MgCl_2$; 20 mM sodium succinate), mixed and incubated in 30 °C for 10 min. Next, antimycin A, rotenone, KCN and
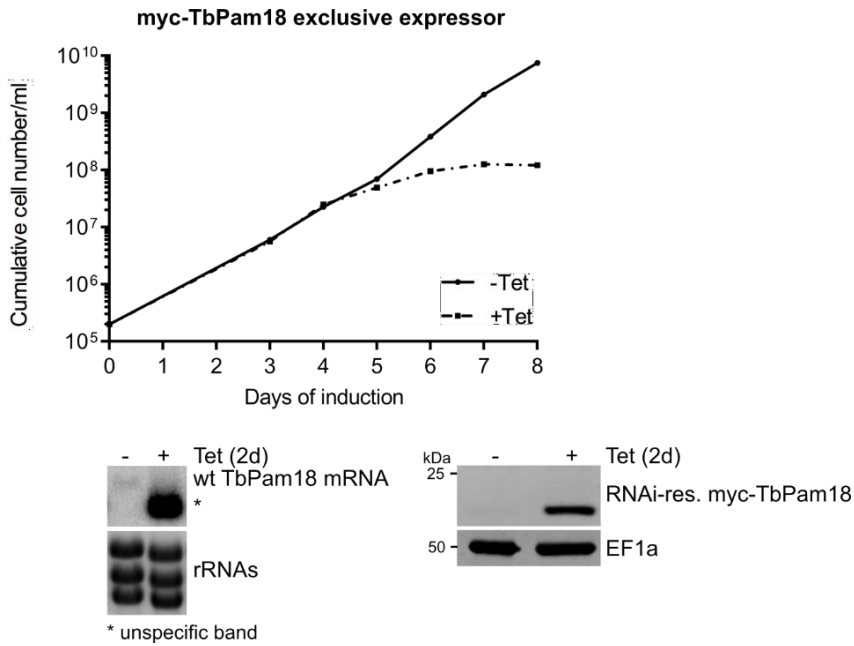
2,6-dichlorophenolindophenol were separately added to a final concentration of 2 μg/ml, 2 μg/ml, 2 mM and 50 μM, respectively. The background reaction was monitored at 600 nm. The reaction itself was started upon the addition of coenzyme Q_2 to a final concentration of 65 μM, and was followed at 600 nm for 5 min.

The cytochrome c reductase activity (complex III) was followed in 1 ml of reaction buffer (40 mM NaPi, pH 7.4; 0.5 mM EDTA, pH 8.5; 20 mM sodium malonate; 50 μM cytochrome c; 0.005% [w/v] dodecyl maltoside). Simultaneously, 2 μl of the mitochondrial lysate and 2 μl of 2,3-dimethoxy-5-methyl-6-dodecyl-1,4-benzoquinol were added and the reaction was monitored at 550 nm for 1 min. The cytochrome c oxidase activity (complex IV) was measured in 1 ml of reaction buffer (40 mM NaPi, pH 7.4; 0.5 mM EDTA, pH 8.5; 20 μM cytochrome c; 30 μM ascorbic acid; 0.005% [w/v] dodecyl maltoside). Ten μl of the mitochondrial lysate was added to the buffer and the reaction was monitored at 550 nm for 10 min.

Membrane potential (Δψm) measurements

The Δψm was determined using the red-fluorescent stain tetramethylrhodamine ethyl ester TMRE (Thermo Fisher Scientific Waltham, MA). Cells in the exponential growth phase were stained with 60 nM of the dye for 30 minutes at 27°C. Cells were pelleted (1.300g, 10 minutes, RT), resuspended in 2 mL of PBS (pH 7.4), and immediately analysed by flow cytometry (BD FACS Canto II Instrument). Treatment with the protonophore FCCP (20 μM) was used as a control for mitochondrial membrane depolarization. For all samples, 10,000 events were collected. Data were evaluated using BD FACS Diva software (BD Biosciences, San Jose, CA).

Supplementary figures

myc-TbPam18 exclusive expressor

(Upper panel: Growth curve — y-axis: Cumulative cell number/ml (10^5 to 10^{10}); x-axis: Days of induction (0–8); legend: -Tet, +Tet)

*(Lower left: Tet (2d), - / +; wt TbPam18 mRNA; *; rRNAs; * unspecific band)*

(Lower right: kDa, Tet (2d), - / +; 25 — RNAi-res. myc-TbPam18; 50 — EF1a)

Figure S1 - myc-TbPam18 is not a functional protein: Upper panel: Growth curve of uninduced (-Tet) and induced (+Tet) cells expressing RNAi-resistant (RNAi-res.) myc-TbPam18 in the back- ground of RNAi against the wild type (wt) TbPam18 (myc-TbPam18 exclusive expressor). Lower panels: Left: Northern blot of total RNA extracted from uninduced (-) and two days induced (+) cells, probed for wt TbPam18. Asterisk (*) indicates a prominent RNA product that results from the RNAi against TbPam18. Ethidium bromide-stained rRNAs serve as loading control. Right: Immunoblot analysis of whole cell-extracts of uninduced (-) and two days induced (+) cells, probed for RNAi-res. myc-TbPam18 and EF1a as loading control.
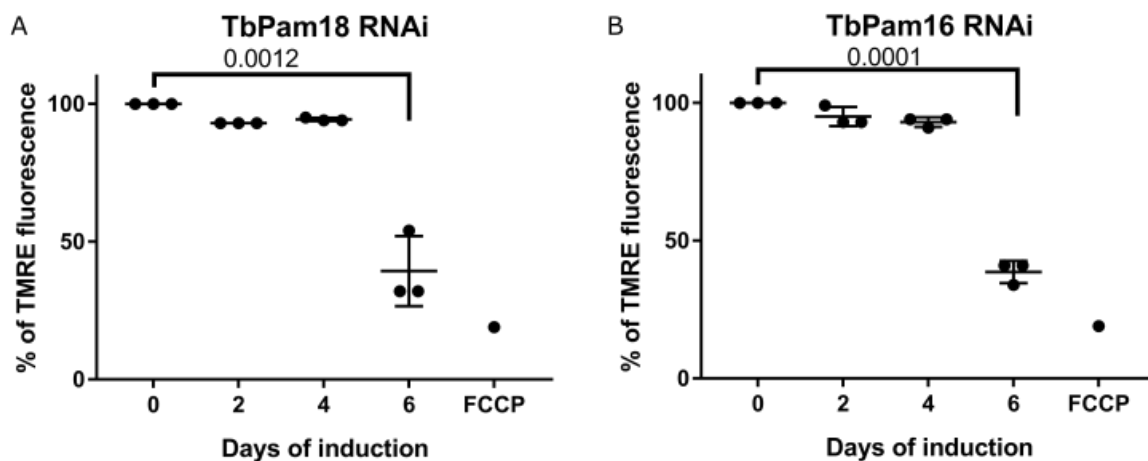
Figure S2 – Ablation of TbPam18, TbPam16 affects the membrane potential (Δψm) as a later effect. A) Flow cytometry measurement of Δψm of TbPam18 at days 0 (uninduced), 2, 4 and 6 after induction. B) Δψm variation measured for TbPam16 at days 0 (uninduced), 2, 4 and 6 after induction. Active mitochondria were stained by positively charged TMRE. A protonophore FCCP serves as a control for membrane depolarization.

Figure S3 - Multiple sequence alignments of Pam18 homologues: (A) Sequence alignment of N-terminal regions of Pam18 homologues of 13 representative eukaryotes. **(B)** Sequence alignment of N-terminal regions of Pam18 homologues of nine representative trypanosomatids. In (A) and (B) His-Pro-Asp (HPD) motifs are highlighted in red and His-Ser-Asp (HSD) motifs in green. Multiple sequence alignments were generated using Clustal Omega[95].

RNAi-res. TbPam18

ATG GCA GCT CCG CTT GCG GCA TTG GTG CTG CTT GGA GGA GCA TAT TAT ATT TTC AGG TTG GCA CCA CGT ATT ACA CAA CGC GTG TCT ATG GCT CAG GGT CTT ACA TGT GCT GCT AAT CGT CAA CTT CGT CCA TAC CGT CGT TAC GAA GGT GGA TTT GAA AAG TCA ATG ACA AAG CGA GAA GCT CTT CTT CTT CTG GGT TTT ACA GAG GAC GTG GCA TCA GGA GGT TTT CTG TCA CTG CCG TCT GAC GAA GAG ATA AAG ACG CGT TAT TAC GGA TTG ATG AAA CAG CTT CAC TCA GAC GTT GAT GGT AGC CCA TAC ATT GCT GCA AAG TTG AAC GAG GCT CGT GAC ATA CTT GGT AAA AAA TAA

RNAi-res. ΔN-TbPam18

ATG TTC CGC TGT GTC GTC CGT TTT GGT GCC AAA GAC ATC CGT TTT GTG TCT ATG GCT CAG GGT CTT ACA TGT GCT GCT AAT CGT CAA CTT CGT CCA TAC CGT CGT TAC GAA GGT GGA TTT GAA AAG TCA ATG ACA AAG CGA GAA GCT CTT CTT CTT CTG GGT TTT ACA GAG GAC GTG GCA TCA GGA GGT TTT CTG TCA CTG CCG TCT GAC GAA GAG ATA AAG ACG CGT TAT TAC GGA TTG ATG AAA CAG CTT CAC TCA GAC GTT GAT GGT AGC CCA TAC ATT GCT GCA AAG TTG AAC GAG GCT CGT GAC ATA CTT GGT AAA AAA TAA

RNAi-res. Tb/ScPam18

ATG GCA GCT CCG CTT GCG GCA TTG GTG CTG CTT GGA GGA GCA TAT TAT ATT TTC AGG TTG GCA CCA CGT ATT ACA CAA CGC GTG TCT ATG GCT CAG GGT CTT ACA TGT GCT GCT AAT CGT CAA CTT CGT CCA TAC CGT TTC TTG AAA GGC GGA TTT GAC CCG AAA ATG AAT TCT AAA GAG GCT CTA CAG ATT TTG AAT TTG ACA GAA AAT ACA TTG ACT AAA AAA AAG TTG AAA GAG GTT CAT AGG AAA ATT ATG TTA GCT AAT CAT CCT GAC AAA GGT GGT TCT CCA TTT TTG GCC ACT AAG ATA AAC GAA GCT AAG GAC TTT TTG GAA AAA AGG GGT ATT AGC AAA TAA

RNAi-res. TbPam16

ATG CGT CGT ATT ATG TCA CCA CGC GTT ATG TGC GAG GTA AAA TTT GGT AGC CGT CCA GCT CCA CTG GCC TGC AGC CGT ATG TTC TTT ATT CCA CCA CAG CTG GCG AAA TTG ATT GTG ACA TCT GGA TTA CTG ATA GTG AAA GCA TTC CTG GTG GCT CAC CAA CAG GAA GCT AAA AGA TTG CGT GAA GAA GAA GAA AAA GAA GGT CAC TCT GCT ACA AAC GCA CAG GTT GGT ACG GGA AGC GCA GCT CTG ATG ACG TCA TCA GAG GCT TTG CAA ATT CTG GGT CTG CAG CCA AAC ATG TCA GTG CCG TTG ACG GCT GAA TCA GAC CGG CAA CTG GCT GCA GTT CGA TTT GAA CAC TTG TTT GCC ATC GCG ACG CGA TGT AAG AAC GTG TTC CTG CAA GGG AAG TTG AGT GGT GCA TAC CGT GTT TGC GTG GAT CCA GAA TGG GAC TTG AAA GAC GAG GTT AAG GAT AGT CAC GGA AAC TCA AGG GGT AAC GAT GCA ATG TGG TAA

RNAi-res. ΔN-TbPam16

ATG TTC CGC TGT GTC GTC CGT TTT GGT GCC AAA GAC ATC CGT TTT CAC CAA CAG GAA GCT AAA AGA TTG CGT GAA GAA GAA GAA AAA GAA GGT CAC TCT GCT ACA AAC GCA CAG GTT GGT ACG GGA AGC GCA GCT CTG ATG ACG TCA TCA GAG GCT TTG CAA ATT CTG GGT CTG CAG CCA AAC ATG TCA GTG CCG TTG ACG GCT GAA TCA GAC CGG CAA CTG GCT GCA GTT CGA TTT GAA CAC TTG TTT GCC ATC GCG ACG CGA TGT AAG AAC GTG TTC CTG CAA GGG AAG TTG AGT GGT GCA TAC CGT GTT TGC GTG GAT CCA GAA TGG GAC TTG AAA GAC GAG GTT AAG GAT AGT CAC GGA AAC TCA AGG GGT AAC GAT GCA ATG TGG TAA

Figure S4 - Synthetic TbPam18 and TbPam16 genes: RNAi-resistant (RNAi-res.) TbPam18 and TbPam16 DNA sequences in black with changed nucleotides highlighted in red. In the ΔN-TbPam18 and ΔN-TbPam16 constructs, the first 81 and 156 nucleotides, respectively, were replaced by the first 45 nucleotides of TbmHsp60, which encode the mitochondrial targeting sequence (MTS) of the protein (green). To generate the Tb/ScPam18 fusion protein, the first 295 nucleotides of RNAi-res. TbPam18 were fused to the last 213 nucleotides of ScPam18 (blue).

Supplementary tables

Table S1 - List of 50 most downregulated mitochondrial proteins in the TbPam18 SILAC RNAi experiment: Mitochondrial ribosomal proteins (MRPs) are assigned to the large subunit (LSU) or small sub- unit (SSU) of the mitoribosome[3]. Proteins of the oxidative phosphorylation (OXPHOS) pathway have been assigned to their respective complexes (I, II, III, IV and V)[42].

	Accession number	Gene names	Fold reduction	≥1.5-fold reduced in TbPam16 RNAi	MRPs (LSU or SSU)	OXPHOS (com-plex)
1	Tb927.6.2470	TIM54	5.44	X		
2	Tb927.1.1390	Hypothetical protein, conserved	3.62	X		
3	Tb927.7.3510	mtLAF21	3.47			
4	Tb927.7.3460	KRIT2	3.42	X	LSU	
5	Tb927.5.3410	MRPL9	3.38	X	LSU	
6	Tb927.10.380	PPR5	3.34	X	LSU	
7	Tb927.3.820	LSU ribosomal protein, mitochondrial, putative	3.34	X	LSU	
8	Tb927.10.6200	Hypothetical protein, conserved	3.32	X		
9	Tb927.6.3930	KRIT1	3.29	X	LSU	
10	Tb927.10.1870	LSU ribosomal protein, mitochondrial, putative	3.18	X	LSU	
11	Tb927.11.10400	mS70	3.17	X	SSU	
12	Tb11.v5.0700	Hypothetical protein, conserved	3.17			
13	Tb927.1.1200	MRPS15	3.10	X	SSU	
14	Tb927.4.1070	MRPL13	2.95	X	LSU	
15	Tb927.3.1410	COXVII	2.94	X		IV
16	Tb927.6.4080	Ribosomal protein, putative	2.92	X	LSU	
17	Tb927.11.10170	MRPL20	2.89	X	LSU	
18	Tb927.5.3980	MRPL15	2.88	X	LSU	
19	Tb927.11.8040	LSU ribosomal protein, mitochondrial, putative	2.87	X	LSU	
20	Tb927.5.3110	MRPL49	2.80	X	LSU	
21	Tb927.11.5530	LSU ribosomal protein, mitochondrial, putative	2.79	X	LSU	
22	Tb927.7.7010	LSU ribosomal protein, mitochondrial, putative	2.76	X	LSU	

23	Tb927.10.7380	LSU ribosomal protein, mitochondrial, putative	2.75			
24	Tb927.11.6000	MRPL4	2.68	X	LSU	
25	Tb927.11.7070	ATP-dependent DEAD/H RNA helicase, putative	2.67			
26	Tb927.7.3030	Hypothetical protein, conserved	2.63	X		
27	Tb927.11.10570	LSU ribosomal protein, mitochondrial, putative	2.58	X	LSU	
28	Tb927.8.5860	MRPL17	2.54	X	LSU	
29	Tb927.3.2370	Hypothetical protein, conserved	2.53	X		
30	Tb927.10.600	MRPL29	2.53	X	LSU	
31	Tb927.6.1700	Hypothetical protein, conserved	2.53		LSU	
32	Tb11.0400	P27	2.50	X		IV
33	Tb927.7.7090	Hypothetical protein, conserved	2.49	X		
34	Tb927.7.2760	MRPL22	2.45	X	LSU	
35	Tb927.5.4120	LSU ribosomal protein, mitochondrial, putative	2.42	X	LSU	
36	Tb927.11.1630	LSU ribosomal protein, mitochondrial, putative	2.39	X	LSU	
37	Tb927.10.8320	COXIX	2.33	X		IV
38	Tb927.9.12850	mtLAF5	2.32	X		
39	Tb927.10.280	COXVI	2.32	X		IV
40	Tb927.11.11630	LSU ribosomal protein, mitochondrial, putative	2.31	X	LSU	
41	Tb927.9.8290	MRPL30	2.31	X	LSU	
42	Tb927.9.9150	mtLAF13	2.28			
43	Tb927.10.13300	MRPS8	2.22	X	SSU	
44	Tb927.11.12040	Cytochrome c oxidase component, putative	2.20	X		IV
45	Tb927.4.1810	MRPL33	2.18	X	LSU	
46	Tb927.1.4480	WD repeat and HMG-box DNA-binding protein, putative	2.15			
47	Tb927.10.13600	Cardiolipin-dependent protein 18kDa	2.11	X		IV
48	Tb927.8.3170	KRIPP9	2.11	X		
49	Tb927.11.4650	MRPL52	2.08	X	LSU	
50	Tb927.6.1410	NADH dehydrogenase1	2.01	X		I

		alpha subcomplex subunit, putative				

Table S2: Lists of 50 most downregulated mitochondrial proteins in the TbPam16 SILAC RNAi experiment. Mitochondrial ribosomal proteins (MRPs) are assigned to the large subunit (LSU) or small sub- unit (SSU) of the mitoribosome3. Proteins of the oxidative phosphorylation (OXPHOS) pathway have been assigned to their respective complexes (I, II, III, IV and V)[42].

	Accession number	Gene names	Fold reduction	≥1.5-fold reduced in TbPam18 RNAi	MRP (LSU or SSU)	OXPHOS (com-plex)
1	Tb927.1.1390	Hypothetical protein, conserved	4.75	X		
2	Tb927.3.2370	Hypothetical protein, conserved	4.18	X		
3	Tb927.6.2470	TIM54	4.17	X		
4	Tb927.3.1410	COXVII	4.16	X		IV
5	Tb927.5.3410	MRPL9	3.99	X	LSU	
6	Tb927.10.380	PPR5	3.96	X	LSU	
7	Tb927.11.9830	LSU ribosomal protein,mitochondrial, putative	3.89	X	LSU	
8	Tb927.1.1200	MRPS15	3.77	X	SSU	
9	Tb11.0400	P27	3.76	X		IV
10	Tb927.11.6000	MRPL4	3.74	X	LSU	
11	Tb927.5.2810	Hypothetical protein, conserved	3.73	X	LSU	
12	Tb927.1.1210	conserved protein, unknown function	3.68	X	LSU	
13	Tb927.7.3460	KRIT2	3.66	X	LSU	
14	Tb927.10.6200	Hypothetical protein, conserved	3.65	X		
15	Tb927.10.11350	LSU ribosomal protein,mitochondrial, putative	3.60	X	LSU	
16	Tb927.10.600	MRPL29	3.59	X	LSU	
17	Tb927.7.7090	Hypothetical protein, conserved	3.57	X		
18	Tb927.5.3110	MRPL49	3.57	X	LSU	
19	Tb927.6.3930	KRIT1	3.56	X	LSU	
20	Tb927.10.280	COXVI	3.45	X		IV
21	Tb927.11.12040	Cytochrome c oxidase component, putative	3.39	X		IV

22	Tb927.10.7380	LSU ribosomal protein,mitochondrial, putative	3.36	X		
23	Tb927.9.13540	hypothetical protein, conserved	3.29			
24	Tb927.7.2760	MRPL22	3.28	X	LSU	
25	Tb927.4.2720	ATP dependent DEAD-box helicase	3.26	X		
26	Tb927.11.10170	MRPL20	3.26	X	LSU	
27	Tb927.5.3980	MRPL15	3.22	X	LSU	
28	Tb927.11.1630	LSU ribosomal protein,mitochondrial, putative	3.20	X	LSU	
29	Tb927.10.7910	MERS3	3.17	X		
30	Tb927.7.7010	LSU ribosomal protein,mitochondrial, putative	3.16	X	LSU	
31	Tb927.10.8320	COXIX	3.09	X		IV
32	Tb927.5.4120	LSU ribosomal protein,mitochondrial, putative	3.06	X	LSU	
33	Tb927.8.3450	Hypothetical protein, conserved	3.06	X		
34	Tb927.11.7790	MRPS16	3.05	X	SSU	
35	Tb927.7.3030	Hypothetical protein, conserved	3.05	X		
36	Tb927.3.820	LSU ribosomal protein,mitochondrial, putative	3.04	X	LSU	
37	Tb927.11.10400	mS70	3.02	X	SSU	
38	Tb927.4.1070	MRPL13	2.97	X	LSU	
39	Tb927.9.8290	MRPL30	2.96	X	LSU	
40	Tb927.11.3640	MRPL27	2.96		LSU	
41	Tb927.4.4620	COXVIII	2.94	X		IV
42	Tb927.10.6090	tRNA pseudouridine synthase A, putative	2.92			
43	Tb927.10.13600	Cardiolipin-dependent protein 18kDa	2.88	X		IV
44	Tb927.10.13300	MRPS8	2.80	X	SSU	
45	Tb927.10.4880	Hypothetical protein, conserved	2.75	X		IV
46	Tb927.5.1510	mS23	2.74	X	SSU	
47	Tb927.11.2930	Membrane-associated protein, putative	2.73	X		
48	Tb927.11.4650	MRPL52	2.70	X	LSU	
49	Tb927.6.4200	LSU ribosomal protein,mitochondrial, putative	2.70	X	LSU	

50	Tb927.3.5600	Hypothetical protein, conserved	2.69			

References

1. Verner, Z., Basu S., Benz C., Dixit S., Dobáková E., Faktorová D., Hashimi H., Horáková E., Huang Z., Paris Z., Peña-Diaz P., Ridlon L., Týč J., Wildridge D., Zíková A., Lukeš J. (2015) Malleable mitochondrion of *Trypanosoma brucei*. *Int Rev Cell Mol Biol.* **315:** 73–151.

2. Jensen, R.E., Englund, P.T. (2012) Network News: The replication of kinetoplast DNA. *Annu Rev Microbiol.* **66**: 473–491.

3. Ramrath D.J.F., Niemann M., Leibundgut M., Bieri P., Prange C., Horn E.K., Leitner A., Boehringer D., Schneider A., Ban N. (2018) Evolutionary shift toward protein-based architecture in trypanosomal mitochondrial ribosomes. *Science*. **362**: eaau7735.

4. Read L.K., Lukeš J., Hashimi H. (2016) Trypanosome RNA editing: the complexity of getting U in and taking U out. *Wiley Interdiscip Rev RNA*. **7**: 33–51.

5. Stuart K.D., Schnaufer A., Ernst N.L., Panigrahi A.K. (2005) Complex management: RNA editing in trypanosomes. *Trends Biochem Sci*. **30**: 97–105.

6. Hajduk, S., Ochsenreiter, T. (2010). RNA editing in kinetoplastids. *RNA Biol*. **7**: 229–236.

7. Cooper S., Wadsworth E.S., Ochsenreiter T., Ivens A., Savill N.J, Schnaufer A. (2019) Assembly and annotation of the mitochondrial minicircle genome of a differentiation-competent strain of *Trypanosoma brucei*. *Nucleic Acids Res.* **47**: 11304–11325.

8. Chen J., Rauch C.A., White J.H., Englund P.T., Cozzarelli, N.R. (1995) The topology of the kinetoplast DNA network. *Cell*. **80**: 61–69.

9. Shapiro T.A. (1993) Kinetoplast DNA maxicircles: networks within networks. *PNAS*. **90**: 7809–7813.

10. Povelones M.L. (2014) Beyond replication: division and segregation of mitochondrial DNA in kinetoplastids. *Mol Biochem Parasitol*. **196**: 53–60.

11. Drew M.E., Englund, P.T. (2001) Intramitochondrial location and dynamics of *Crithidia fasciculata* kinetoplast minicircle replication intermediates. *J Cell Biol.* **153**: 735– 744.

12. Ryant K.A., Englund, P.T. (1989) Synthesis and processing of kinetoplast DNA minicircles in*Trypanosoma equiperdum*. *Mol Cell Biol*. **9**: 3212–3217.

13. Melendy T., Sheline C., Ray, D.S. (1988) Localization of a type II DNA topoisomerase to two sites at the periphery of the kinetoplast DNA of *Crithidia fasciculata*. *Cell*. **55**: 1083–1088.

14. Ryan K.A., Englund P.T. (1989) Replication of kinetoplast DNA in *Trypanosoma equiperdum*: Minicircle H strand fragments which map at specific locations. *J Biol Chem*. **264**: 823–830.

15. Carpenter L.R., Englund P.T. (1995) Kinetoplast maxicircle DNA replication in *Crithidia fasciculata* and *Trypanosoma brucei*. *Mol Cell Biol.* **15**: 6794–6803.

16. Hoeijmakers J.H.J., Weijers P.J. (1980) The segregation of kinetoplast DNA networks in *Trypanosoma brucei*. *Plasmid.* **4**: 97–116.

17. Ogbadoyi E.O., Robinson D.R., Gull K. (2003) A high-order transmembrane structural linkage is responsible for mitochondrial genome positioning and segregation by flagellar basal bodies in trypanosomes. *Mol Biol Cell.* **14**: 1769–1779.

18. Schneider A., Ochsenreiter T. (2018) Failure is not an option - mitochondrial genome segregation in trypanosomes. *J Cell Sci*. **131**: jcs221820.

19. Gluenz E., Povelones M.L., Englund P.T., Gull K. (2011) The kinetoplast duplication cycle in *Trypanosoma brucei* is orchestrated by cytoskeleton-mediated cell morphogenesis. *Mol Cell Biol.* **31**: 1012–1021.

20. Peikert C.D., Mani J., Morgenstern M., Käser S., Knapp B., Wenger C., Harsman A., Oeljeklaus S., Schneider A., Warscheid B. (2017) Charting organellar importomes by quantitative mass spectrometry. *Nat Commun*. **8**: 15272.

21. Harsman A., Schneider A. (2017) Mitochondrial protein import in trypanosomes: Expect the unexpected. *Traffic.* **18**: 96–109.

22. 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.

23. Žárský V., Doležal P. (2016) Evolution of the Tim17 protein family. *Biol Direct.* **11**: 54.

24. Ferramosca A., Zara V. (2013) Biogenesis of mitochondrial carrier proteins: Molecular mechanisms of import into mitochondria. *Biochim et Biophys Acta.* **1833**: 494–502.

25. Zimmermann R., Neupert W. (1980) Transport of proteins into mitochondria. Posttranslational transfer of ADP/ATP carrier into mitochondria *in vitro*. *Eur J Biochem*. **109**: 217–29.

26. Hansen K.G., Herrmann J.M. (2019) Transport of Proteins into Mitochondria. P*rotein J*. **38**: 330–342.

27. Schulz C., Schendzielorz A., Rehling P. (2015) Unlocking the presequence import pathway. *Trends Cell Biol*. **25**: 265–75.

28. Marom M., Azem A., Mokranjac D. (2011) Understanding the molecular mechanism of protein translocation across the mitochondrial inner membrane: still a long way to go. *Biochim Biophys Acta*. **1808**: 990–1001.

29. Craig E.A. (2018) Hsp70 at the membrane: Driving protein translocation. *BMC Biol*. **16**: 11.

30. Kang P.J., Ostermann J., Shilling J., Neupert W., Craig E.A., Pfanner N. (1990) Requirement for hsp70 in the mitochondrial matrix for translocation and folding of precursor proteins. *Nature*. **348**: 137–43.

31. Horst M., Oppliger W., Rospert S., Schönfeld H.J., Schatz G., Azem A. (1997) Sequential action of two hsp70 complexes during protein import into mitochondria. *EMBO J*. **16**: 1842–9.

32. D'Silva P.D., Schilke B., Walter W., Andrew A., Craig E.A. (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.

33. Truscott K. N., Voos W., Frazier A. E., Lind M., Li Y., Geissler A., Dudek J., Müller H., Sickmann A., Meyer H.E., 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.

34. Frazier A.E., Dudek J., Guiard B., Voos W., Li Y., Lind M., Meisinger C., Geissler A., Sickmann A., Meyer H.E., Bilanchone V., Cumsky M.G., Truscott K.N., Pfanner N., Rehling P. (2004) Pam16 has an essential role in the mitochondrial protein import motor. *Nat Struct Mol Biol*. **11**: 226–33

35. Banerjee R., Gladkova C., Mapa K., Witte G., Mokranjac D. (2015) Protein translocation channel of mitochondrial inner membrane and matrix-exposed import motor communicate via two-domain coupling protein. *Elife*.**4**: e11897.

36. Laloraya S., Gambill B.D., Craig E.A. (1994) A role for a eukaryotic GrpE-related protein, Mge1p, in protein translocation. *Proc Natl Acad Sci U S A*. **91**: 6481–5.

37. Laloraya S., Dekker P.J.T. Voos W., Craig E.A., Pfanner N. (1995) Mitochondrial GrpE modulates the function of matrix Hsp70 in translocation and maturation of preproteins. *Mol Cell Biol*. **15**: 7098–7105.

38. Schneider H.C., Westermann B., Neupert W., Brunner M. (1996) The nucleotide exchange factor MGE exerts a key function in the ATP-dependent cycle of mt-Hsp70-Tim44 interaction driving mitochondrial protein import. *EMBO J.* **15**: 5796–5803.

39. Harsman A., Oeljeklaus S., Wenger C., Huot J.L., Warscheid B., Schneider A. (2016) The non-canonical mitochondrial inner membrane presequence translocase of trypanosomatids contains two essential rhomboid-like proteins. *Nat Commun.* **7**: 13707.

40. von Känel C., Muñoz-Gómez S.A., Oeljeklaus S., Wenger Ch., Warscheid B., Wideman J.G., Harsman A., Schneider A. (2020) Homologue replacement in the import motor of the mitochondrial inner membrane of trypanosomes. *Elife.* **9**: e52560.

41. Tschopp F., Charrière F., Schneider A. (2011) *In vivo* study in *Trypanosoma brucei* links mitochondrial transfer RNA import to mitochondrial protein import. *EMBO Rep.* **12**: 825–32.

42. Zíková A., Verner Z., Nenarokova A., Michels P.A.M., Lukeš J. (2017) A paradigm shift: The mitoproteomes of procyclic and bloodstream *Trypanosoma brucei* are comparably complex. *PLoS Pathog.* **13**: e1006679.

43. Amodeo S., Jakob M., Ochsenreiter T. (2018) Characterization of the novel mitochondrial genome replication factor MiRF172 in *Trypanosoma brucei*. *J Cell Sci.* **131**: jcs211730.

44. Týč J., Klingbeil M.M., Lukeš J. (2015) Mitochondrial heat shock protein machinery hsp70/hsp40 is indispensable for proper mitochondrial DNA maintenance and replication. *mBio.* **6**: e02425–14.

45. Sullenberger C., Hoffman B., Wiedeman J., Kumar G., Mensa-Wilmot K. (2021) Casein kinase TbCK1.2 regulates division of kinetoplast DNA, and movement of basal bodies in the African trypanosome. *PLoS One.* **16**: e0249908.

46. Grams J., Morris J.C., Drew M.E., Wang Z., Englund P.T., Hajduk S.L. (2002) A trypanosome mitochondrial RNA polymerase is required for transcription and replication. *J Biol Chem.* **277**: 16952–9.

47. Trikin R., Doiron N., Hoffmann A., Haenni B., Jakob M., Schnaufer A., Schimanski B., Zuber B., Ochsenreiter T. (2016) Correction: TAC102 Is a Novel Component of the Mitochondrial Genome Segregation Machinery in Trypanosomes. *PLoS Pathog.* **12**: e1005750.

48. Sykes S.E., Hajduk S.L. (2013) Dual functions of α-ketoglutarate dehydrogenase E2 in the Krebs cycle and mitochondrial DNA inheritance in *Trypanosoma brucei*. *Eukaryot Cell.* **12**: 78–90.

49. Beck K., Acestor N., Schulfer A., Anupama A., Carnes J., Panigrahi A.K., Stuart K. (2013) *Trypanosoma brucei* Tb927.2.6100 is an essential protein associated with kinetoplast DNA. *Eukaryot Cell*. **12**: 970–8.

50. Grewal J.S., McLuskey K., Das D., Myburgh E., Wilkes J., Brown E., Lemgruber L., Gould M.K., Burchmore R.J., Coombs G.H., Schnaufer A., Mottram J.C. (2016) PNT1 Is a C11 Cysteine Peptidase Essential for Replication of the Trypanosome Kinetoplast. *J Biol Chem*. **291**: 9492–500.

51. Clement S.L., Mingler M.K., Koslowsky D.J. (2004) An intragenic guide RNA location suggests a complex mechanism for mitochondrial gene expression in *Trypanosoma brucei*.*Eukaryotic Cell.* **3**: 862–869.

52. Bakshi R.P., Shapiro T.A. (2004) RNA interference of *Trypanosoma brucei* topoisomerase IB: both subunits are essential. *Mol Biochem Parasitol*. **136**: 249–55.

53. Bodley A.L., Chakraborty A.K., Xie S., Burri C., Shapiro T.A. (2003) An unusual type IB topoisomerase from African trypanosomes. *PNAS.* **100**: 7539–7544.

54. Milman N., Motyka S.A., Englund P.T., Robinson D., Shlomai J. (2007) Mitochondrial origin-binding protein UMSBP mediates DNA replication and segregation in trypanosomes. *Proc Natl Acad Sci U S A*. **104**: 19250–5.

55. Hines J.C., Ray D.S. (2010) A mitochondrial DNA primase is essential for cell growth and kinetoplast DNA replication in *Trypanosoma brucei*. *Mol Cell Biol*. **30**: 1319–28.

56. Liu, B., Wang j., Yaffe N., Lindsay M., Zhao Z., Zick A., Shlomai J., Englund P.T. (2009) Trypanosomes have six mitochondrial DNA helicases with one controlling kinetoplast maxicircle replication. *Mol Cell.* **35**: 490–501.

57. Mokranjac D., Bourenkov G., Hell K., Neupert W., Groll M. (2006) Structure and function of Tim14 and Tim16, the J and J-like components of the mitochondrial protein import motor. *EMBO J*. **25**: 4675–85.

58. Krogh A., Larsson B., von Heijne G., Sonnhammer E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* **305**: 567–580.

59. Blum M., Chang H.Y., Chuguransky S., Grego T., Kandasaamy S., Mitchell A., Nuka G., Paysan-Lafosse T., Qureshi M., Raj S., Richardson L., Salazar G.A., Williams L., Bork P., Bridge A., Gough J., Haft D.H., Letunic I., Marchler-Bauer A., Mi H., Natale D.A., Necci M., Orengo C.A., Pandurangan A.P., Rivoire C., Sigrist C.J.A., Sillitoe I., Thanki N., Thomas P.D., Tosatto S.C.E., Wu C.H., Bateman A., Finn R.D. (2021) The

InterPro protein families and domains database: 20 years on. *Nucleic Acids Res*. **49**: D344–D354.

60. Kampinga H.H., Andreasson C., Barducci A., Cheetham M.E., Cyr D., Emanuelsson C., Genevaux P., Gestwicki J.E., Goloubinoff P., Huerta-Cepas J., Kirstein J., Liberek K., Mayer M.P., Nagata K., Nillegoda N.B., Pulido P., Ramos C., De Los Rios P., Rospert S., Rosenzweig R., Sahi C., Taipale M., Tomiczek B., Ushioda R., Young J.C., Zimmermann R., Zylicz A., Zylicz M., Craig E.A., Marszalek J. (2019) Function, evolution, and structure of J-domain proteins. *Cell Stress Chaperones*. **24**: 7–15.

61. Walsh P., Bursać D., Law Y.C., Cyr D., Lithgow T. (2004) The J-protein family: modulating protein assembly, disassembly and translocation. *EMBO Rep*. **5**: 567–71.

62. Nolan D.P., Voorheis H.P. (1992) The mitochondrion in bloodstream forms of *Trypanosoma brucei* is energized by the electrogenic pumping of protons catalysed by the F1F0-ATPase. *Eur J Biochem*. **209**: 207–16.

63. Schnaufer A., Clark-Walker G.D., Steinberg A.G., Stuart K. (2005) The F1-ATP synthase complex in bloodstream stage trypanosomes has an unusual and essential function. *EMBO J*. **24**: 4029–40.

64. Bhat G.J., Koslowsky D.J., Feagin J.E., Smiley B.L., Stuart K. (1990) An extensively edited mitochondrial transcript in kinetoplastids encodes a protein homologous to ATPase subunit 6. *Cell*. **61**: 885–94.

65. Dean S., Gould M.K., Dewar C.E., Schnaufer A.C. (2013) Single point mutations in ATP synthase compensate for mitochondrial genome loss in trypanosomes. *Proc Natl Acad Sci U S A*. **110**: 14741–6.

66. Wirtz E., Leal S., Ochatt C., Cross G.A. (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol*. **99**: 89–101.

67. Kapeller I., Milman N., Yaffe N., Shlomai J. (2011) Interactions of a replication initiator with histone H1-like proteins remodel the condensed mitochondrial genome. *J Biol Chem*. **286**: 40566–74.

68. Lai D.H., Hashimi H., Lun Z.R., Ayala F.J., Lukes J. (2008) Adaptations of *Trypanosoma brucei* to gradual loss of kinetoplast DNA: *Trypanosoma equiperdum* and *Trypanosoma evansi* are petite mutants of *T. brucei*. *Proc Natl Acad Sci U S A*. **105**: 1999–2004.

69. Yaffe N., Rotem D., Soni A., Porath D., Shlomai J. (2021) Direct monitoring of the stepwise condensation of kinetoplast DNA networks. *Sci Rep*. **11**: 1501.

70. Hoffmann A., Käser S., Jakob M., Amodeo S., Peitsch C., Týč J., Vaughan S., Zuber B., Schneider A., Ochsenreiter T. (2018) Molecular model of the mitochondrial genome segregation machinery in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A*. **115**: 1809–1818.

71. Schimanski B., Aeschlimann S., Stettler P., Käser S., Gomez-Fabra Gala M., Bender J., Warscheid B., Vögtle F.N., Schneider A. (2022) p166 links membrane and intramitochondrial modules of the trypanosomal tripartite attachment complex. *PLoS Pathog*. **18**: e1010207.

72. Kampinga H.H., Craig E.A. (2010) The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nat Rev Mol Cell Biol*. **11**: 579–92.

73. Ajit Tamadaddi C., Sahi C. (2016) J domain independent functions of J proteins. *Cell Stress Chaperones*. **21**: 563–70.

74. Pulido P., Leister D. (2018) Novel DNAJ-related proteins in *Arabidopsis thaliana*. *New Phytol*. **217**: 480–490.

75. Tamadaddi C., Verma A.K., Zambare V., Vairagkar A., Diwan D., Sahi C. (2022) J-like protein family of *Arabidopsis thaliana*: the enigmatic cousins of J-domain proteins. *Plant Cell Rep*. **41**: 1343–1355.

76. Bentley S.J., Jamabo M., Boshoff A. (2019) The Hsp70/J-protein machinery of the African trypanosome, *Trypanosoma brucei*. *Cell Stress Chaperones*. **24**: 125–148.

77. Panigrahi, A.K., Ogata, Y., Ziková, A., Anupama, A., Dalley, R.A., Acestor, N., Myler, P.J., and Stuart, K.D. (2009) A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome. *Proteomics*. **9**: 434–450.

78. Acestor N., Ziková A., Dalley R.A., Anupama A., Panigrahi A.K., Stuart K.D. (2011) *Trypanosoma brucei* mitochondrial respiratome: composition and organization in procyclic form. *Mol Cell Proteomics*. **10**: M110.006908.

79. Acestor N., Panigrahi A.K., Ogata Y., Anupama A., Stuart K.D. (2009) Protein composition of *Trypanosoma brucei* mitochondrial membranes. *Proteomics*. **9**: 5497–508.

80. Niemann M., Wiese S., Mani J., Chanfon A., Jackson C., Meisinger C., Warscheid B., Schneider A. (2013) Mitochondrial outer membrane proteome of *Trypanosoma brucei* reveals novel factors required to maintain mitochondrial morphology. *Mol Cell Proteomics*. **12**: 515–28.

81. Louw C.A., Ludewig M.H., Mayer J., Blatch G.L. (2010) The Hsp70 chaperones of the Tritryps are characterized by unusual features and novel members. *Parasitol Int.* **59**: 497–505.

82. Urbaniak M.D., Guther M.L., Ferguson M.A. (2012) Comparative SILAC proteomic analysis of *Trypanosoma brucei* bloodstream and procyclic lifecycle stages. *PLoS One*. **7**: e36619.

83. Gunasekera K., Wüthrich D., Braga-Lagache S., Heller M., Ochsenreiter T. (2012) Proteome remodelling during development from blood to insect-form *Trypanosoma brucei* quantified by SILAC and mass spectrometry. *BMC Genomics*. **13**: 556.

84. Butter F., Bucerius F., Michel M., Cicova Z., Mann M., Janzen C.J. (2013) Comparative proteomics of two life cycle stages of stable isotope-labeled *Trypanosoma brucei* reveals novel components of the parasite's host adaptation machinery. *Mol Cell Proteomics*. **12**: 172–9.

85. Schneider A. (2020) Evolution of mitochondrial protein import - lessons from trypanosomes. *Biol Chem*. **401**: 663–676.

86. Brun R., Schönenberger. (1979) Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. *Short communication. Acta Trop*. **36**: 289–92.

87. Hirumi H., Hirumi K. (1989) Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *J Parasitol*. **75**: 985–9.

88. Bochud-Allemann N., Schneider A. (2002) Mitochondrial substrate level phosphorylation is essential for growth of procyclic *Trypanosoma brucei*. *J Biol Chem*. **277**: 32849–54.

89. Oberholzer M., Morand S., Kunz S., Seebeck T. (2006) A vector series for rapid PCR-mediated C-terminal in situ tagging of *Trypanosoma brucei* genes. *Mol Biochem Parasitol*. **145**: 117–20.

90. 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.

91. Abràmoff M.D., Magalhaes P.J., Ram S.J. (2004) Image Processing with ImageJ. *Biophotonics International*. **11**: 36–42.

92. Chomczynski P., Sacchi N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*. **162**: 156–9.

93. Liu B., Molina H., Kalume D., Pandey A., Griffith J.D., Englund P.T. (2006) Role of p38 in replication of *Trypanosoma brucei* kinetoplast DNA. *Mol Cell Biol*. **26**: 5382–93.

94. Lamour N., Rivière L., Coustou V., Coombs G.H., Barrett M.P., Bringaud F. (2005) Proline metabolism in procyclic *Trypanosoma brucei* is down-regulated in the presence of glucose. *J Biol Chem*. **280**: 11902–10.

95. Madeira F., Pearce M., Tivey A.R.N., Basutkar P., Lee J., Edbali O., Madhusoodanan N., Kolesnikov A., Lopez R. (2022) Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res*. **50**: W276–9.

96. Verner Z., Cermáková P., Skodová I., Kováčová B., Lukeš J., Horváth A. (2014) Comparative analysis of respiratory chain and oxidative phosphorylation in *Leishmania tarentolae*, *Crithidia fasciculata*, *Phytomonas serpens* and procyclic stage of *Trypanosoma brucei*. *Mol Biochem Parasitol*. **193**: 55–65.

97. Bradford M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*. **72**: 248–54.

Extended Results:

Mitochondrial localization of TbPam18 (Tb927.8.6310) and TbPam16 (Tb927.9.13530)

In order to study these two proteins, we performed endogenous tagging with a construct bearing C-terminal V5 tag. According to the TrypTag *in-situ* tagging database (Dean *et al*., 2015), both TbPam proteins are genuine components of the *T. brucei* mitochondrion.

Localization of the *in situ*-tagged lines using immunofluorescence showed a mitochondrial localization for TbPam18. This localization was subsequently also determined using digitonin fractionation, where the V5-tagged signal was observed in the mitochondrial fraction. From TbPam16, *in* situ V5 tagging was not as successful and we observed high signal levels when detected by fluorescence microscopy. Verification of the localization using digitonin fractionation was already more successful and the V5 signal was observed in the mitochondrial fraction (Fig. 10).
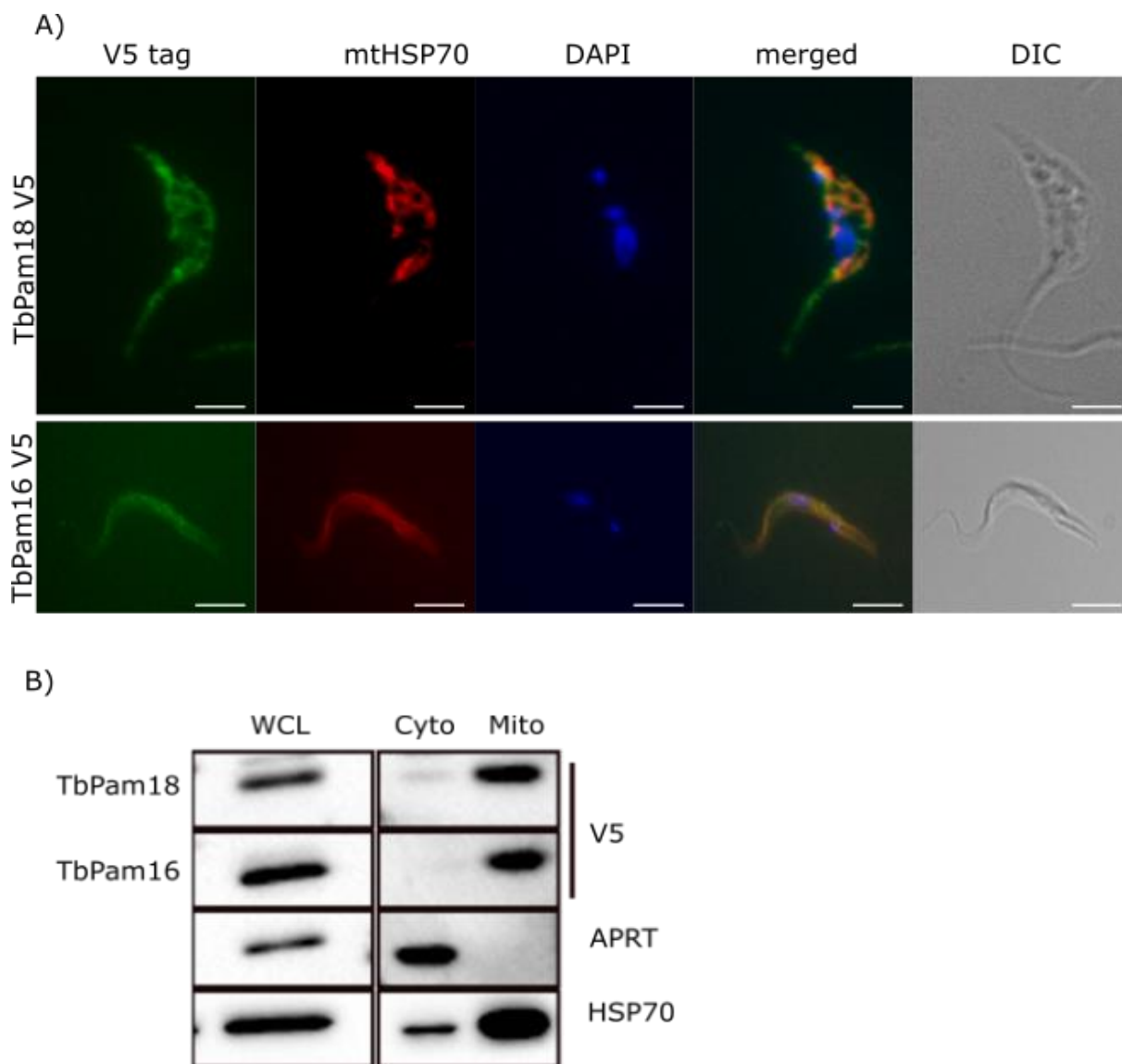
Figure10: Intracellular localization of TbPam18 and TbPam16. A) *In situ* C-terminally V5-tagged proteins were expressed in *T. brucei* and their localization was inspected using immunofluorescence microscopy. Monoclonal α-V5 rabbit and α-mHsp70 mouse antibodies were used. DNA was stained with DAPI. DIC, differential interference contrast. Scale bars, 5 μm. B) Western blot analysis over cytoplasmic (Cyt) and mitochondrial (Mito) fractions obtained by digitonin-based subcellular fraction of V5-tagged transfected procyclic forms. α-APRT rabbit and α-mtHSP70 mouse were used as a cytoplasmic and mitochondrial marker. WCL: Whole cell lysate.

kDNA and cell-division related phenotype

We received TbPam18 RNAi and TbPam16 RNAi cell lines from the laboratory of A. Schneider (University of Bern). In a previous study it was shown that the ablation of TbPam18 and TbPam16 influences cell growth on day 4 of RNAi induction (von Känel *et al*., 2020). The absence of a tag did not allow monitoring the protein levels. Hence, the phenotype was followed in days post-RNAi induction when the effect of decreased levels of targeted proteins should have been apparent.

In *T. brucei*, kDNA divides once per cell cycle, immediately prior to the nuclear DNA synthesis, the replication of kDNA thus being in synchrony with the nuclear division (Woodward *et al*., 1990). In an unsynchronized culture, most cells are in a 1N1K stage, which means that they contain one nucleus and one kDNA. Replication starts by duplication of the basal body, followed by the division of kDNA (Kdiv), then the kDNA network splits, producing a 1N2K cell. Finally, the nucleus divides and the ensuing 2N2K cell undergoes cytokinesis (Concepción-Acevedo *et al*., 2012; Chowdhury *et al*., 2008). As kDNA replication in *T. brucei* cannot be synchronized, unsynchronized cultures were analysed, with cells being individually assigned to their respective categories. At least 200 cells were counted per each time point (Fig. 10A). TbPam18 and TbPam16 are mitochondrial proteins, so we expected that their ablation would affect mitochondrial morphology. To our surprise, we did not observe this effect, but an unexpected effect on the cell cycle of *T. brucei* was detected.

On the sixth day of induction, a 15% decrease in the number of cells with dividing kinetoplast was observed in both cell lines. There was also a 15% increase for TbPam18 and 20% increase for TbPam16 in the number of cells, which completely lack the kinetoplast (0K1N). The presence of multinucleated cells 0K2N, 0K3N suggested a defect in the cell cycle and cytokinesis (Fig. 11). Surprisingly, a decrease in the kDNA area was observed in the RNAi-induced cells (see the results in the attached manuscript).
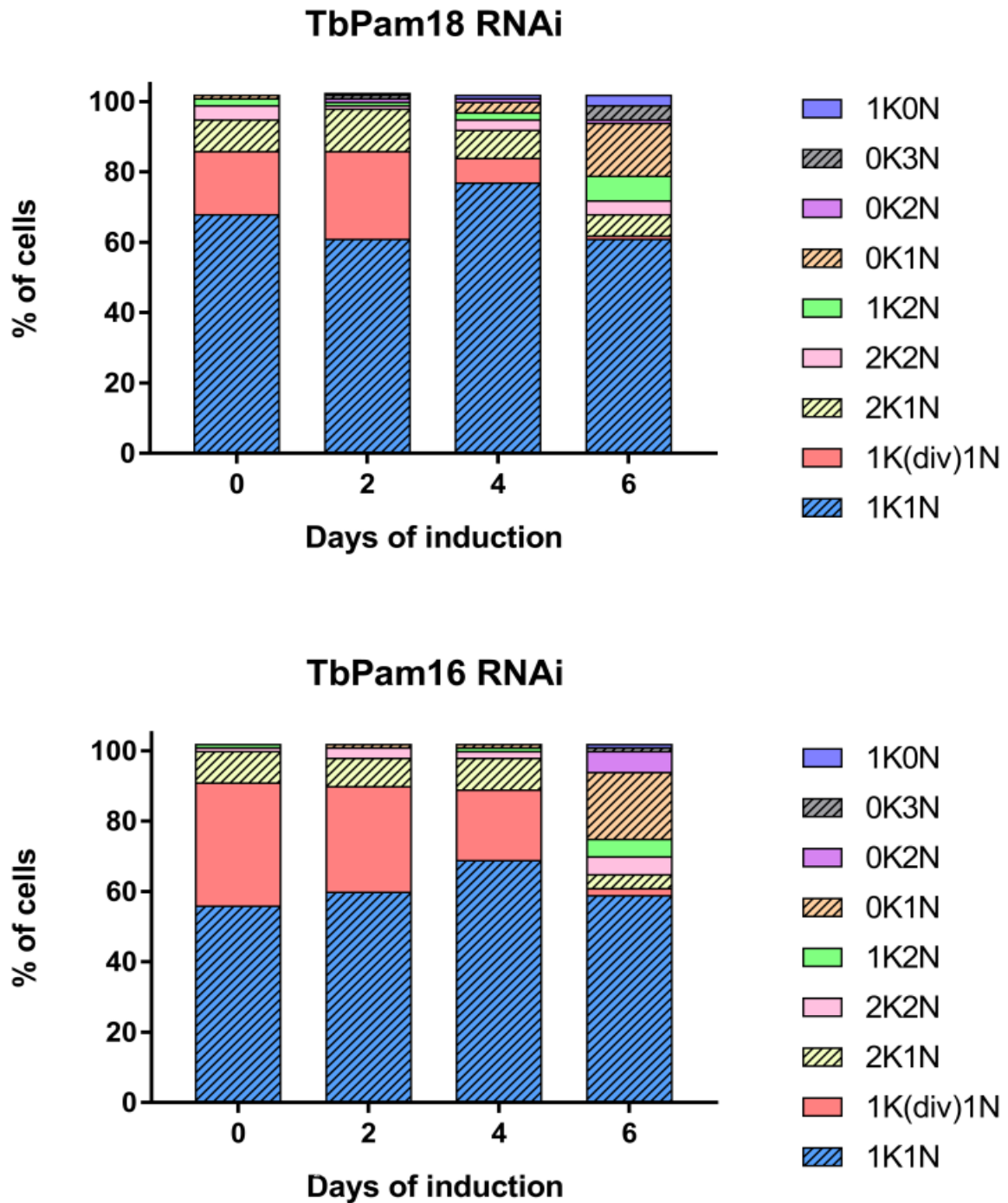
Figure 11. Distribution of cell cycle stages after depletion of TbPam18 and TbPam16. A) Quantification of relative nuclei/kinetoplasts numbers for non-induced (day 0) and induced (days 2, 4 and 6 post-addition of Tet; 1µg/ml) on DAPI stained procyclic cells. IFA quantification for each day of induction was performed by measuring at least 200 cells per time point.

In vivo mitochondrial translation assay

Generated RNAi strain allowed the observation of a phenotype that became apparent after day 4 post-induction. Ablation of TbPam18 and TbPam16 caused reduction of the kDNA maxicircles but not minicircles. Maxicircles of *T. brucei* and other trypanosomatids encode, among others, apocytochrome b (Cyb) of complex III and cytochrome *c* oxidase subunit I (COI) of complex IV. These two polypeptides can be separated by a two-dimensional SDS-PAGE, and their synthesis can thus be investigated *in vivo* after inhibition of the cytosolic translation with cycloheximide (Horváth *et al.*, 2000). Day 0, 2 and 4 after the induction, an aliquot (1 × 10^7 cells) was withdrawn from each RNAi-induced culture and labelled with [35S]methionine-cysteine. The COI and Cyb products are visible as two prominently labelled spots (Fig. 12A-B).

While the signal for TbPam16 was comparable in uninduced and RNAi-induced cells (Fig. 3B), the effect of the TbPam18 RNAi on the COI and Cyb synthesis was visible already at day 2 post-RNAi induction, and the difference was well pronounced at day 4 p.i (Fig. 12A). These gels were also stained with Coomassie brilliant blue as loading control (insets, Fig. 12A–B).
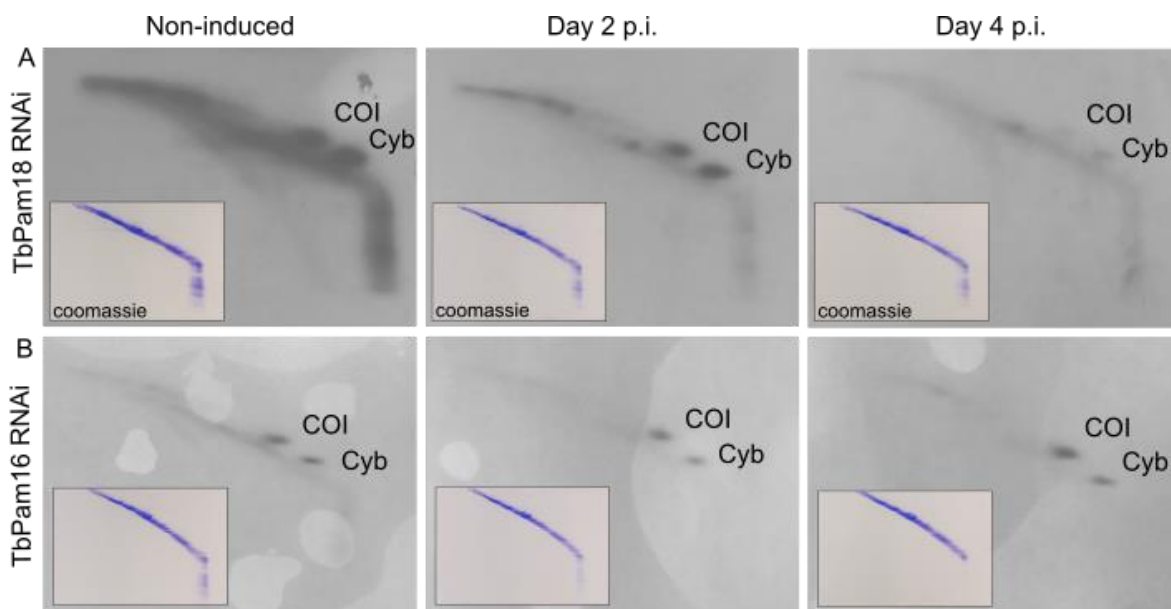
Figure 12 – Ablation of TbPam18 RNAi (A) nor TbPam16 RNAi (B) leads to reduction in mitochondrial protein synthesis. Autoradiographs of 2D gel electrophoresis of total mitochondrial [35S] methionine–cysteine labelled proteins in the presence of 100 µg/ml cycloheximide, used to assess *de novo* mitochondrial translation. Cells in aliquots of the non-induced cultures (left column of panels),

the cultures at day 2 postinduction (middle column of panels and the cultures at day 4 postinduction (right column of panels). Positions of the mitochondrial polypeptides COI and Cyb are indicated.

To determine if this reduction in protein synthesis impacts the function of respiratory complexes, we measured the activity of complex III and complex IV in isolated mitochondria, by spectrophotometric assay (see the results in the attached manuscript).

References for additional results

Concepción-Acevedo J., Luo J., Klingbeil M.M. (2012) Dynamic localization of *Trypanosoma brucei* mitochondrial DNA polymerase ID. *Eukaryot Cell.* **11**: 844–855.

Dean S., Sunter J., Wheeler R.J., Hodkinson I., Gluenz E., Gull K. (2015) A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol*. **5**: 140197.

Englund P.T. (1979) Free minicircles of kinetoplast DNA in *Crithidia fasciculata*. *J Biol Chem.* **254**: 4895–4900.

Horváth A., Berry E.A., Huang L.S., Maslov D.A. (2000) *Leishmania tarentolae*: a parallel isolation of cytochrome bc(1) and cytochrome c oxidase. *Exp Parasitol*. **96**: 160–7.

Chowdhury A.R., Zhao Z., Englund P.T. (2008) Effect of hydroxyurea on procyclic *Trypanosoma brucei*: an unconventional mechanism for achieving synchronous growth. *Eukaryot Cell.* **7:** 425– 428.

Jensen, R.E., Englund, P.T. (2012) Network News: The replication of kinetoplast DNA. *Annu Rev Microbiol.***66**: 473–491.

von Känel C., Muñoz-Gómez S.A., Oeljeklaus S., Wenger Ch., Warscheid B., Wideman J.G., Harsman A., Schneider A. (2020) Homologue replacement in the import motor of the mitochondrial inner membrane of trypanosomes. E*Life*. **9**: e52560.

Woodward R., Gull K. (1990) Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. *J Cell Sci.* **95**: 49–57.

Chapter Summary

In conclusion, we have studied the function and structure of TbPam18 and TbPam16 in *Trypanosoma brucei*, with the following conclusions:

1. TbPam18 and TbPam16 are localised in the mitochondrion of *T. brucei.*

2. Ablation of TbPam18 RNAi and TbPam16 RNAi affects *T. brucei* cell cycle.

3. Ablation of TbPam18 RNAi nor TbPam16 RNAi leads to reduction in mitochondrial protein synthesis.

4. Ablation of TbPam18 decrease the activity of respiratory complexes and membrane potential.

5. Depletion of TbPam18 and TbPam16 causes loss of kDNA size.

Chapter IV.

Novel protein complex involved in kinetoplast DNA replication and maintenance

Manuscript in preparation

Abstract

The single kinetoplast containing mitochondrial DNA (kDNA) is the defining feature of Kinetoplastids, and the tripartite attachment complex (TAC) connects the basal body with the kDNA coupling its synchronized segregation during cell division. Both these complex subcellular structures are appealing targets of intervention against trypanosomatids driven neglected diseases. Here, we screened the TrypTag *in-situ* tagging database (Dean *et al*., 2015) localization repository and prioritized 10 previously undescribed putative proteins displaying kinetoplast proximal enrichment (KEP) and therefore chosen for further validation. From these, 70% were verified by alternate endogenous tagging for kinetoplast or TAC localization. The essentiality of these proteins was assessed by RNAi knock-down, revealing KP56, KP84 and KP86 as essential for parasite growth in the procyclic stage. These interactors were as well verified for KP associated localization, resulting in an enrichment of proteins associated with KP and TAC regions. Moreover, this work provides a methodological pipeline for the identification of novel KP/TAC associated targets for intervention against trypanosomatids caused diseases.

Materials and Methods

Preparation of cell lines

T. brucei procyclic stage SmOx cell line (Poon *et al.*, 2012) was grown in SDM79 supplemented with 10% fetal bovine serum for most of the experiments. Selected proteins were *in situ*-tagged by a recently developed PCR-based transfection protocol described elsewhere (Dean *et al.*, 2015). Ten proteins identified in the MitoTag study (Pyrih *et al.,* under review) were C-terminally or N-terminally V5-tagged (see Table 1) using a previously modified pPOTv4 vector (Peña-Diaz *et al.*, 2017). For transfection, 10^7 exponentially growing cells were washed and re-suspended in Human T-Cell Nucleofector transfection solution (Lonza) according to the manufacturer's instructions and transfected using the AMAXA program X-014. 12 to 16 h post transfection, cells were diluted in 24-well plates and incubated for 2 weeks in the presence of hygromycin (Invitrogen) at 50 µg/ml final concentration or blasticidin (InvivoGen-BARIA) at 10 mg/ml. Plates were evaluated for the expression of V5-tagged proteins by Western blotting.

Primer	*In situ*-tagging	Sequence
KP56 Fw	N-term	AATACCACCCGCTGTAGTGCAAGATCATTTGTAGCTGT ATTCGTCATGTGGACGGAAGGACCTCATAGCGATACGT AAAAGTATAATGCAGACCTGCTGC
KP56 Rev		CCACAAAGTGTAATAGAAGGAGTAAGTGGGATGACAG TGCACGTATCAGTGGGCATGAGCCGTACGGTGATGCG ATGCATACTACCCGATCCTGATCCAGATCCTGATCCGG ATCCCGTGCTATCAAGACCGAGGAGGGG
KP86 Fw	C-term	GAGATACTCTCACAGACTTGCTGCAGAGAAATCCCAAC GCCAGCAGTGGGACCTCCCCTACAACCGACACACAGC ATAACGGTTCTGGTAGTGGTTCC
KP86 Rev		GAGCATGCCCTCGCTATGTACTCGCACTTTTATTATATG TTATTATTTGTTTCCGTTTAAACTCTATGCTGTATGTCT ATCCAATTTGAGAGACCTGTGC
KP84 Fw	N-term	TATAGAGAGAGAGTTTCATACGCATTTCTTACGTAA CAAAGGCGGAAACCAAAGCACACCAAAAAAAAAAAA AGAGGAGTATAATGCAGACCTGCTGC
KP84 Rev		CTAGCGTTTCCCCAACATTGCAAACGGTTCCACACCGC AATGGCATCCTCCATAACATTCTGAGGTACCTGTGTTT CCATACTACCCGATCCTGATCCAGATCCTGATCCGGAT CCCGTGCTATCAAGACCGAGGAGGGG
KP37 Fw	C-term	AGAAAGTTTGCAAGCGATTGGAGGAAAATCGCATACC ACTCCACCGTGTTACAGTCGAAAGGTTGGAAGCACTCA AGTTGGGTTCTGGTAGTGGTTCC
KP37 Rev		AAAAAGGAACCCCCAGCTGTACAACCTCCCCAAGGGA GTTGGGGGAGGGGGAAAATATGGCGGGACCCCACCAT ACGCGGCCAATTTGAGAGACCTGTGC

Table 1: Primer sequences for *in situ*-tagging.

RNAi construct preparation

For RNAi, gene fragments were PCR amplified from *T. brucei* genomic DNA using primers listed in Table 2. The PCR amplicon was extracted from gel and purified, then cloned into the pET/100D-TOPO expression vector (Invitrogen). The resulting expression plasmid encoding was transformed into the E. coli BL21(DE3) Star strain (Novagen) grown on LB agar. 35 µl of X-gal (5-bromo-4-chloro-3-indolyl-beta-galactopyranoside) and 5 µl IPTG (isopropylthio-β-galactoside) were added to ampicillin plates for blue-white selection. Successfully transformed *E. coli* containing modified vector (white colonies) were selected for growth in LB media. DNA isolation was performed according to the QIAprep Spin Miniprep Kit from Qiagen. We chose the p2T7-177 plasmid to clone our insert from the TOPO vector. Plasmids were digested with indicated restriction enzymes (HindIII and BamHI) and ligated into the p2T7-177 vector pre-digested with the same enzymes. All plasmids were linearized using NotI, electroporated, and selected (phleomycin 50 µg/ml final concentration) as described for *in situ*-tagging. RNAi was induced by the addition of tetracycline (1 µg/ml) to the growth medium, and cell densities were measured using a Beckman Coulter Z2 counter every 24 h over a period of 6 days after RNAi induction.

Primer	Sequence
KP56 RNAi Fw	**CGCAAGCTT**CTGATTCAGTGGTCACGGCT
KP56 RNAi Rev	**CGCGGATCC**AAGCCAACACTCCGCTACTC
KP86 RNAi Fw	**CGCAAGCTT**AGGCTGATGAGAAAGCGGAG
KP86 RNAi Rev	**CGCGGATCC**TGCGACGGTTGAAAAACACC
KP84 RNAi Fw	**CGCAAGCTT**ACCGTCATCATTCCTCTCGC
KP84 RNAi Rev	**CGCGGATCC**GAACGCACAGATGGGTTGAC
KP37 RNAi Fw	**CGCAAGCTT**GCACCTCCCGTGGATACATT
KP37 RNAi Rev	**CGCGGATCC**AGGCTACAACCACCACGAAG

Table 2: primers for RNAi: 5´part of Gene of Interrest + HindIII (Fw – blue) and BamHI (Rev – red) restriction sites + CGC overhang, finally joined with p2T7-177 backbone

RNAi phenotype analysis

Cultures were grown in biological triplicate in the presence (Dox+) or absence (Dox-) of doxycycline (Sigma-Merck). Cell density was counted using the Beckman Coulter Z2 Cell and Particle Counter every 24 hours and cells were subsequently diluted to 2×10^6 cells/ml, maintaining cultures in the exponential phase of growth. For Western blot analysis, monoclonal α-V5 (Life Technologies), α-HSP70 antibodies (Sigma-Aldrich), and secondary HRP-conjugated α-mouse IgG antibodies (Sigma-Aldrich) were used, with signal visualized by Clarity Western ECL Blotting Substrate (Bio-Rad).

Immunofluorescence

10^6 *T. brucei* cells were fixed with 4% paraformaldehyde in phosphate buffered saline (PBS) for 30 min at RT. The cells were subsequently centrifugated at 1,300 x g for 5 min and resuspended in 200 µl PBS and placed on a glass slide for 30 min. The slides were washed once with PBS and placed in methanol for 30 min at -20°C. The slides were washed three times with PBS and blocked with 5% milk inside a humid chamber for 1 h before being incubated with V5-mouse primary antibody (Invitrogen; 2.5:1000 dilution) overnight at 4°C. The slides were washed three times with PBS and incubated with secondary Alexa Fluor 488 goat anti-mouse IgG antibody (1:1000 dilution; Invitrogen) for 1 h at RT. The slides were once again washed three times with PBS before a final wash with ddH$_2$O, air-dried, and mounted in ProLong Gold antifade reagent with 4',6-diamidine-2'-phenylindole dihydrochloride (DAPI; Life Technologies). Images were captured with the BX51 Olympus Fluorescence Microscope.

Mitochondrial sub-fractionation

All steps were performed on ice, with precooled buffers to 4 °C. 0.015 % Digitonin in SOTE was heated in a water bath at 85 °C until it dissolved (up to 1 h). 10^8 cells were collected and pelleted at 12 °C at 1300 g for 10 min. The pellet was resuspended in 1 ml PBS-G (PBS + 5mM glucose), pellet was spined in a precooled centrifuge at 4 °C at 1300 g for 10 min. The pellet was resuspended in 1 ml SOTE (20 mM Tris-HCl pH 7.5, 0.6 M Sorbitol, 2 mM EDTA) with 0.015% Digitonin and it was incubated on ice for 5 min. Sample was spined in a precooled centrifuge at 4 °C at 5000 g for 5 min. The pellet consists of mitochondria was washed once with 1 ml SoTE, centrifuged and supernatant, cytosol fraction, was removed.

Individual fractions were evaluated for the localisation by Western blotting using specific antibodies. Cytosolic fraction was determined by monoclonal rabbit α-enolase (P. A. M. Michels) and mitochondrial fraction was determined via mouse α-mtHsp70 (kindly provided by Alena Zíková).

Results

Kinetoplast candidate selection

Ten initial kinetoplast candidates were selected for analysis, five of which were novel kinetoplast proteins identified in the MitoTag study (Pyrih *et al.*, under review). A sixth novel identified kinetoplast protein from MitoTag (Tb927.10.12180) was not pursued due to simultaneous investigation from a collegial lab. Another five selected candidates (Tab. 3) were categorised by MitoTag as exhibiting artificial <u>K</u>inetoplast <u>P</u>roximal <u>E</u>nrichment (KPE), but upon closer inspection, appeared to be genuine kinetoplast proteins worthy of further investigation. In order to establish whether these ten kinetoplast protein candidates are genuine constituents of the kinetoplast, we employed a variety of methods including alternate tagging, and knockdown studies (Fig. 14).

Of the initial ten candidates, six showed a distinct kinetoplast signal when tagged with a V5-epitope (two are not part of this study), while four were excluded from further analysis due to tagging difficulties (Fig. 13). One candidate (Tb927.7.4810) proved resistant to generating a tagged cell line, two proteins (Tb927.11.2360, Tb927.11.14120) produced cytosolic localization, while another candidate (Tb927.4.2780) upon V5 tagging showed an extremely faint signal, which only upon the generation of an overexpressed cell line produced a clearly identifiable kinetoplast signal (Fig. 14).

Protein Accession Number	New Annotation	Previous Annotation
Tb927.8.4040	KP56	Endonuclease G, putative
Tb927.10.15660	KP86	Hypothetical protein, conserved
Tb927.7.5320	KP84	Hypothetical protein, conserved
Tb927.6.4510	KP37	Hypothetical protein, conserved

**Table 3: Initial kinetoplast candidates.** Four initial kinetoplast candidates.
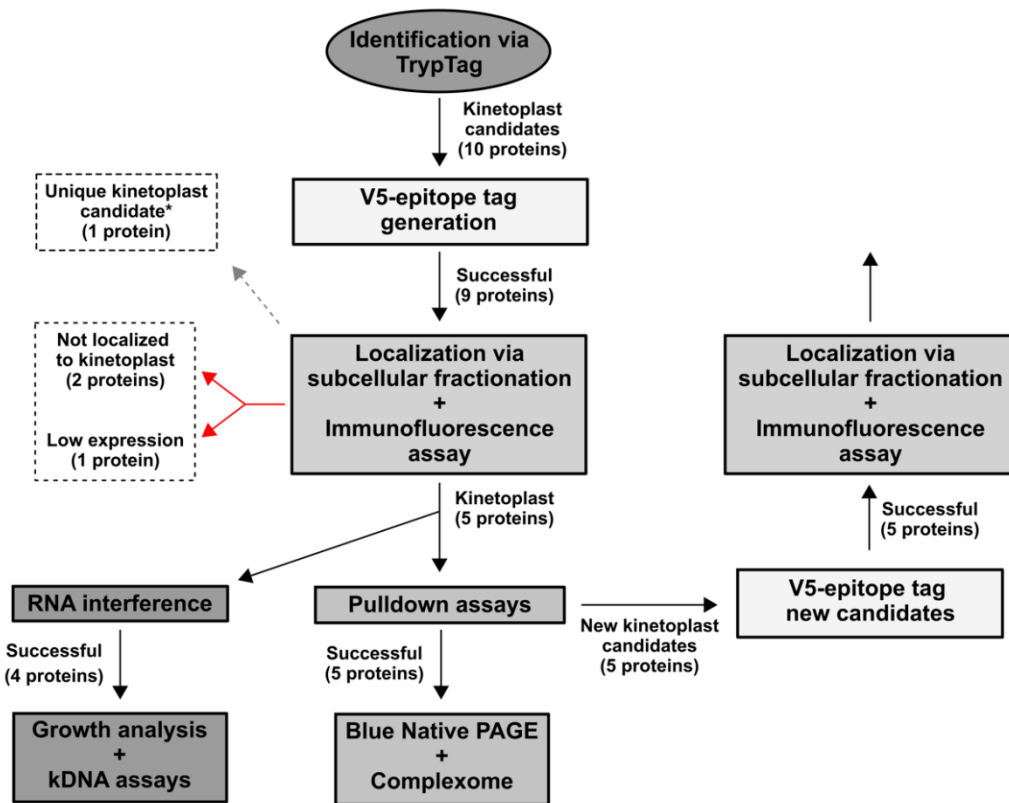
Figure 13: Schematic representation depicting methodology workflow. Ten protein candidates showed promising kinetoplast signals based the MitoTag study (Pyrih *et al.*, under review). Of these candidates, nine were successfully tagged with a V5-epitope, with three proteins excluded from further investigation due either lack of localization to the kinetoplast or low endogenous expression levels. A fourth candidate was additionally excluded due to a unique kinetoplast signal that warrants standalone investigation (asterisk). Four proteins showed colocalization to the kinetoplast and thus were subject to further investigation via RNA interference and pulldown assays.

A further four kinetoplast candidates (Tab. 3) would later be selected for validation, based on candidate sequence functional annotation, mitochondrial presence from previous *T. brucei* mitochondrial proteomes (Pyrih *et al.*, under review; Panigrahi *et al.*, 2009; Peikert *et al.*, 2017), the reported procyclic growth phenotypes from a former high throughput screen (Horn, 2021) as well as advice from collaborators.

Localisation

Based on results of digitonin fractionation (Fig. 14), we can confidently classify, KP86, KP84 and KP37 among the proteins found in the mitochondrial/organellar fraction of cells. KP56 V5 cell line failed to exhibit any signal from western analysis. The expression levels of each novel kinetoplast associated protein differed from one another, with KP56 being lowly expressed compared to the other four. However, Immunofluorescence analysis of V5 tagged lines demonstrated kinetoplast localization of all 4 initially selected kinetoplast candidates.
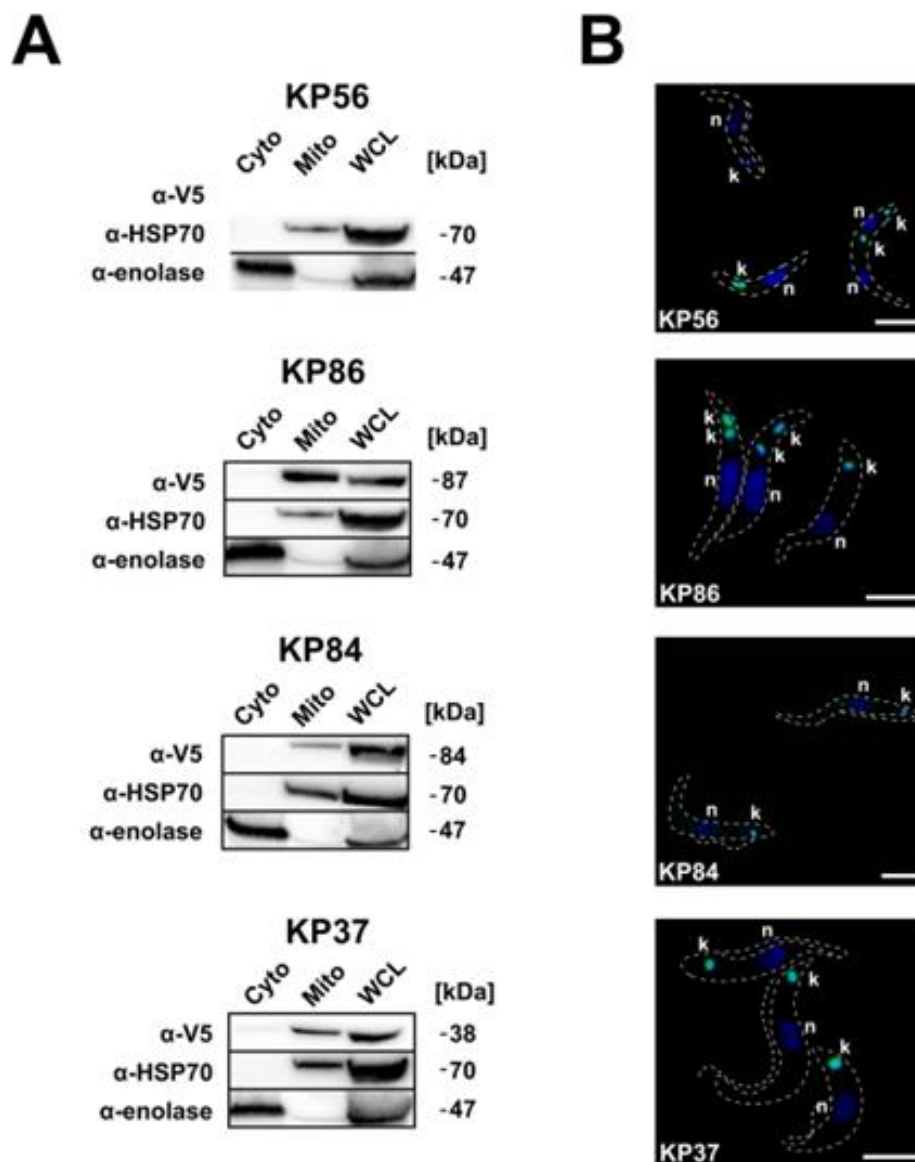
Figure. 14. Intracellular localization of selected KP proteins. A) V5-epitope tagged kinetoplast protein candidates showed enrichment within the mitochondrial fraction after separation of crude mitochondria from the cytosol. Western blot analysis over cytoplasmic (Cyt) and mitochondrial (Mito) fractions obtained by digitonin-based subcellular fraction of V5-tagged transfected procyclic forms.

Enolase and HSP-70 were used as a cytoplasmic and mitochondrial marker, respectively. WCL: Whole cell lysate. B) *In situ* C-terminally and N-terminally V5-tagged proteins were expressed in *T. brucei* and their localization was inspected using immunofluorescence microscopy. Nuclear (n) and kinetoplast DNA (k) were stained with DAPI (blue) and the respective V5-epitope tagged protein shown in green. The occurrence of cyan color shows localization of the tagged protein to the kinetoplast. Monoclonal α-V5 mouse and antibody was used. DNA was stained with DAPI. Scale bars, 10 μm.

**Cell viability upon KP targeted RNA interference**

In order to study whether these five novel kinetoplast associating proteins were necessary for cell viability, inducible RNA interference (RNAi) targeting the respective tagged KP protein was introduced. Growth curves revealed that depletion of KP56 and KP86 resulted in moderate growth phenotypes while KP84 appeared more drastic, with severe growth arrest occurring 48 hours after inducing RNAi. Noteworthy, KP37 did no exhibit any negative growth phenotype upon depletion (Fig. 15).

KP56 RNAi
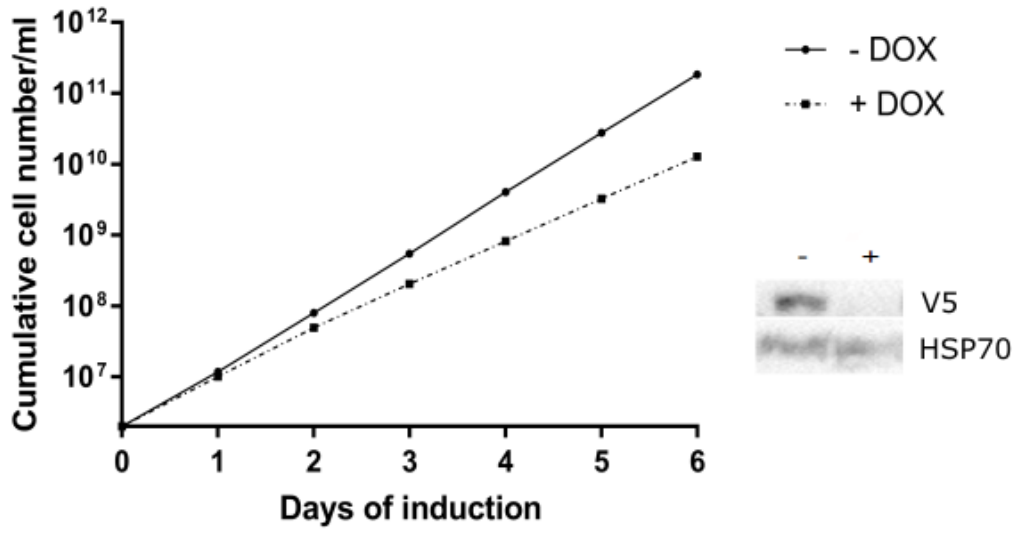
KP86 RNAi
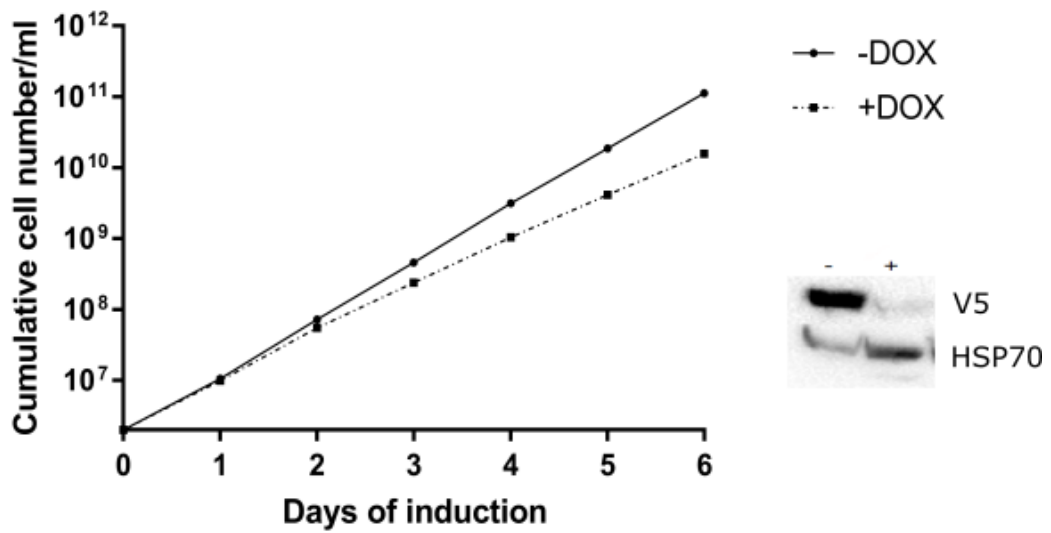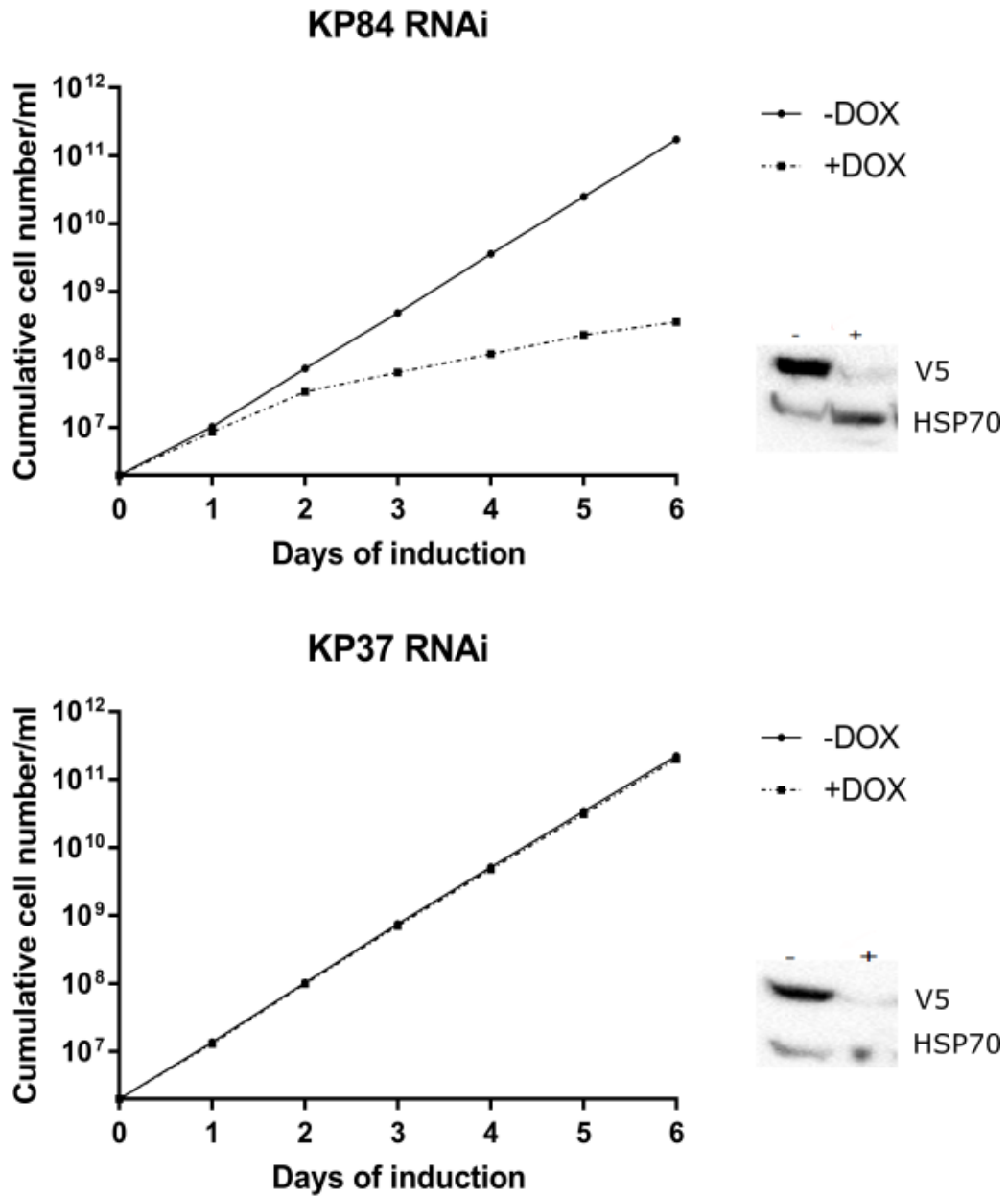
Figure 15: Ablation of KP56, KP86, KP84 nor KP37 affects the cell growth. Cell concentrations of non-induced (-) and RNAi-induced (+) cells are indicated. The experiment was performed in biological triplicate. Error bars represent standard deviations. Protein levels detected by Western blot analysis three days post induction. α-HSP70 antibody serves as a loading control.

Discussion

The principle finding of the MitoTag study was the unexpected demonstration that an mNeonGreen tag (along with other large globular epitopes) produced an artificial localization proximal to the kinetoplast in mitochondrial proteins, which were typically found in either the matrix or integral inner membrane (Pyrih *et al*., under review). Amongst a plethora of artificially sub-localized mitochondrial proteins identified from MitoTag, several novel proteins were also identified exhibiting genuine kinetoplast signals, that appeared distinct from an artificial signal pattern. To remove ambiguity for these novel kinetoplast candidates, it was viewed as a necessary follow-up procedure to employ a smaller V5 tag to redetermine localization, which specifically does not induce such artificial interactions, as a measure of validation for identified kinetoplast candidates (Pyrih *et al*., under review). The localization of all four chosen candidates to the kinetoplast via V5 tagging thus demonstrates these proteins as new components of the kinetoplast.

To our knowledge, the kinetoplast consists of 57 established proteins, thus the identification of six new kinetoplast-associated proteins is significant for such an intensely studied structure of the trypanosome mitochondrion (Pyrih *et al.*, under review; Panigrahi *et al.*, 2009; Peikert *et al.*, 2017; Horn, 2021). The kinetoplast has long been considered an attractive drug target, due to its unique presence amongst only the kinetoplastid clade, for which these novel proteins expand on its potential for therapeutic development. It was noted that several established kinetoplast proteins exhibited a staining pattern indistinguishable from an artificial interaction signal when tagged with mNeonGreen (Pyrih *et al*., under review), and consequently, it remains a possibility that certain other unidentified kinetoplast proteins exist within the MitoTag dataset.

The demonstration of growth phenotypes in several RNAi cell lines generated from these novel proteins also demonstrates that they are not spurious (useless) components of the kinetoplast but are in fact important for the viability of the cell. For future studies, it will be useful to determine the function of these proteins beyond the growth phenotypes they induce. We recommend first investigating the status of kinetoplast DNA upon RNAi induction. The measurement of certain mitochondrial metabolites may additionally hold promise in elucidating protein function considering the new metabolic potential within the kinetoplast (Pyrih *et al*., under review).

References

Dean S., Sunter J., Wheeler R.J., Hodkinson I., Gluenz E., Gull K. (2015) A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol*. **5**: 140197.

Horn D. (2022) Genome-scale RNAi screens in African trypanosomes. *Trends Parasitol*. **38**: 160–173.

Panigrahi, A.K., Ogata, Y., Zíková, A., Anupama, A., Dalley, R.A., Acestor, N., Myler, P.J., and Stuart, K.D. (2009) A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome. *Proteomics*. **9**: 434–450.

Peikert C.D., Mani J., Morgenstern M., Käser S., Knapp B., Wenger C., Harsman A., Oeljeklaus S., Schneider A., Warscheid B. (2017) Charting organellar importomes by quantitative mass spectrometry. *Nat Commun*. **8**: 15272.

Peña-Diaz P., Vancová M., Resl C., Field M.C., Lukeš J. (2017) A leucine aminopeptidase is involved in kinetoplast DNA segregation in *Trypanosoma brucei*. *PLoS Pathog*. **13**: e1006310.

Poon S.K., Peacock L., Gibson W., Gull K., Kelly S. (2012) A modular and optimized single marker system for generating *Trypanosoma brucei* cell lines expressing T7 RNA polymerase and the tetracycline repressor. *Open Biol*. **2**: 110037.

Pyrih J., Hammond M., Alves A., Dean S., Sunter J., Wheeler R., Gull K., Lukeš J. (Under review) Comprehensive sub-mitochondrial protein map of the parasitic protist *Trypanosoma brucei* reveals novel aspects of organellar biology.

Schneider A. (2018). Mitochondrial protein import in trypanosomatids: Variations on a theme or fundamentally different? *PLoS Pathog*. **14**: 1007351.

Chapter Summary

In conclusion, we screened the TrypTag localization repository and prioritized 10 previously undescribed putative proteins displaying kinetoplast proximal enrichment (KEP) in *Trypanosoma brucei*, with the following conclusions:

1. Work provides a methodological pipeline for the identification of novel KP associated targets.

2. KP56, KP86, KP84 and KP37 are bona fide novel kinetoplast associated proteins.

3. KP56, KP86, KP84 nor KP37 are essential for parasite growth in the procyclic stage.

5. Concluding remarks

The primary goal of the thesis is characterized the novel mitochondrial proteins in *Trypanosoma brucei*. For ZapE, a spectrum of seemingly unrelated functions ranging from apoptosis to cell division has been proposed (Marteyn *et al*., 2015; Cesnekova *et al*., 2016a; Cesnekova *et al*., 2016b; Germany *et al*., 2018). Curiously, it was not yet addressed how a single protein may perform such a plethora of functions. Here we propose a hypothesis, which connects most of the observed phenotypes into a single putative role of the ZapE protein. One of the three identified putative interaction partners of the ZapE2 paralog in *T. brucei* is Oxa1, a highly conserved protein required for both the insertion of mitochondrially-encoded subunits of the respiratory chain complexes (Koli *et al.,* 2018). We suggest that in eukaryotes Oxa1 and ZapE operate together, and that a depletion of the latter partner triggers a complex phenotype.

ZapE seems to be invariably present in the aerobic mitochondria, a key feature of which is the presence of numerous subunits of the respiratory chain complexes encoded by the mitochondrial genome (Smith *et al*., 2015). Their protein products are inserted with the assistance of the membrane insertase Oxa1 into the corresponding complexes residing in the inner membrane (Herrmann and Bonnefoy, 2004). In *T. brucei* in which both ZapE paralogs were downregulated, the activities of Oxa1 substrates, in particular complexes I and IV, significantly increased. As Oxa1 participates in the insertion of several subunits of the respiratory complexes into the inner mitochondrial membrane, we propose that ZapE may negatively regulate the function of Oxa1. We attribute the lack of growth phenotype to the fact that a small fraction of the target proteins escaped RNAi, but alternatively it is possible, that proteins might not be required for normal cell growth under tested conditions.

We then investigated how the interaction of ZapE with Oxa1 could be linked with other described roles of the former protein, which was associated with the FtsZ-dependent bacterial division, as both its overexpression and ablation resulted in elongated bacterial cells (Marteyn *et al*., 2015). We explored a possible co-occurrence of ZapE with the FtsZ machinery, but none was found, so it is likely that ZapE participates in different processes. In phylogenetic trees, eukaryotic ZapE homologs form a sister group to α-proteobacterial ZapE, implying their mitochondrial origin, a conclusion further supported by the absence of ZapE in plastids and cyanobacteria.

Additionally, we tested the suitability of BioID2 for the mapping of the mitochondrial proteins. We identified 117 proteins, all located to the organelle with high confidence. Thus, the BioID2 technique produces a very clean, albeit incomplete *T. brucei* mitochondrial proteome, which is estimated to contain ~1,200 proteins (Peikert *et al.* 2014; Dean *et al.*, 2015).

In second chapter of this thesis, we show that at least four distantly related eukaryotic lineages (Heterolobosea, Hemimastigophora, Alveida, and *Goniomonas* spp.) contains homologs of the bacterial Ffh and FtsY proteins that are unrelated to the previously known cyanobacteria-derived cpSRP54 and cpFtsY functioning in the plastids. Two lines of evidence (proteomic data and expression in a heterologous system) conclusively demonstrate that the respective proteins from *N. gruberi* function within the mitochondrion. Considering additional bioinformatic evidence for the mitochondrial localization of their homologs in other eukaryotes, we labelled these proteins as mtFfh and mtFtsY. Phylogenetic analyses indicate their common origin and are consistent with the vertical inheritance of the gene pair from a common ancestor of the respective eukaryotic lineages. Meanwhile, it is noteworthy to compare the evolutionary patterns of the SRP system in the plastids and mitochondria. Except for euglenophytes, the plastidial system is ubiquitous (Záhonová *et a*l., 2018), attesting to its tight integration into the molecular fabric of this cyanobacterium-derived organelle. In contrast, the mitochondrial version has been dispensed with on multiple occasions. Moreover, in some taxa, the plastidial system retains its RNA component (Träger *et al*., 2012), whereas the available evidence suggests that the corresponding SRP RNA had most likely been present in the alphaproteobacterial ancestor of the mitochondrion yet was lost prior to the LECA.

Nevertheless, the existence of a eukaryote-specific component of the mitochondrial SRP system remains an open possibility that needs to be addressed by more direct approaches. Since none of the eukaryotes carrying the mitochondrial SRP system is presently amenable to genetic manipulations, it is difficult to address its composition and function by experimental approaches. Assuming functional conservation dating back to bacterial ancestors of the mitochondrion as the most parsimonious alternative, the dissected system is involved in cotranslational membrane protein targeting. Hence, we evaluated the ability of the N-terminal sequences of the mitochondrial-encoded *N. gruberi* proteins that bear characteristics of a signal peptide to navigate a fused reporter fluorescent protein into the ER of genetically tractable *T. brucei.*

Furthermore, we identified MAP67, a novel mitochondrial protein that also occurs broadly in eukaryotes and was most likely already present in the LECA. Its obvious

evolutionary relationship to Ffh raises the possibility that it is a highly modified ortholog of mtFfh. However, the four mtFfh bearing lineages are interspersed among taxa with MAP67 and at least one of them, the genus *Goniomonas*, bearing both genes. Therefore, we propose that MAP67 and mtFfh coexisted in early eukaryotes and their current distribution reflects extensive differential loss. Unfortunately, MAP67 is not sufficiently similar to the M domains of Ffh/SRP54 to make a conventional phylogenetic analysis meaningful, but the most parsimonious explanation of its origin is that it emerged from a duplicated copy of mtFfh by an internal deletion that removed its N and G domains. Presently, we can only speculate about the function of MAP67, but it has already been shown to be essential in two model apicomplexans, *T. gondii* (TGGT1_254230; Sidik *et al.*, 2016) and *Plasmodium falciparum* (PF3D7_1004900).

The third chapter dealt with the characterization of TbPam16 and TbPam18 proteins that are bona fide orthologs of the PAM subunits Pam18 and Pam16 in other eukaryotes. Unexpectedly, in *T. brucei* they are neither involved in mitochondrial protein import nor part of PAM. In order to study these two proteins, we confirmed their mitochondrial localisation, that was suggest in TrypTag *in-situ* tagging database (Dean *et al.*, 2015). Depletion of TbPam18 and TbPam16 leads to downregulation of levels of MRP and components of OXPHOS complexes I, III, IV and V. Interestingly, ablation of TbPam18 and TbPam16 also leads to upregulated steady-state levels of DNA topoisomerase IB subunits and UMSBP1. Topoisomerase IB, is associated with the nuclear and mitochondrial genome replication, yet its exact role remains unclear (Jensen and Englund, 2012; Bakshi and Shapiro, 2004; Bodley *et al.*, 2003). During replication, the minicircles are disconnected from the kDNA network (Jensen and Englund, 2012). This observation could explain the progressive shrinkage of the kDNA disk and also why the total number of minicircles, as measured by Southern blots, remains stable several days after the loss of maxicircles. The presence of multinucleated cells 0K2N, 0K3N suggested a defect in the cell cycle and cytokinesis.

The function of trypanosomal Pam18 in the PAM was replaced by TbPam27 (von Känel *et al.* 2020). Here, we present data demonstrating that instead of being lost, TbPam18 and TbPam16 underwent architectural and sequence-specific alterations. Interestingly, the two proteins are still interaction partners, and their stability is interdependent (von Känel *et al.* 2020). However, they are involved in a completely unrelated function, the maxicircle replication or maintenance. Therefore, we link this function, for the first time, to integral IM proteins. Our findings once again underline that the presence of an orthologue does not proof

that its function is also conserved. While we can only speculate at this point, what the exact function of TbPam18 and TbPam16 might be.

While the signal for TbPam16 was comparable in uninduced and RNAi-induced cells (Fig. 3B), the effect of the TbPam18 RNAi on the COI and Cyb synthesis was visible already at day 2 post-RNAi induction, and the difference was well pronounced at day 4 p.i (Fig. 12A). These gels were also stained with Coomassie brilliant blue as loading control (insets, Fig. 12A–B).*T. brucei* arguably possesses the most well-investigated protist mitochondrion, yet despite multiple previous proteomics surveys, our genome-wide protein tagging analysis identified 337 novel mitoproteins, in addition to a strong agreement with previous proteomic analyses (Panigrahi *et al.*, 2009; Peikert *et al.*, 2017; Horn, 2021). Localization by protein tagging implicitly relies on the behaviour of a mutant protein and as such this may not always reproduce wild-type localizations. However, aberrations can equally be informative: differences between N- and C-terminal tagging allowed global mapping of the position of targeting sequences, and the KPE artefact stemming from a globular tag allowed global and highly accurate mapping of mitochondrial sub-compartments (Pyrih *et al.*, under review).

The theme of the fourth chapter was: Novel protein complex involved in kinetoplast DNA replication and maintenance. The principle finding of the MitoTag study was the unexpected demonstration that an mNeonGreen tag (along with other large globular epitopes) produced an artificial localization proximal to the kinetoplast in mitochondrial proteins, which were typically found in either the matrix or integral inner membrane (Pyrih *et al.*, under review). Amongst a plethora of artificially sub-localized mitochondrial proteins identified from MitoTag, several novel proteins were also identified exhibiting genuine kinetoplast signals, that appeared distinct from an artificial signal pattern. To remove ambiguity for these novel kinetoplast candidates, it was viewed as a necessary follow-up procedure to employ a smaller V5 tag to redetermine localization, which specifically does not induce such artificial interactions, as a measure of validation for identified kinetoplast candidates (Pyrih *et al.*, under review). The localization of all four chosen candidates to the kinetoplast via V5 tagging thus demonstrates these proteins as new components of the kinetoplast.

The demonstration of growth phenotypes in several RNAi cell lines generated from these novel proteins also demonstrates that they are not spurious (useless) components of the kinetoplast but are in fact important for the viability of the cell. For future studies, it will be useful to determine the function of these proteins beyond the growth phenotypes they induce. We recommend first investigating the status of kinetoplast DNA upon RNAi induction. The

measurement of certain mitochondrial metabolites may additionally hold promise in elucidating protein function considering the new metabolic potential within the kinetoplast (Pyrih *et al*., under review).

References

Bakshi, R.P., Shapiro, T.A. (2004) RNA interference of *Trypanosoma brucei* topoisomerase IB: both subunits are essential. *Mol Biochem Parasitol.* **136**: 249–255.

Bodley A.L., Chakraborty A.K., Xie S., Burri C., Shapiro, T.A. (2003) An unusual type IB topoisomerase from African trypanosomes. *PNAS.* **100**: 7539–7544.

Cesnekova J., Rodinova M., Hansikova H., Houstek J., Zeman J., Stiburek L. (2016a) The mammalian homologue of yeast Afg1 ATPase (lactation elevated 1) mediates degradation of nuclear-encoded complex IV subunits. *Biochem J.* **473**: 797–804.

Cesnekova J., Spacilova J., Hansikova H., Houstek J., Zeman J., Stiburek L. (2016b) LACE1 interacts with p53 and mediates its mitochondrial translocation and apoptosis. *Oncotarge*t. **7**: 47687–47698.

Dean S., Sunter J., Wheeler R.J., Hodkinson I., Gluenz E., Gull K. (2015) A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol.* **5**: 140197.

Germany E.M., Zahayko N., Huebsch M.L., Fox J.L., Prahlad V., Khalimonchuk O. (2018) The AAA ATPase Afg preserves mitochondrial fidelity and cellular health by maintaining mitochondrial matrix proteostasis. *J Cell Sci.* **131**: jcs219956.

Herrmann J.M., Bonnefoy N. (2004) Protein export across the inner membrane of mitochondria: The nature of translocated domains determines the dependence on the Oxa1 translocase. *J Biol Chem.* **279**: 2507–2512.

Hines J. C., Ray, D.S. (2010) A mitochondrial DNA primase is essential for cell growth and kinetoplast DNA replication in *Trypanosoma brucei*. *Mol Cell Biol.* **30**: 1319–1328.

Jensen, R.E., Englund, P.T. (2012) Network News: The replication of kinetoplast DNA. *Annu Rev Microbiol.* **66**: 473–491.

Kolli R., Soll J., Carrie C. (2018) Plant mitochondrial inner membrane protein insertion. *Int J Mol Sci.* **19**: 641

Marteyn B.S., Karimova G., Fenton A.K., Gazi A.D., West N., Touqui L., Prevost M.C., Betton J.M., Poyraz O., Ladant D., Gerdes K., Sansonetti P.J., Tang C.M. (2014) ZapE is a

novel cell division protein interacting with FtsZ and modulating the Z-ring dynamics. *MBio*. **5**: e00022-14.

Milman N., Motyka S.A., Englund P.T., Robinson D., Shlomai, J. (2007) Mitochondrial origin-binding protein UMSBP mediates DNA replication and segregation in trypanosomes. *PNAS*. **104**: 19250–19255.

Peikert C.D., Mani J., Morgenstern M., Käser S., Knapp B., Wenger C., Harsman A., Oeljeklaus S., Schneider A., Warscheid B. (2017) Charting organellar importomes by quantitative mass spectrometry. *Nat Commun*. **8**: 15272.

Sidik S.M., Huet D., Ganesan S.M., Huynh M.H., Wang T., Nasamu A.S., Thiru P., Saeij J.P.J., Carruthers V.B., Niles J.C., Lourido S. (2016) A Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes. *Cell*. **166**: 1423–1435.

Smith D.R., Keeling P.J. (2015) Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. *Proc Natl Acad Sci*. **112**: 10177–10184.

Träger C., Rosenblad M.A., Ziehe D., Garcia-Petit C., Schrader L., Kock K., Richter C.V., Klinkert B., Narberhaus F., Herrmann C., Hofmann E., Aronsson H., Schünemann D. (2012) Evolution from the prokaryotic to the higher plant chloroplast signal recognition particle: the signal recognition particle RNA is conserved in plastids of a wide range of photosynthetic organisms. *Plant Cell*. **24**: 4819–36.

Týč J., Klingbeil M.M., Lukeš J. (2015) Mitochondrial heat shock protein machinery hsp70/hsp40 is indispensable for proper mitochondrial DNA maintenance and replication. *mBio*. **6**: e02425–14.

von Känel C., Muñoz-Gómez S.A., Oeljeklaus S., Wenger Ch., Warscheid B., Wideman J.G., Harsman A., Schneider A. (2020) Homologue replacement in the import motor of the mitochondrial inner membrane of trypanosomes. *Elife*. **9**: e52560.

Záhonová K., Füssy Z., Birčák E., Novák Vanclová A.M.G., Klimeš V., Vesteg M., Krajčovič J., Oborník M., Eliáš M. (2018) Peculiar features of the plastids of the colourless alga *Euglena longa* and photosynthetic euglenophytes unveiled by transcriptome analyses. *Sci Rep*. **8**: 17012.

6. Curriculum vitae

Mgr. Vendula Rašková

CONTACT

☎ 00420 606 653 467

✉ raskovave@gmail.com

ⓘD 0000-0002-1746-1482

DATE OF BIRTH

17/August/1990

PLACE OF BIRTH

Ostrava, Czech Republic

NATIONALITY

Czech

LANGUAGE

English ★★★★☆

German ★★☆☆☆

CURRENT POSITION

PhD student & Research assistant

University of South Bohemia Faculty of Science & Biology Centre, Czech Academy of Sciences

OFFICE ADDRESS

Branišovská 1160/31
370 05 České Budějovice
Czech Republic

AREAS OF EXPERTISE

Molecular biology

Microscopy

Cell Culture

Parasitology

EDUCATION

2017–ongoing – **Ph.D. in Molecular and Cell Biology and Genetics**

Faculty of Science, University of South Bohemia, Czech Republic

Thesis topic: Characterisation of novel mitochondrial proteins in *T. brucei*

2013–2015 – **MSc. in Molecular and Cell Biology**

Faculty of Science, Palacký University Olomouc, Czech Republic

Thesis topic: Proteomics of cervical mucus

2010–2013 – **BSc. in Molecular and Cell Biology**

Faculty of Science, Palacký University Olomouc, Czech Republic

Thesis topic: Proteomic profile of methylprednisolone treated leukemia derived CEM cell line and comparison of protein changes with known effects of drug

WORK EXPERIENCE

06/2015–08/2016 - **Laboratory assistant**

Laboratory of Clinical Biochemistry and Hematology, SPADIA LAB a.s., (Nový Jičín, Czech Republic)

04/2017–07-2017 – **Research assistant**

Biomedical Research Centre, University Hospital in Hradec Králové (Hradec Králové, Czech Republic)

10/2017–present – **Research assistant**

Laboratory of Molecular Biology of Protists, Institute of Parasitology, Biology Centre, Czech Academy of Sciences (České Budějovice, Czech Republic)

03/2023-present – **Research specialist**

Pavel Krejčí Research Group, Department of Biology, Theoretical Departments, Faculty of Medicine, Masaryk University (Brno, Czech Republic)

RESEARCH STAY

Comenius University in Bratislava, Faculty of Natural Sciences, Department of Biochemistry, RNDr. Ingrid Škodová-Sveráková, Ph.D.; 3. 8. – 14. 8. 2020, 2. 8. – 20. 8. 2021. Biochemical analysis of the respiratory chain in selected strains of *T. brucei*.

REFERENCES

prof. RNDr. Julius Lukeš, CSc.

Head of Laboratory of Molecular Biology of Protists,

Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branišovská 1160/31 370 05 České Budějovice Czech Republic

📞 00420 387 775 416

✉ jula@paru.cas.cz

MSc. Ignacio Miguel Durante, Ph.D.

Postdoc in Laboratory of Molecular Biology of Protists,

Institute of Parasitology, Biology Centre, Czech Academy of Sciences, Branišovská 1160/31 370 05 České Budějovice Czech Republic

📞 00420 387 775 445

✉ ignacio.durante@paru.cas.cz

RNDr. Ingrid Sveráková, PhD

Assistant professor,

Department of Biochemistry, Faculty of Natural Sciences, Comenius University, Ilkovičova 6 842 15 Bratislava Slovakia

📞 00421 903 452 931

✉ skodovaister@gmail.com

PUBLICATIONS

Pyrih J.*, **Rašková V.***, Škodová-Sveráková I., Pánek T., Lukeš J. (2020) ZapE/Afg1 interacts with Oxa1 and its depletion causes a multifaceted phenotype. *PLoS ONE* 15: e0234918. [IF=3.752] DOI: 10.1371/journal.pone.0234918.   *Equal contribution

Durante I., Butenko A., **Rašková V**., Charyyeva A., Svobodová M., Yurchenko V., Hashimi H., Lukeš J. (2020) Large-scale phylogenetic analysis of trypanosomatid adenylate cyclases reveals associations with extracellular lifestyle and host-pathogen interplay. *Genome Biology and Evolution* 12: 2403–2416. [IF=3.416] DOI: 10.1093/gbe/evaa226.

Pyrih J., Pánek T., Durante I., **Rašková V**., Cimrhanzlová K., Kriegová E., Tsaousis A., Eliáš M., Lukeš J. (2021) Vestiges of the bacterial signal recognition particle-based protein targeting in mitochondria. *Molecular Biology and Evolution* 38: 3170–3187. [IF=8.800] DOI: 10.1093/molbev/msab090.

Von Känel C., Warscheid B., **Rašková V.**, Durante I., Škodová-Sveráková I., Oljeklaus S., Lukeš J., Schneider A. Orthologies of presequence translocase-associated motor subunits are essential for kinetoplast DNA replication in procyclic form of *Trypanosoma brucei* (in preparation).

Cadena L.R., Hammond M., Svobodová M., Benz C., **Rašková V**., Hashimi H., Durante I.M., Lukeš J. Novel protein complex involved in kinetoplast DNA replication and maintenance (in preparation).

TEACHING ACTIVITIES

Co-Supervisor: 2x BSc. thesis

Courses taught: Practice in Molecular Biology (2017), Methods of Molecular Biology – lectured in English (2018, 2019)

Science popularization:

Researcher´s Night in Biology Centre, CAS

Day of the open doors at the Institute of Parasitology, Biology Centre, CAS

Tutoring: Chemistry, Biology, Physics

INTERNATIONAL CONFERENCES

Raskova V., Skarka A., Schmidt M., Benek O., Hrabinova M., Musilek K. Significance and role of cyclophilin D in neurodegenerative diseases. EMBL Symposium: Mechanisms of Neurodegeneration 14.6–17.6. 2017, Heidelberg, Germany (poster).

Rašková V., Pyrih J., Lukeš J. A novel bacterial cell division protein ZapE and its role in the mitochondrion of *Trypanosoma brucei*. Jírovec´s Protozoological days, 30.4–4.5. 2018, Kostelec nad Černými lesy, Czech Republic (poster).

Rašková V., Pyrih J., Škodová-Sveráková I., Lukeš J. ZapE/Afg1 is Oxa1-dependent regulator of respiratory complexes in *Trypanosoma brucei.* Jírovec´s Protozoological days, 22.4–26.4. 2019, Kostelec nad Černými lesy, Czech Republic (oral presentation).

Rašková V., von Känel C., Durante I., Škodová-Sveráková I., Schneider A., Lukeš J. Characterisation of the Pam related proteins in *Trypanosoma brucei*. Jírovec´s Protozoological days, 28.6–2.7. 2021, Nové Hrady, Czech Republic (oral presentation).