High confidence glycosomal membrane protein inventory unveils trypanosomal Peroxin PEX15

Chethan K. Krishna[1*], Hirak Das[2*], Lisa Hohnen[1], Wolfgang Schliebs[1], Silke Oeljeklaus[2], Bettina Warscheid[2,3#], Vishal C. Kalel[1#], Ralf Erdmann[1#]

Author affiliations:

[1] Department of Systems Biochemistry, Institute of Biochemistry and Pathobiochemistry, Faculty of Medicine, Ruhr University Bochum, Bochum, Germany.

[2] Faculty of Chemistry and Pharmacy, Biochemistry II, Theodor Boveri-Institute, University of Würzburg, Würzburg, Germany.

[3] Signalling Research Centers BIOSS and CIBSS, University of Freiburg, Freiburg, Germany.

[*]These authors contributed equally

[#] **Correspondence:**

Vishal.kalel@ruhr-uni-bochum.de

Bettina.warscheid@uni-wuerzburg.de

Ralf.erdmann@ruhr-uni-bochum.de

Abstract

Trypanosomatids are a subgroup of kinetoplastids that share the features of harboring a kinetoplast and a flagellum. Infections by these parasites pose a major concern to human health, causing diseases like Chagas disease, Human African Trypanosomiasis (HAT), and Leishmaniasis, affecting over 12 million people worldwide. Trypanosomatid parasites possess a unique peroxisome-like organelle called Glycosomes, which are essential for parasite survival. Formation and function of glycosomes depend on the cytosolic receptors and various peroxisomal membrane proteins (PMPs). These PMPs either act as Peroxin (PEX) proteins in the glycosome biogenesis or play a crucial role in facilitating the movement of solutes/molecules across the glycosomal membrane. Given the significance of PMPs, an inventory of the glycosomal membrane proteins will provide insights into their functional importance, and also aid in the identification of unknown or parasite specific Peroxins. In this study, we describe the isolation of glycosomes, its membrane protein enrichment, and mass-spectrometry-based proteomic analysis followed by protein-correlation profiling to establish a high-confidence inventory of 57 glycosomal membrane proteins, with 44 additional putative candidates. Four novel proteins from this inventory were validated, which includes two tail-anchored (TA) proteins, a homolog of human PXMP4, and a Macrodomain containing protein. Using a structure-based approach, we identified that the novel TA protein is the long-sought *Trypanosoma* PEX15. Despite its low sequence similarity, it exhibits structural and topological similarities with its yeast (Pex15) and human counterpart (PEX26). We show that PEX15 is an integral membrane protein that localizes to the glycosome and interacts with PEX6. RNAi knockdown of PEX15 in bloodstream form (BSF) trypanosomes demonstrates that it is essential for glycosome biogenesis. Considering the low degree of conservation with its human counterpart, PEX15 is a promising molecular target for drug development.

Keywords

peroxisome, glycosome, Cell fractionation, mass spectrometry, peroxin, PMP, PEX15

1 Introduction

Kinetoplastids are a group of flagellated protists with a DNA-containing region called kinetoplast found within their single mitochondrion (Stuart et al. 2008). Trypanosomatids are a subgroup of kinetoplastids that share the features of having a kinetoplast and a flagellum (Kaufer et al. 2017). The insect-borne infections of trypanosomatid parasites cause different human diseases. These include Chagas disease, Human African Trypanosomiasis (HAT), and leishmaniasis, which are caused by *Trypanosoma cruzi*, *T. brucei*, and *Leishmania* species, respectively, posing a major concern for public health (Perez-Molina and Molina 2018, Kushwaha and Capalash 2022). The incidence of trypanosomiasis affects approximately 6-7 million people and a population of around 75 million at risk (Kourbeli et al. 2021, De Rycker et al. 2023). Leishmaniasis (all forms of diseases) exhibits a prevalence of over 6 million cases and over 600 million people at risk, with around 1 million new cases of cutaneous leishmaniasis and 50,000-90,000 cases of visceral leishmaniasis annually (Oryan and Akbari 2016, Kourbeli et al. 2021, De Rycker et al. 2023) (DNDi 2023, WHO 2023). Presently, treatment options are available for the early stages of HAT (Kande Betu Ku Mesu et al. 2021) and, to some extent for visceral leishmaniasis, with candidates in clinical trials (Nagle et al. 2020). However, treatment for Chagas disease and other forms of leishmaniasis still faces challenges due to significant side effects and toxicity (Perez-Molina and Molina 2018). There is a requirement to develop new drug targets to counter these infections.

Trypanosomatid parasites have specialized peroxisomes called glycosomes that compartmentalize glycolytic enzymes (Opperdoes and Borst 1977, Opperdoes et al. 1984). This compartmentation is unique because, in other organisms, glycolysis occurs completely in the cytosol. The compartmentation is crucial for the parasite survival because the mislocalization of glycolytic enzymes, which lack feedback is lethal for the parasites (Haanstra et al. 2014, Kalel et al. 2018). Glycosomes also contain enzymes for other pathways like gluconeogenesis (Morales et al. 2016), pentose phosphate pathway, pyrimidine biosynthesis, and purine salvage pathways (Michels and Gualdron-Lopez 2022). Depending on their life cycle stage, *Trypanosoma* also has enzymes for ether-lipid biosynthesis (Michels and Gualdron-Lopez 2022) inside glycosomes. The organelle's various functions make it essential for the different life cycle stages of the parasites. Like peroxisomes, glycosomes do not possess DNA, and hence their proteins are synthesized in the cytosol, and imported post-translationally into the organelles (Lazarow and Fujiki 1985, Michels et al. 2005). The biogenesis of the organelle depends on a machinery of proteins collectively referred to as Peroxins (PEX proteins). The Peroxisome/Glycosome biogenesis depends on two distinct machineries for the targeting of proteins. The first machinery is responsible for peroxisomal matrix protein import, while the second machinery is responsible for the formation of the peroxisomal membrane by targeting and inserting peroxisomal membrane proteins (PMPs) (Agrawal and Subramani 2016). The matrix proteins contain type 1 or type 2 peroxisomal targeting signals (PTS1/PTS2) at the extreme C-terminus or close to the N-terminus, respectively (Gould et al. 1989, Swinkels et al. 1991). PEX5 is the cytosolic receptor for PTS1-containing proteins and a co-receptor of PEX7, the receptor for PTS2-containing proteins (Galland et al. 2007). Peroxisomal membrane proteins (PMPs) are involved in the formation and function of glycosomes, and they play a

crucial role in facilitating the movement of solutes/molecules across the glycosomal membrane. Eleven of the 16 known parasite PEX proteins are PMPs (Gualdron-Lopez et al. 2013, Kalel et al. 2018). Knowledge on components of the glycosomal import machinery (peroxins) will be clinically important, particularly as novel drug targets. By interfering with these peroxins, the import of matrix proteins and in turn the biogenesis of glycosomes will be blocked. This will disrupt various glycosome-associated metabolic pathways, resulting in parasite death (Kalel et al. 2017, Li et al. 2021). Given the significance of PMPs as described above, an inventory of glycosomal membrane proteins will provide insights into their functional importance, translocation mechanisms, and role in organelle biogenesis.

In trypanosomes not all peroxins that are known in other species have been identified, like PEX15/PEX26. Yeast Pex15, Human PEX26, and Plant APEM9 are tail-anchored peroxisomal proteins that recruit the heterohexameric complex of AAA+ ATPases PEX1 and PEX6 from the cytosol to the peroxisomal membrane. This complex along with the RING finger peroxins (PEX2, 10 and 12) form an exportomer complex, which functions in receptor recycling (Birschmann et al. 2003, Matsumoto et al. 2003, Goto et al. 2011). Despite having a low sequence identity, AlphaFold structure predictions suggest that these anchor proteins share a common structure of a cytosolic helical bundle followed by a linker to the transmembrane domain at the C-terminal (Goto et al. 2011, Judy et al. 2022). Using AlphaFold2-multimer models it has been predicted that the N-terminal helical bundle of these anchor proteins binds to PEX6 across different organisms (Judy et al. 2022). Both AAA+ ATPases PEX1 and PEX6 have been identified in trypanosomes (Krazy and Michels 2006, Mahadevan et al. 2023). However, the protein anchoring the PEX1–PEX6 complex, the suspected PEX15/PEX26 at the glycosomal membrane, remains unknown.

In this study, we describe the isolation of glycosomes, enrichment of glycosomal membranes, and mass-spectrometry-based proteomic analysis followed by protein-correlation profiling to establish a high-confidence inventory of the glycosomal membrane proteome of *Trypanosoma brucei* (procyclic form). Four proteins were validated and their subcellular localization to the glycosomes demonstrated by immunofluorescence microscopy. We report on the identification of *Trypanosoma* PEX15, Peroxisomal membrane protein 4 (a homolog of *Hs*PXMP4), and Macrodomain 2 (PTS1 containing membrane protein). Using a structure-based approach, we identified *Trypanosoma* PEX15. Despite its low sequence similarity, it exhibits structural and topological similarities with its yeast (Pex15) and human counterparts (PEX26). We demonstrate that PEX15 is an integral membrane protein localized to the glycosome and confirmed its interaction with PEX6. RNAi knockdown of PEX15 in the bloodstream form (BSF) trypanosomes demonstrates its essential role in the glycosome biogenesis, further supporting the PEX15-PEX6 interaction as an attractive target for structure-based drug design.

2 Materials and methods

2.1 *Trypanosoma* cell culture and transfection

In this study, *Trypanosoma* procyclic form (PCF) 29-13 and bloodstream form (BSF) 90-13 cell lines, co-expressing T7 RNAP and TetR, were used. The BSF cells were cultured in HMI-

11 medium (Hirumi and Hirumi 1989) at 37 °C in a humidified incubator with 5% CO_2, while PCF cells were grown in SDM-79 medium (Brun and Schonenberger 1979) at 28 °C. Both cultures were supplemented with 10% fetal bovine serum (FBS). PCF cultures were maintained at 1×10^6– 30×10^6 cells/ml, while the BSF cultures were kept in the logarithmic growth phase (below 2×10^6 cells/ml). Transfections were performed with NotI-linearized plasmid constructs, which were genomically integrated into the rRNA locus within the genome of the respective cell lines. Clones were selected using the antibiotics as previously described in (Kalel et al. 2015). Transfection of PCF and BSF cells was performed as described previously in (Krazy and Michels 2006) and (Kalel et al. 2015), respectively. The antibiotic concentrations used for the selection of clones with respective plasmid constructs are as follows: PCF cells, pGN1/pGC1 (10 µg/ml Blasticidin); and BSF cells, pGN1/pGC1 (5 µg/ml Blasticidin) or p2T7-177 (2.5 µg/ml Phleomycin).

2.2 Cell lysis, subcellular fractionation, and glycosome enrichment

Procyclic *Trypanosoma* cells were harvested by centrifugation at 1500 g (Rotor SX4400, Beckman Coulter) for 10 minutes at 4 °C (≈1.9×10^{10} cells were used in this study) and were resuspended in the homogenization medium (250 mM sucrose, 25 mM Tris-HCl, 25 mM NaCl, 1 mM EDTA, pH 7.6, supplemented with 1X protease inhibitor cocktail i.e. PIC, Merck #4693132001). Further, the cells were ruptured with silicon carbide in the pre-cooled mortar, followed by the differential centrifugation of the homogenate which yields an organelle-enriched fraction (Krishna et al. 2023). This organelle-enriched fraction generally contains mitochondrial-derived vesicles, lysosomes, glycosomes, and endoplasmic reticulum (ER) vesicles. Density gradient centrifugation (20 – 46% OptiPrep^{TM}-Sucrose solution) was performed to isolate the enriched glycosomal fraction, which results in the separation of different cellular components according to their buoyant density. Following centrifugation, a clear separation of two bands in the gradient tube is visible, with the lower one indicating the enriched intact glycosomes, while the other upper band contains organelles such as mitochondria, ER, and others (Scheme shown in **Figure 1A**, **left panel**). The fractions of 0.5 mL each were collected by puncturing the tube at the bottom. Subsequently, the samples from each fraction were prepared and analyzed by SDS-PAGE, followed by immunoblotting. The glycosomal fractions 12-13 with a volume of 0.5 mL each (**Figure 1A**, **right panel**) were collected, concentrated by centrifugation at 20,800 g for 30 min at 4 °C (Rotor F45-30-11, Eppendorf) and the resulting pellet was then used for the enrichment of glycosomal membrane proteins. The detailed description of the protocol followed for isolating enriched glycosomes is described in (Krishna et al. 2023).

2.3 Enrichment of glycosomal membrane proteins

The glycosomal pellet was resuspended in ~350 µL of low-salt buffer (5 mM Tris-HCl, pH 7.6, 1 mM EDTA, and 1 × PIC) and was incubated for 30 min at 4 °C. Upon centrifugation for 30 min at 4 °C (96,000 g, Rotor TLA-100, Beckman Coulter), the resulting supernatant yields soluble glycosomal matrix proteins while the pellet contains membrane-associated proteins. Further, the pellet was resuspended in ~300 µL of alkaline carbonate buffer (100 mM Na_2CO_3, pH 11.5) and incubated for 60 min at 4 °C. Following the incubation, the suspension was layered on a ~120 µL of cushion (250 mM sucrose, 100 mM Na_2CO_3, pH 11.5) and was

centrifuged for 60 min at 4 °C (150,000 g, Rotor TLA-100, Beckman Coulter). The resulting supernatant contains matrix proteins and peripheral membrane proteins, while the carbonate-resistant pellet is the sediment of integral membrane proteins. The treatment with the carbonate buffer was repeated to reduce the contamination of matrix proteins. After each centrifugation step, supernatant and pellet samples were collected for immunoblot and mass spectrometry analysis. The detailed protocol followed for the enrichment of glycosomal membrane protein is described in (Krishna et al. 2023).

2.4 Membrane protein enrichment of the density gradient fractions for profiling

Another set of samples (Procyclic *Trypanosoma* cells) were processed for subcellular fractionation as described in section **2.2.** Following ultracentrifugation, the fractions of 0.5 mL each were collected by puncturing the tube at the bottom. Subsequently, fractions numbered from 9 to 22 were processed by centrifugation (20,800 g for 30 min at 4 °C, Rotor F45-30-11, Eppendorf), and the resulting pellet was further resuspended in 160 µL of alkaline carbonate buffer (100 mM Na_2CO_3, pH 11.5). Following the incubation for 60 min at 4 °C, the samples were layered on a 60 µL of cushion (250 mM sucrose, 100 mM Na_2CO_3, pH 11.5) and were centrifuged for 60 min at 4 °C (150,000 g, Rotor TLA-100, Beckman Coulter). The yielded pellet was analyzed by immunoblotting and mass spectrometry (**Supplementary Figure 1**).

2.5 Proteolytic in-solution digestion

Protein concentrations in the samples were initially determined using a bicinchoninic acid (BCA) assay kit. Approximately 20 µg of each sample underwent acetone precipitation via incubation in 80% acetone at -20 °C overnight. The samples were centrifuged at 12,000 × g for 10 min to facilitate precipitation, followed by air drying. The dried pellet was resuspended in a solution comprising 8 M urea and 50 mM ammonium bicarbonate (ABC). Cysteine residues were both reduced and alkylated by sequential incubation steps. First, in 5 mM Tris (2-carboxyethyl) phosphine (TCEP) dissolved in 10 mM ABC at 37 °C for 30 min, and then in 55 mM iodoacetamide mixed with 10 mM ABC for 30 min at room temperature, shielded from light. Alkylation was terminated by adding dithiothreitol (DTT) to achieve a final concentration of 25 mM. The samples were diluted to 4 M urea using 50 mM ABC and digested with LysC at 37 °C for 4 h for enzymatic digestion. Further dilution to 1 M urea was performed prior to overnight digestion with trypsin at 37 °C. Proteolytic reactions were arrested by acidifying the samples with trifluoroacetic acid (TFA) to a final concentration of 1% (v/v).

2.6 Liquid chromatography-mass spectrometry (LC-MS) and data analysis

For the LC-MS analysis, nano high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry (Nano-HPLC-ESI-MS/MS) was conducted using an Orbitrap Elite hybrid mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) in conjunction with UltiMate 3000 RSLCnano HPLC systems (Thermo Fisher Scientific, Dreieich, Germany). The RSLC system was outfitted with a C18 pre-column (nanoEase M/Z Symmetry C18; 20 mm length, 0.18 mm inner diameter) and an analytical C18 reversed-phase nano LC column (nanoEase M/Z HSS C18 T3; 250 mm length, 75 mm inner diameter, 1.8 mm particle size, 100 Å packing density). A binary solvent system was employed for peptide separation, composed of 4% (v/v) dimethyl sulfoxide (DMSO) and 0.1% (v/v) formic acid as

solvent A, and a mixture of 30% (v/v) acetonitrile (ACN), 48% (v/v) methanol, 4% (v/v) DMSO, and 0.1% (v/v) FA as solvent B. Peptide mixtures equivalent to 1 µg of protein were loaded onto the pre-column using solvent A, followed by a 5-min wash and pre-concentration at a 10 µL/min flow rate. A gradient elution was then applied at a 300 nL/min flow rate: 1–7% solvent B over 5 minutes, 7–50% B across 245 min, a surge to 95% B in 85 min, and a hold at 95% B for an additional 5 min. Eluted peptides were directed to a fused silica emitter for electrospray ionization facilitated by a Nanospray Flex ion source with a DirectJunctionTM adaptor, applying a spray voltage of 1.8 kV and a capillary temperature of 200 °C. Data-dependent acquisition parameters were set as follows: MS precursor scans at m/z 370–1700 with a resolution of 120,000 (at m/z 400); automatic gain control (AGC) set at 1×10^6 ions; a maximum injection time (IT) of 200 ms; a TOP20 method for low-energy collision-induced dissociation of precursor ions with a normalized collision energy of 35%, activation q of 0.25, and an activation time of 10 ms; AGC for MS/MS scans set at 5×10^3 with a maximum IT of 150 ms; and a dynamic exclusion time of 45 seconds.

For the analysis of mass spectrometric data, the MaxQuant software package (version 2.3.1.0) (Cox and Mann 2008), with its integrated Andromeda search engine (Cox et al. 2011), was employed to facilitate protein identification and quantification. Raw mass spectrometric datasets were interrogated against two FASTA files encompassing the full protein sequences for *T. brucei* strains Lister 427 (TREU427) and 927 (TREU927), obtained from the TriTrypDB database (version 8.1) to retain protein identification numbers of both strains in the dataset. Protein identification and subsequent quantification relied on the detection of one or more unique peptides. ‘Trypsin/P’ and ‘LysC/P’ were designated as the proteolytic enzymes, permitting a maximum of three missed cleavages. The 'iBAQ' parameter was activated for protein quantification. In terms of other variable settings, MaxQuant defaults were retained, which included the fixed modification of carbamidomethylation for cysteine residues and variable modifications involving N-terminal acetylation and methionine oxidation.

2.7 Cloning: Yeast and *Trypanosoma* expression plasmid constructs and cloning strategies are listed in **Table 1**, and oligonucleotide sequences are listed in **Table 2**. GFP-*Tb*PEX15 was cloned by the FastCloning (In-fusion cloning) method, as described in (Li et al. 2011). Mutations in *Tb*PEX15$_{1\text{-}360aa}$ and *Tb*PEX15 ΔTM$_{1\text{-}320aa}$ were generated by overlap extension PCR. Sequences of the constructs, mutations, and gene fragment deletions were verified for all constructs by automated Sanger sequencing.

2.8 Microscopy

Trypanosoma stable cell lines (PCF) encoding various tetracycline-inducible GFP tagged constructs [GFP, Peroxisomal membrane protein 4 (PMP4)-GFP, GFP-PEX15, GFP-hypothetical protein , and GFP-Macro domain 2 (MDo2)] were induced with 1 µg/mL tetracycline or treated with DMSO alone as negative control. Cells were sedimented and fixed by resuspension in 4% paraformaldehyde in PBS (phosphate-buffered saline, pH 7.4) at 4 °C for 15 min. Fixed *Trypanosoma* cells were processed for imaging as described previously (Krishna et al. 2023). The cells were stained with α-aldolase antibody (1:500) and α-GAPDH (1:500), glycosomal markers. Glycosome stained cells were visualized and imaged (five Z-

stacks) with a Carl Zeiss microscope using Zen 3.6 software (blue edition). All the images were processed by the deconvolution method and were merged. Further, the processed images were analyzed using Zeiss Zen 3.2 software (blue edition).

2.9 Cell fractionation

For the cell fractionation, *Trypanosoma* cells (PCF) encoding GFP tagged constructs (GFP, GFP-PTS1 (Dawidowski et al. 2017), PMP4-GFP, GFP-PEX15, GFP-HP, and GFP-MDo2) were expressed upon tetracycline induction (1 µg/mL). The cells with the expressed protein of interest were first treated with digitonin at a final concentration of 0.1 mg/mg of protein for 3-5 min at 37 °C. After this, the samples were centrifuged at 20,800 g for 15 min at 4 °C (Rotor F45-30-11, Eppendorf), yielding cytosol enriched sample and organellar pellet. The organellar pellet was further treated with alkaline carbonate buffer (100 mM Na_2CO_3, pH 11.5) and was incubated for 60 min at 4 °C. Following the incubation, the suspension was layered on a cushion (250 mM sucrose, 100 mM Na_2CO_3, pH 11.5) and is centrifuged for 60 min at 4 °C (150,000 g, Rotor TLA-100, Beckman Coulter). The resulting supernatant contains the released matrix proteins and peripheral membrane proteins, while the carbonate-resistant pellet is the sediment of integral membrane proteins. Sample collected from each step were analyzed with immunoblot.

2.10 Bioinformatics analysis and structure modelling

The protein sequences of Human PEX26, Yeast PEX15, and Plant APEM9, which are all orthologs of PEX15, were obtained from the UniProt database with IDs Q7Z412 (*H. sapiens*), Q08215 (*S. cerevisiae*), and Q8W4B2 (*Arabidopsis thaliana*). Additionally, the Human PXMP4 protein sequence was also obtained from the UniProt database (ID: Q9Y6I8). The parasite protein sequences of identified PEX15 were obtained from the TriTrypDB database with accession IDs Tb427_100021800 (*T. brucei*), TcCLB.508059.20 (*T. cruzi*), and LdBPK_210110.1 (*L. donovani*). In addition to this, the protein sequences of other characterized glycosomal membrane proteins in *T. brucei* include PMP4 (Peroxisomal membrane protein 4- Tb427_090008200), HP (Hypothetical protein- Tb427_010033000) and the MDO2 (Macrodomain 2- Tb427_110110000), all of which were retrieved from the TriTrypDB database. Transmembrane segments were predicted using the Phobius web tool (https://www.ebi.ac.uk/Tools/pfa/phobius/) (Madeira et al. 2022). The domain architecture of the proteins was predicted with the InterPro Scan (https://www.ebi.ac.uk/interpro/) (Paysan-Lafosse et al. 2023). Multiple sequence alignment was performed using the Clustal Omega tool (https://www.ebi.ac.uk/Tools/msa/clustalo/) (Madeira et al. 2022) and visualized using Jalview (v 2.1.1.0) software with a ClustalX color scheme. Percentage identity and similarity matrix of proteins were calculated using the SIAS tool with BLOSUM62 matrix (http://imed.med.ucm.es/Tools/sias.html).

The 3D structure of the proteins was obtained from the AlphaFold protein structure database (Jumper et al. 2021, Varadi et al. 2022). Since the structures of proteins from the *T. brucei* 427 strain were not available in the AlphaFold database, these proteins were modeled using the AlphaFold2 server (Jumper et al. 2021, Mirdita et al. 2022). The protein IDs modelled using the AlphaFold2 server are Tb427_100021800 and Tb427_010033000. Additionally, *Tb*PEX15

(Tb427_100021800) was also modeled using the Phyre2 tool (One 2 One threading) with the crystal structure of *Sc*PEX15 (PDB Id: 5VXV) as a template (http://www.sbg.bio.ic.ac.uk/phyre2) (Kelley et al. 2015, Brooke M. Gardner et al. 2018). The structures modeled using the AlphaFold2 server and Phyre2 tool were validated with the Ramachandran plots generated using the PROCHECK tool from the Saves server (https://saves.mbi.ucla.edu/) (Laskowski et al. 1993). All the structures were analyzed and visualized using UCSF Chimera software 1.17.3 (https://www.rbvi.ucsf.edu/chimera) (Pettersen et al. 2004).

2.11 Yeast two-hybrid analysis (Y2H)

S. cerevisiae wild-type strain PCY2 was grown in double dropout SD synthetic media (without tryptophan and leucine). Yeast cells were transformed by the traditional Lithium-acetate method (Gietz and Woods 2002). Y2H studies were based on the Yeast protocols handbook (Clontech, Protocol No. PT3024-1, Version No. PR742227). Full-length or truncation (ΔTM) of *Tb*PEX15 with and without mutations were cloned in the pPC97 vector containing GAL4-DNA Binding Domain (BD) as described in **Table 1**. Co-transformation of various two-hybrid plasmids, i.e., BD-*Tb*PEX15 and AD-*Tb*PEX6 (Mahadevan et al. 2023) constructs, were performed in the Wildtype PCY2 strain. The clones were selected on SD synthetic medium without tryptophan and leucine. A filter-based β-galactosidase assay and liquid culture assay using ONPG were performed in three replicates as described in the Yeast protocols handbook (Clontech). As a positive control for the study, the interaction between GAL4-AD and GAL4-BD fused to *Sc*PEX5 (Kerssen et al. 2006), and *Sc*PCS60 (Effelsberg et al. 2016) was used. The GAL4-AD fusion of *Tb*PEX6 and the various GAL4-BD fusions of *Tb*PEX15 (full-length and ΔTM) WT/mutants were also tested for autoactivation, serving as a negative control.

2.12 RNA interference and digitonin fractionation

The double-stranded (stem-loop) *Tb*PEX15 RNAi construct was cloned into a *Trypanosoma* expression plasmid p2T7-177 (tetracycline-inducible system), as described in **section 2.6**. It was then genomically integrated into the BSF *Trypanosoma* cells by transfection, as described in **section 2.1**. With the positive clones, a survival analysis was performed. RNAi was induced upon tetracycline (1 μg/ml) addition to the transfected *Trypanosoma* cells, cultured at 2×10^5 cells/mL density. DMSO served as a negative control for the study. Following the incubation for 24 h, the cell count was recorded, and the cells were diluted back to the 2×10^5 cells/mL with the addition of tetracycline. The culture with a density below 2×10^5 cells/mL was incubated further without the dilution and the addition of tetracycline (Kalel et al. 2019). The growth of the above-transfected cell line treated with DMSO, and tetracycline was monitored for four days. The experiment was performed with three biological replicates. The cumulative growth curve on a logarithmic scale was plotted using GraphPad Prism 10. However, the PEX15 mRNA or protein expression level reduction must be confirmed by Quantitative Real-time PCR (qRT-PCR) or immunoblot (ongoing). The DMSO and RNAi-induced cells were harvested, and the cell lysates corresponding to 2 million cells per well were analyzed by immunoblot to assess the level of PEX5 with day 1 and day 2 samples. In addition, the day 2 samples were also processed for immunofluorescence as described in **section 2.4** and

biochemical fractionation of cells using digitonin to assess the effect on glycosomal protein import.

Biochemical fractionation using digitonin was performed using the protocol adapted from (Chou et al. 2023). The harvested samples were resuspended in PBS (phosphate-buffered saline, pH 7.4) with 250 mM sucrose and 1x PIC (Roche). Protein concentration was estimated using the Bradford method with the day 2 harvested samples, i.e., DMSO-treated and RNAi-induced cells. Following the protein estimation, protein corresponding to ≈25 µg from the DMSO-treated and RNAi-induced samples were treated with increasing amounts of digitonin from 0.05 mg to 1 mg of digitonin/mg of protein (diluted in PBS with 250 mM sucrose). Cells treated with 1% Triton-X 100 represent the complete release of proteins by dissolving all membranes and serve as a positive control. After adding digitonin, the suspension was vortexed vigorously and incubated at 37 °C for 2-3 min with shaking (Thermo mixer-600 rpm). The incubated sample was centrifuged at 20,800 g for 15 min at 4 °C (Rotor F45-30-11, Eppendorf), and the resulting supernatant was analyzed by immunoblotting using various antibodies. The markers used are enolase (cytosolic marker); GAPDH, phosphofructokinase, glycerol-3-phosphate dehydrogenases (G3PDH), aldolase, hexokinase, GAPDH (glycosomal matrix markers); MtHSP70 (mitochondrial matrix marker)

2.13 Immunoblotting

The samples or cell lysate to be analyzed were denatured in Laemmli buffer for 5 min at 95 °C and separated by SDS-PAGE. After the gel electrophoresis, immunoblotting was performed as described in (Krishna et al. 2023). The following primary antibodies were used in this study: mouse anti-GFP (Sigma, 1:2,000), MtHSP70 (1:1,000); rabbit anti-*Trypanosoma* Aldolase (1:20,000), Enolase (1:20,000), Hexokinase (1:10,000), GAPDH (1:10,000), GIM5 (1:2,000), PEX11 (1:2,000), PFK (1:10,000), G3PDH (1:10,000), VDAC (1:10,000), BiP (1:5,000), PEX5 (1:20,000) and anti-Human alpha-tubulin (Abcam, 1:2,000) which shows cross-reactivity with *Trypanosoma* tubulin. All the above antibodies were prepared in PBS-T buffer (PBS, pH 7.4, 0.05% Tween-20) with 3% BSA except for VDAC, MtHSP70, and alpha-tubulin, which were prepared with 2.5% and 5% fat free milk, respectively. The corresponding secondary antibodies i.e., goat anti-rabbit IRDye 680 or goat anti-mouse IRDye 800CW (LI-COR Biosciences, 1:15,000) were used. The immunoblots were scanned using the LI-COR Odyssey Infrared Imaging System and analyzed with Image Studio version 5.2.

2.14 Quantification and statistical analysis

<u>Processing of membrane-enrichment sample data:</u> For the analysis of data from glycosomal membrane protein enrichment experiments, iBAQ intensities were used to measure proteins identified in different glycosomal fractions. These included the whole glycosomal fraction (Total), supernatants obtained after the low-salt buffer (SN0) and carbonate buffer treatments (SN1 and SN2), and the membrane-enriched pellet (MP). Proteins were included in the analysis if identified in at least two of the five samples per replicate. This criterion was met by 2851, 2844, and 2377 proteins in the first, second, and third replicates, respectively. To standardize the iBAQ intensities, each value was normalized by the maximum intensity observed for that

protein across all sample fractions. The 'IterativeImputer' function in Python was employed for instances where intensity values were missing (Pedregosa et al. 2011, Buitinck et al. 2013).

Processing of density-gradient membrane pellet data: Data from the membrane pellets of density-gradient centrifugation, spanning fractions #11 to #22, were analyzed using iBAQ intensities. Given that protein abundance can differ between various subcellular fractions, proteins were included in the analysis only if they were identified in at least two of the 12 density-gradient fractions. This criterion was met by 3932 proteins in the single available replicate. The iBAQ intensities were normalized, and missing values were imputed in the same manner as previously described. The processed data were then used for candidate validation.

Classification of proteins by protein correlation profiling (PCP): A curated list of marker proteins was initially compiled to serve as a reference set, each known for their distinct subcellular localization within *T. brucei* (**Supplementary Table S1**). These included proteins from the glycosomal membrane, glycosomal matrix, mitochondria, acidocalcisome, ER, cytoplasm, and other compartments. Utilizing this list, Pearson correlation coefficients (PCC) were computed between the mean iBAQ intensities of marker proteins localized in the glycosomal membrane and the iBAQ intensities of proteins throughout the dataset. The Pearson correlation was calculated with the 'pearsonr' function from the 'scipy.stats' package. Each protein trace was centered by subtracting the mean iBAQ intensity prior to computing the correlation. A composite similarity score was then derived according to the formula: Similarity Score = −log10 (Pearson's correlation × Magnitude Difference)
This procedure was performed for three independent biological replicates of glycosomal PMP enrichment, and the rank sum method (one-sided test) (Breitling and Herzyk 2005, Pedregosa et al. 2011, Buitinck et al. 2013), as implemented in the R package 'RankProd' (version 3.11) (Del Carratore et al. 2017), was applied to calculate p-values and mean correlation to the marker profiles. Proteins with >80% correlation to the mean profile of glycosomal membrane marker proteins were classified into two classes based on defined p-value cutoffs: Class 1 with p-value < 0.01 and Class 2 with p-value <0.05. For the single replicate of the membrane pellets from density-gradient fractions, similarity scores were calculated as described above and stored for each protein.

Definition of the glycosomal membrane protein inventory: To compile an inventory of glycosomal membrane proteins based on protein correlation profiling of samples in the membrane-enrichment steps as well as that of membrane pellets separated by density-gradient centrifugation, we proceeded with two-step selection criteria. In the first step, we categorize the proteins into two classes based on the significance threshold described above. In the second step, based on the similarity scores from membrane pellets from density-gradient fractions, we selected a cutoff of >80% similarity with the profile of glycosomal membrane marker proteins. Since only a certain percentage of a multi-localized protein may be present in glycosomes, and the majority may be present in the shared 'partner' organelle, we emphasize to visually inspect the profile of any candidate protein to see if there is a small peak in the glycosomal peak fraction, especially if the candidate protein may be multi-localized.

Microscopy: Microscopy images were quantified using Pearson's correlation coefficient, which was calculated with colocalization tool of Zen 3.6 pro (blue edition). Statistical significances were calculated using a one-way ANOVA (mixed) by Dunnett's multiple comparisons test with each row represents matched or repeated measures.

3 Results

3.1 Isolation of Glycosomes and enrichment of glycosomal membrane proteins

Several studies reported on the glycosomal proteome of *T. brucei*, *T. cruzi* and *Leishmania* species (Colasante et al. 2013, Guther et al. 2014, Jardim et al. 2018, Acosta et al. 2019). The total glycosomal proteome is dominated by highly abundant glycosomal matrix enzymes, particularly the glycolytic pathway. We adopted a systematic approach to define the glycosomal membrane proteome, with specific goal to identify low abundant PMPs, which may function as parasite specific peroxins, metabolite transporters, or proteins involved in membrane contact sites. To this end, we utilized procyclic form *Trypanosoma* cells to isolate intact and highly enriched glycosomes. The isolation method is based on the controlled rupture of the cells by silica carbide, followed by differential and density gradient centrifugation. Further, the glycosomal membrane proteins were enriched from the purified glycosomes by successive treatments with low-salt, and alkaline carbonate buffer extractions. The detailed protocol followed for isolating enriched glycosomes is described in (Krishna et al. 2023) and the scheme is outlined in **Figure 1A** (left panel). Following density gradient centrifugation, fractions were analyzed by immunoblot analysis as shown in **Figure 1A** (right panel). The results show that glycosomal matrix and membrane markers are strongly enriched in fractions 12 and 13, indicating the presence of intact glycosomes without the contamination of cytosolic proteins (Enolase). However, the glycosomal fractions contained minimal amounts of protein of mitochondrial (matrix and outer membrane), ER (matrix), and cytoskeleton (Acosta et al. 2019). Nonetheless, categorizing these proteins from other organelles in the enriched fraction of glycosomes as 'contaminants' needs further investigation, because previous studies have demonstrated the involvement of mitochondria and endoplasmic reticulum vesicles in the peroxisomal biogenesis in other organisms (Hettema et al. 2014, Kim 2017, Acosta et al. 2019). The markers of mitochondrion and ER were highly enriched in fraction no. 18 with very minimal amount of glycosomal proteins, possibly due to rupture of glycosomes during mechanical lysis with silica carbide. In addition, modest mitochondrial and ER markers, including tubulin, were present in fractions 14 to 17.

Further, the enriched glycosomal fractions (fractions 12 and 13) used for the preparation of glycosomal membrane proteins, as outlined in **Figure 1B** (left panel). To this end, the purified glycosomes were sequentially treated with low-salt buffer and carbonate buffer at pH 11.5, which resulted in the removal of glycosomal matrix proteins and proteins loosely associated to the membrane and yielding the carbonate-resistant pellet that contained integral membrane proteins (**Figure 1B**, middle panel). Further, the samples collected from each step of the enrichment process were evaluated with high-resolution liquid chromatography-tandem mass spectrometry (LC-MS) and advanced statistical data analysis was performed to identify

candidates (**Figure 1B**, right panel), which is described below (See **section 3.2**). Another set of samples was proceeded the same way as indicated in **Supplementary Figure 1**.

3.2 Proteomic Data Analysis: Defining the glycosomal membrane protein inventory using protein correlation profiling.

To systematically define the inventory of *T. brucei* glycosomal membrane proteins, we analyzed two datasets to identify and validate putative candidates (**Figure 2**). To account for the differential abundance of unique peptides in smaller proteins, especially membrane proteins of low abundance, we used intensity-based absolute quantification values (iBAQ-values) of proteins on all MS-based quantification (Schwanhausser et al. 2011, Tyanova et al. 2016). In the first dataset, we filtered the dataset to have iBAQ values in at least two samples across the membrane-enrichment steps (Total to MP) to calculate relative enrichment between them and identified more than 2800 protein groups. Next, we established a protein correlation profiling method (Andersen et al. 2003, Foster et al. 2006, Wiese et al. 2007) that can account for the change in the relative abundance of membrane proteins during each step of the membrane-enrichment protocol based on measured iBAQ values. To achieve this, we created manually a highly curated list of organelle marker proteins based on the available literatures (Wang et al. 2015, Huang et al. 2023) (**Supplementary Table 1**) and visualized their profiles based on iBAQ values individually to assess patterns arising from changes in abundance across each enrichment step (**Figure 2**). Based on the distinct profile of glycosomal membrane markers, we calculated a similarity score based on the Pearson correlation (dCorr) value of each protein in the dataset with that of the mean normalized iBAQ values of the glycosomal membrane markers. In each of the replicates, we identified 565, 437, and 321 values with more than 80% dCorr values to that of the mean glycosomal marker profile. We then performed a rank-product test to calculate the mean ratio and p-values (Breitling and Herzyk 2005). We classified the proteins into two classes based on defined p-value cutoffs: Class 1 with p-value < 0.01 and Class 2 with p-value <0.05. We identified 57 proteins in Class 1, which we term ‘high-confident’ glycosomal membrane proteins, including 44 putative candidates (**Table 3**), and 65 proteins in Class 2, which we term as ‘additional’ candidates that may be multi-localized proteins and require further investigation. Overall, we covered a total of approximately 4400 protein groups combined across all samples and identified 122 proteins in *T. brucei* to describe the comprehensive inventory of glycosomal membrane proteins.

3.3 Classification and validation of shortlisted PMPs from the inventory

The High-confident glycosomal membrane protein inventory was compared with the available glycosomal (Guther et al. 2014) and mitochondrial outer membrane proteomics (Niemann et al. 2013) to assess the inventory quality (**Supplementary Figure 2**). In this study, out of the 57 high-confident (class 1) glycosomal membrane proteins identified, 17 proteins were also observed in glycosomal proteome reported by Güther et al. as the high-confident glycosomal proteins, which include known PMPs like PEX10, PEX12, and GAT1 (Guther et al. 2014). Additionally, two proteins were found in the mitochondrial outer membrane proteomics study reported by Niemann et al (Niemann et al. 2013). This indicates the reliability and accuracy of the new inventory. These two proteins, reported in the above-mentioned proteomic studies and this study, are *Tb*Msp1 and putative 3-oxo-5-alpha-steroid 4-dehydrogenase. The *Tb*Msp1, the

second most significantly enriched glycosomal membrane marker protein in this study, which was recently identified and reported to localize to glycosomes and the mitochondrial outer membrane (Gerber et al. 2023). However, 3-oxo-5-alpha-steroid 4-dehydrogenase could be a putative dual localizing protein requiring further investigation. To further distinguish from the contaminant, the Rank-product volcano plots of the PMP inventory from this study were compared with the curated marker protein list (**Supplementary Figure 2B**).

From the class 1 high-confident potential glycosomal membrane proteins list based on the rank-product test (**Table 3**); we shortlisted four candidates for the preliminary characterization (**Figure 3A**). The shortlisted candidates were initially subjected to a BLAST search against yeast (*S. cerevisiae*), human, and plant (*A. thaliana*) protein databases, of which only one protein came up as the only hit against Human PXMP4 which will be referred to from here on as *Trypanosoma* peroxisomal membrane protein 4 (*Tb*PMP4; Tb427_090008200). In contrast, the other three proteins currently annotated as hypothetical proteins in Trypanosoma database (TriTrypDB; www.tritrypdb.org) are conserved in trypanosomatids only. Further, bioinformatic analysis was performed with these proteins, such as transmembrane domain (TMD) prediction using Phobius web tool and domain architecture prediction using InterPro Scan. The *Tb*PMP4 has four TMD with a conserved Tim17 superfamily domain and shows a sequence identity of 36.9% with *Hs*PXMP4 (**Supplementary Figure 3A**). One of the three hypothetical proteins has a Macrodomain with a single TMD close to the N-terminal region and will be termed as a Macrodomain 2 protein (MDO2). Previously, in both *T. brucei* and *T. cruzi*, a Macrodomain protein, i.e., MDO1, is identified and characterized (Haikarainen and Lehtio 2016). Both *T. brucei* MDO1 and MDO2 contain PTS1 signals and are also identified in high-confidence glycosomal proteome (Guther et al. 2014). However, only MDO2 contains a transmembrane domain, making it a second glycosomal membrane protein that contains PTS1 signal apart from PEX13.1 in *Trypanosoma* (Verplaetse et al. 2009). Further, the remaining two hypothetical proteins harbor a single TMD close to their C-terminus, i.e., tail-anchored (TA) proteins. These proteins do not possess any conserved domain, making them potential candidates for *Trypanosoma* PEX15. They are referred to as PEX15 candidate 1 and PEX15 candidate 2 at this stage, but after subsequent characterization they are termed as *Trypanosoma* PEX15 (this study) and Hypothetical Protein HP . The domain architecture of all the four shortlisted proteins is shown in **Figure 3B**.

Further, we performed immunofluorescence microscopy analysis to assess the localization of the GFP-tagged shortlisted proteins to the glycosomes of *T. brucei*. *Trypanosoma* stable cell lines (PCF) encoding various tetracycline-inducible GFP tagged constructs, i.e., PMP4-GFP, GFP-PEX15 candidate 1/candidate 2, and GFP- MDO2 were induced with 1 µg/mL tetracycline or treated with DMSO alone as negative control. Expression of tetracycline-inducible GFP-tagged constructs was confirmed immunoblotting. The analysis shows that the tagged proteins are running at the predicted molecular weight except the GFP-HP running at ~70 kDa (**Supplementary Figure 3B**). Glycosomal localization was assessed by the analysis of colocalization of the fluorescent GFP-fusion of the above constructs with the glycosomal marker enzyme aldolase, which was monitored by immunofluorescence microscopy. Overexpressed GFP-tagged PMP constructs (PMP4, PEX15 candidate 1/ candidate 2, i.e., HP,

and MDO2) colocalized with the glycosomal marker, implying that the new proteins are indeed glycosomal (**Figure 3C**). However, upon overexpression of the GFP-tagged PMP4 and the PEX15 candidate 1 and 2, the glycosomes appeared to cluster. Such a clustering has also been observed previously upon over-expression of other glycosomal membrane proteins, like PEX11 (Lorenz et al. 1998, Krishna et al. 2023) and PEX16 of *T. brucei* (Kalel et al. 2015). Pearson's correlation coefficients were used to quantify the colocalization, and the results indicate a significant correlation for GFP-tagged PMPs with aldolase (**Figure 3D**). Further, the membrane-pellet fractions were sequentially treated with digitonin and alkaline carbonate buffer to investigate whether the shortlisted proteins are indeed carbonate resistant integral membrane proteins. The resulting samples were subjected to immunoblot analysis using GFP, Enolase (cytosol marker), GIM5, and PEX11 (glycosomal membrane marker) antibodies. Enolase, the cytosolic marker was present in all supernatants following the digitonin treatment, indicating the complete release of cytosol fraction. The presence of GIM5 and the PEX11 protein in the carbonate-resistant pellet served as positive control for the preparation of integral membrane proteins. The GFP-tagged PMP constructs, i.e., PMP4, PEX15 candidate 1/ candidate 2, and MDO2, were predominantly present in the carbonate-resistant pellet, confirming them as integral membrane proteins (**Figure 3C**, right panel). Taken together, these results validate that the inventory's shortlisted PMP candidates are indeed glycosomal integral membrane proteins.

3.4 Identification of PEX15 in trypanosomatids

Most of the Peroxins or PEX proteins in trypanosomes were discovered by their sequence similarity with the known PEX proteins of other organisms. However, identification of *Trypanosoma* PEX15 by BLAST search failed, as the yeast PEX15 and its counterpart from other organisms, like Human PEX26 and Plant APEM9, themselves share very low amino acids sequence identity. But they do share a characteristic feature of transmembrane domain close to the C-terminus (Tail-anchored i.e. TA proteins) as well as high structural similarity of the cytosol facing portion (Goto et al. 2011). AlphaFold structure predictions suggest that the anchor proteins of the PEX1-PEX6 complex in Yeast, Human, and plants are similar in structure (Judy et al. 2022). With this as starting information, the structures of the two *Trypanosoma* PEX15 candidates were modeled using the AlphaFold2 server and compared with the Yeast PEX15 structure available in the AlphaFold protein structure database (**Figure 4A**). Of the two *Trypanosoma* PEX15 candidates, candidate 1 shows structural similarity with *Sc*PEX15 with a common N-terminal helical bundle known to interact with PEX6 across organisms. In contrast, candidate 2, with three alpha helices, lacks structural similarity compared to *Sc*PEX15 and was not considered further as the PEX15 candidate (**Figure 4A**). Hereafter, PEX15 candidate 1 will be referred to as *Trypanosoma* PEX15, while candidate 2 will be referred to as hypothetical protein .

Further, we performed multiple sequence alignments of *Trypanosoma* and Yeast PEX15 proteins using the Clustal Omega tool. As expected, the result revealed a low sequence identity of 11.94% and a similarity of 23.05% between the proteins (**Figure 4B**). In addition to the structural comparison of Yeast and *Trypanosoma* PEX15 using AlphaFold, a comparative modeling of *Trypanosoma* PEX15 was performed using the Phyre2 tool with the available

crystal structure of yeast PEX15 as a template (B. M. Gardner et al. 2018). The resulting model of *Trypanosoma* PEX15 is superimposed with the crystal structure of *Sc*PEX15 and revealed a high structural similarity (**Figure 4C**). The modelled structure of PEX15 using the AlphaFold2 server and Phyre2 were validated using the Ramachandran plot, where ≥98 % of the residues were in the allowed region, indicating the excellent quality of predicted structures (**Supplementary Figure 4**). Thus, the structure analysis by *de novo* as well as homology modeling strongly suggests that the identified protein is *Trypanosoma* PEX15.

Finally, the amino acid sequence of *Trypanosoma* PEX15 and its counterpart from yeast, human, and plant were retrieved and aligned. The *Trypanosoma* PEX15 showed a low sequence identity of less than 15.94% and less than 26% similarity between proteins compared to the proteins from these organisms (**Supplementary Figure 5 A-B**). Furthermore, we compared the 3D structure of *Trypanosoma* PEX15 with its counterparts from yeast, human, and plants (**Supplementary Figure 5 C-D**). In line with the literature (Judy et al. 2022), all the structures mentioned above share a feature of N-terminal helical bundle known to interact with PEX6 across different organisms, providing additional evidence for the identified *Trypanosoma* PEX15. The syntenic orthologs of *Tb*PEX15 are present in all *Trypanosoma* species. However due to the lack of synteny for *Leishmania* species, we identified *Leishmania* PEX15 ortholog using a BLAST search with the *Tb*PEX15 sequence. We retrieved the sequences of the *Tb*PEX15 and its counterparts in *T. cruzi* and *L. donovani* from TriTrypDB and aligned them. The *Tb*PEX15 protein shares 33.05% and 18.24% identity with its counterparts in *T. cruzi* and *L. donovani*, respectively, while the similarity is 43.61% and 28.07%, respectively (**Supplementary Figure 6A**). We also compared the predicted structure of *Trypanosoma* PEX15 with its counterparts in *T. cruzi* and *L. donovani* obtained from the TriTryp specific AlphaFold database (**Supplementary Figure 6B**) (Wheeler 2021). The predicted structures of PEX15 in *T. cruzi* and *L. donovani* feature a shared N-terminal helical bundle, as observed with *Sc*PEX15, *Hs*PEX26, and plant APEM9, facilitating the identification of a pan trypanosomatid PEX15.

3.5 Interaction between PEX15 and PEX6

PEX15 in yeast, PEX26 in human, and APEM9 in plants are anchor proteins for the AAA-ATPase PEX1-PEX6 complex in the respective organisms (Judy et al. 2022). To investigate the same for the identified *Trypanosoma* PEX15, we performed a Yeast two-hybrid assay (Y2H). To this end, *Tb*PEX15 constructs, i.e., full-length$_{1\text{-}360aa}$ and $\Delta TM_{1\text{-}320aa}$ (lacking transmembrane domain) fused to GAL4-BD were assessed for their interaction with full-length *Tb*PEX6 using Y2H (**Figure 5A**). *Tb*PEX15 full-length$_{1\text{-}360aa}$ showed a clear interaction with full-length *Tb*PEX6 in the plate-based assay (sensitive), whereas in the liquid assay (quantitative), the interaction appeared weak compared to the positive control. This could be due to the presence of the transmembrane domains in *Tb*PEX15, which might impede its translocation into the nucleus and subsequent activation of the GAL promoter. However, *Tb*PEX15 $\Delta TM_{1\text{-}320aa}$, i.e., the cytoplasmic part (lacking transmembrane domain) exhibited clear and strong interactions with full-length *Tb*PEX6, in both plate-based assay (sensitive) and liquid assay (quantitative). The negative controls with either of the fusion proteins (GAL4-AD/BD) did not result in auto-activation when tested with the respective GAL4-AD/BD alone domains.

Furthermore, we opted to introduce mutations in *Trypanosoma* PEX15 and assess its effect on the interaction with PEX6. In the yeast system, Birschmann et al. reported that an exchange of the Leucine at position 22 to Serine of *Sc*PEX15 leads to a loss of interaction with *Sc*PEX6 (Birschmann et al. 2003). Close inspection of the corresponding region revealed two leucine residues including L_{22} that are located in the first N-terminal helical region of *Sc*PEX15 (**Figure 5B, left panel**). Interestingly, we observed the presence of two consecutive leucine residues at positions 19 and 20 (19LL20) in the first N-terminal helical region of the *Trypanosoma* PEX15 structure (**Figure 5B, right panel**). Further, these residues were replaced with serine in *Tb*PEX15, and its effect on the interaction with *Tb*PEX6 was assessed using the Y2H assay. Mutation of leucine residues to serine i.e., 19LL20 to 19SS20 in *Tb*PEX15 (full length and ΔTM) led to the complete abolishment of the interaction with *Tb*PEX6 (**Figure 4C**). Similar to the wildtype *Tb*PEX15, the GAL4-BD fusions of mutated *Tb*PEX15 (full-length and ΔTM) did not show any autoactivation.

3.6 PEX15 is essential for trypanosomatid parasites, and its depletion affects glycosome biogenesis.

To investigate the essentiality of PEX15 for *T. brucei* parasite survival, we performed a knockdown experiment using RNA interference (RNAi). Stable cell lines were generated by transfection with a tetracycline-inducible double-stranded RNAi construct (stem-loop), which targets the coding region of PEX15, resulting in depletion of PEX15 expression. With the stable cell line, a growth analysis was performed as described in section 2.12 with tetracycline (1 µg/mL) and an equivalent volume of DMSO that served as a negative control for the study. The growth of the DMSO-treated or tetracycline-induced parasites was monitored for four days. The cumulative cell number was calculated from the cell counts recorded every 24 h. The growth analysis was plotted as cumulative growth curve on a logarithmic scale (**Figure 6A**). Growth defects were observed in the tetracycline-induced RNAi cell line from day 1 onwards, with a ~40% reduction in cell number compared to DMSO-treated cells. A severe growth defect of the RNAi-induced cultures was evident at day 2, where dead cell debris was clearly seen upon microscopic inspection. This demonstrates that depletion of PEX15 expression is lethal for the *Trypanosoma* parasites.

A common feature of the depletion of a peroxin in *T. brucei* is the disruption of glycosome biogenesis, leading to the mislocalization of glycosomal matrix enzymes to the cytosol, which in turn causes parasite death (Haanstra et al. 2008, Kalel et al. 2018). We investigated the mislocalization of glycosomal matrix proteins in the bloodstream form of parasites using microscopy and biochemical fractionation. As a clear growth defect was observed at day 2 of the *Tb*PEX15 RNAi knockdown, these cells were used for immunofluorescence microscopy analysis with antibodies against GAPDH and Aldolase, which are PTS1 and PTS2 containing glycosomal matrix markers, respectively. DMSO-treated cells served as control and show a bright punctate pattern for the glycosomal markers, indicating normal glycosome biogenesis without any protein import defect. In contrast, the tetracycline-induced RNAi cells show partial mislocalization of both GAPDH and Aldolase to the cytosol (**Figure 6B**, **left panel**). The RNAi-treated parasites still contain previously imported glycosomal enzymes, which may mask the diffuse cytosolic staining in microscopy images and hampers a quantification of the import

defect. Therefore, we performed a biochemical fractionation assay, which provides a quantitative method for assessing the mislocalization of glycosomal matrix proteins to the cytosol. To this end, DMSO-treated and tetracycline-induced PEX15 RNAi cells, were treated with increasing digitonin concentration on day 2. Following centrifugation, the supernatants were analyzed by immunoblotting using antibodies against the markers of cytosol, glycosomal matrix, glycosomal membrane, and mitochondrial matrix. Cells treated with 0.05 mg/mg of protein completely released cytosolic protein enolase, serving as a loading control for DMSO-treated and tetracycline-induced cells. Higher concentrations of digitonin (0.5 – 1 mg/mg protein) were required to release glycosomal proteins into the supernatant in the DMSO-treated cells (**Figure 6B**, **right panel**; labeled as -). Only a modest amount of glycosomal proteins, such as phosphofructokinase and hexokinase, were released with lower concentrations of digitonin in the DMSO-treated cells (Kalel et al. 2015). However, in PEX15 RNAi-induced cells, both PTS1-containing (GAPDH, Phosphofructokinase, and glycerol-3-phosphate dehydrogenase) and PTS2-containing (Aldolase, hexokinase) glycosomal matrix enzymes were well detected in the supernatants already at lower digitonin concentrations (0.05 – 0.5 mg/mg protein), confirming the mislocalization of glycosomal matrix proteins to cytosol upon PEX15 RNAi knockdown (**Figure 6B**, **right panel**).

Lastly, we assessed the impact of *Tb*PEX15 RNAi on the PTS1-containing protein receptor i.e., PEX5 recycling process, since PEX1-PEX6 complex anchored to the peroxisomal/glycosomal membrane is essential for this process. Total cell lysates from Day 1 and Day 2 of *Tb*PEX15 RNAi samples (both DMSO and Tetracycline-treated) were subjected to immunoblotting using antibodies against PEX5 and Enolase, a PTS1 cargo receptor and a cytosolic marker, respectively. The results revealed a significant reduction in the levels of PEX5 at both Day 1 and Day 2 samples compared to the DMSO-treated cells (without RNAi), confirming the instability of PEX5 in PEX15-depleted cells. This result is consistent with PEX5 instability in PEX1-depleted cells of *T. brucei* (Mahadevan et al. 2023).

4 Discussion

From inventory to validation: High confidence glycosomal membrane proteome.

In this study, we established a high confidence inventory of integral glycosomal membrane proteins of *Trypanosoma* parasites. To this end, we purified glycosomes and employed two distinct approaches. Purified glycosomes were subjected to carbonate extraction to enrich integral glycosomal membrane proteins or fractions obtained by density gradient fractionation of enriched organelles were extracted with carbonate. Proteomic analysis of samples from both approaches, and protein correlation profiling (Foster et al. 2006) with manually curated list of various *Trypanosoma* organelle markers resulted in the reliable identification of membrane proteins, which are exclusively localized to glycosomes or show a dual localization. Although some non-glycosomal proteins appear in the analysis, their classification as contaminants requires further investigation as peroxisomal proteins frequently show a dual localization (Acosta et al. 2019). Through these combined approaches, we identified 122 putative glycosomal membrane proteins in *T. brucei*. We classified the proteins into two classes: Class 1 (p-value < 0.01) containing 57 'high-confidence' glycosomal membrane proteins, and 65

proteins in Class 2 (p-value <0.05.), which we term as 'additional' candidates that may be multiorganelle-localized proteins and needs further investigation.

The high-confidence glycosomal membrane protein inventory was compared with the available glycosomal (Colasante et al. 2013, Guther et al. 2014) and mitochondrial outer membrane proteomics studies (Niemann et al. 2013). Out of the 57 high-confidence (class 1) glycosomal membrane proteins identified in this study, 17 proteins were previously reported in the glycosomal proteomics study by Güther et al. , including the known PMPs such as PEX10, PEX13, PEX12, and GAT1 (Guther et al. 2014). Furthermore, Colasante et al. reported 37 shortlisted glycosomal membrane proteins of *T. brucei* in their inventory, of which 31 were observed in our list (Colasante et al. 2013). Eight proteins in the Class 1 list, including known PMPs like PEX11, PEX16, and GIM5A/B, were observed, while the remaining 23 proteins were not observed in class 2, indicating that these proteins did not meet the stringent cutoff p-value criteria. Two known mitochondrial outer membrane proteomics were observed in our inventory (Niemann et al. 2013). One is a known dual-localizing protein, Msp1, which was recently shown to localizes to both glycosomes and the mitochondrial outer membrane (Gerber et al. 2023). The other, 3-oxo-5-alpha-steroid 4-dehydrogenase, has been reported to localize to mitochondria (Colasante et al. 2013), but it could also be a putative dual-localizing protein, requiring further investigation from the glycosomal perspective. Taken together, this suggests that the prepared inventory of high confidence glycosomal membrane proteins in *T. brucei* is very reliable, accurate, and comprehensive inventory, which will be a useful resource for understanding the biology of trypanosomatid glycosomes.

Based on the outcome of the rank-product test, we observed PEX11 and GIM5A, known glycosomal membrane proteins that exhibit the highest and lowest enrichment levels, respectively. Several uncharacterized proteins were identified, each with varying enrichment levels within this Class 1 list. Four proteins were shortlisted for further analysis, regardless of their ranks and enrichment levels (**as shown in Figure 3A**). These are PMP4 (homolog of *Hs*PXMP4), PEX15 (ortholog of *Sc*PEX15), Macrodomain 2 (MDO2) protein and the 'Hypothetical protein', an uncharacterized tail-anchored protein. Further analysis of the *Trypanosoma* PMP4 revealed that it as a homolog of the human peroxisomal membrane protein PXMP4 (also known as PMP24 or PMP4) for which the function is still not known (Reguenga et al. 1999). The *Tb*PMP4 was also reported in the PMP inventory of *T. brucei* (Colasante et al. 2013).

MDO2 is the second novel Macro Domain-like protein identified in trypanosomes. Previously, Macrodomain-like protein, i.e., MDO1, was identified and characterized as ADP-ribose hydrolase in *T. brucei* and *T. cruzi*, which are homologous to human O-acetyl-ADP-ribose deacetylases MacroD1 and MacroD2 (Haikarainen and Lehtio 2016). However, the study did not emphasize its subcellular localization. Interestingly, MDO1 protein contains a PTS1-like signal (-SSL) at its C-terminus and is also reported as a highly confident glycosomal protein (Guther et al. 2014). A distinctive feature of MDO2 is that it contains a predicted transmembrane domain in its N-terminal region, making it the second glycosomal membrane protein with a PTS1 signal in *Trypanosoma*, beside PEX13.1 (Verplaetse et al. 2009). However,

the MDO2 proteins are also reported in the Mitochondrial membrane proteome, suggesting this could be a potential dual localizing protein (Acestor et al. 2009). The MDO-like proteins need further investigation to understand their role in glycosomes. Also, understanding the significance of the PTS1 signal in glycosomal membrane proteins and checking its role in the import of glycosomal matrix proteins makes it an exciting area for further investigation.

The two hypothetical tail-anchored proteins with a characteristic transmembrane domain at their C-terminus were present in the Class 1 inventory, making them potential candidates for *Trypanosoma* PEX15. The bioinformatics analysis revealed that these proteins are conserved in trypanosomatids, and do not share any conserved domain with known organisms like human, yeast, and plants. Besides, both proteins were reported as high-confident glycosomal proteins but were not recognized as membrane protein (Guther et al. 2014). This study confirmed their localization to glycosomes and topology as an integral membrane protein (**Figure 3C**). In this study, the *Trypanosoma* PEX15 was identified and characterized, the other tail-anchored protein termed as a 'hypothetical protein' needs further characterization.

Long-sought *Trypanosoma* PEX15: A potential drug target against Trypanosomatids.

Peroxisomal function depends on the efficient import of proteins or enzymes into their matrix (reviewed in (Platta et al. 2013). Peroxisomal matrix proteins are targeted to peroxisomes through cytosolic receptors, which are recognized by the peroxisomal targeting signal (PTS1/PTS2) (Gould et al. 1989, Swinkels et al. 1991). These receptors, along with the cargo, are targeted to the docking complex at the peroxisomal membrane. The import process occurs through a transient pore or a hydrogel-filled pore in an unknown fashion (Erdmann and Schliebs 2005, Meinecke et al. 2016, Gao et al. 2022). Following the cargo release, the monoubiquitinated receptor is recycled back to the cytosol by the AAA-ATPase complex PEX1-PEX6 in an ATP-dependent manner (Platta et al. 2004, Platta et al. 2005) . In the cytosol, the receptor is made available for the next round of import , indicating that the receptor recycling is crucial for the import and biogenesis of peroxisomes. The AAA- ATPase complex PEX1-PEX6 requires an anchor protein at the peroxisomal membrane for extracting the monoubiquitinated receptor PEX5. In the absence of recycling, PEX5 accumulates at the membrane, leading to pexophagy (Nuttall et al. 2014). The known anchor proteins of other species are tail-anchored proteins that are poorly conserved and share low amino acid identity across organisms (Goto et al. 2011). The tail-anchored proteins recruiting the PEX1-PEX6 complex are designated as PEX15 in yeast, PEX26 in human, and APEM9 in plants (Birschmann et al. 2003, Matsumoto et al. 2003, Goto et al. 2011).

In this study, we aimed to identify *Trypanosoma* PEX15 using a structure-based approach because a BLAST search was not suitable, primarily due to the low amino acid sequence conservation. However, analysis of the predicted structures of these anchor proteins using the AlphaFold Protein Structure Database (Jumper et al. 2021, Varadi et al. 2022) suggests that the anchor proteins *Sc*PEX15, *Hs*PEX26, and APEM9 are structurally similar and share a common N-terminal helical bundle known to interact with PEX6 (reviewed in (Judy et al. 2022)). With the above evidence, the top two tail-anchored proteins from the High-confident glycosomal membrane proteins were examined for structural similarity, which resulted in a

potential *Trypanosoma* PEX15 candidate. Further, we also performed homology-based modeling of the potential PEX15 candidate using the available crystal structure of *Sc*PEX15 (B. M. Gardner et al. 2018) and confirmed the similarity by superimposing both structures (Figure 4). Furthermore, we confirmed the interaction of putative *Tb*PEX15 with *Tb*PEX6 by yeast two-hybrid assay. In addition, the Mutation of leucine residues (19th and 20th residues) to serine in the N-terminal region of *Tb*PEX15 loses interaction with *Tb*PEX6, as observed in the yeast system (Birschmann et al. 2003), which was determined again by the structural comparison (**Figure 5**). Functional characterization of PEX15 in trypanosomes revealed that the identified peroxin localizes to glycosomes, and its depletion leads to the mislocalization of glycosomal matrix proteins, causing cell death. These results are consistent with the previously reported peroxins of *T. brucei* and indicate the PEX15 involvement in glycosomal biogenesis (Kalel et al. 2015, Banerjee et al. 2019, Kalel et al. 2019, Mahadevan et al. 2023). In addition, the depletion of PEX15 leads to the degradation of PTS1 receptor PEX5 (Figure 6), which is also observed with *Tb*PEX1 depletion, a protein involved in receptor recycling (Mahadevan et al. 2023). Taken together, *Trypanosoma* PEX15 is essential for parasite survival, especially considering its low degree of conservation in amino acid sequence compared to its human counterpart, making it an attractive target for drug development.

Conclusion

This study provides a high-confidence inventory of glycosomal membrane proteins (PMPs) in *T. brucei*. Given the significance of PMPs in facilitating the transport of solutes and molecules across the glycosomal membrane, as well as in importing matrix proteins, it holds importance for glycosome function and biogenesis, which is essential for the survival of parasites. Targeting glycosomal membrane proteins/transporters would be more suitable and effective in disrupting their function, disrupting glycosome biogenesis and associated pathways, ultimately resulting in parasite death. In addition to the inventory, this study also identified a long-sought *Trypanosoma* PEX15, an anchoring protein involved in receptor cycling and an essential step in matrix protein import. Considering its low degree of conservation in amino acid sequence compared to its human counterpart, it is supposed to be an attractive target for drug development. Along this line, our Y2H mutational analysis of the PEX15-PEX6 interaction support the potential of the PEX15-PEX6 binding interface as an attractive target. These findings hold significant therapeutic implications, particularly in developing drugs targeting unique proteins like PEX15 for combating *Trypanosoma* parasites.

Acknowledgements

This project work has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 812968 (CK, HD, BW and RE). The work was supported by DFG grant FOR1905 to RE. We thank Paul Michels for kindly providing various antibodies and Lavanya Mahadevan for providing the PEX15 and HP cell lines. We would like to acknowledge the use of molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311 and other web tools such as Phyre2, Sequence Identity And Similarity (SIAS) tool.

Author contributions

CK, BW, VK, RE conceived and planned the experiments. CK, HD, LH performed all experiments. WS, SO, BW, VK, RE supervised the work. CK, HD and VK wrote the manuscript with support of RE and BW. All authors read and gave feedback on the manuscript.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Figures:

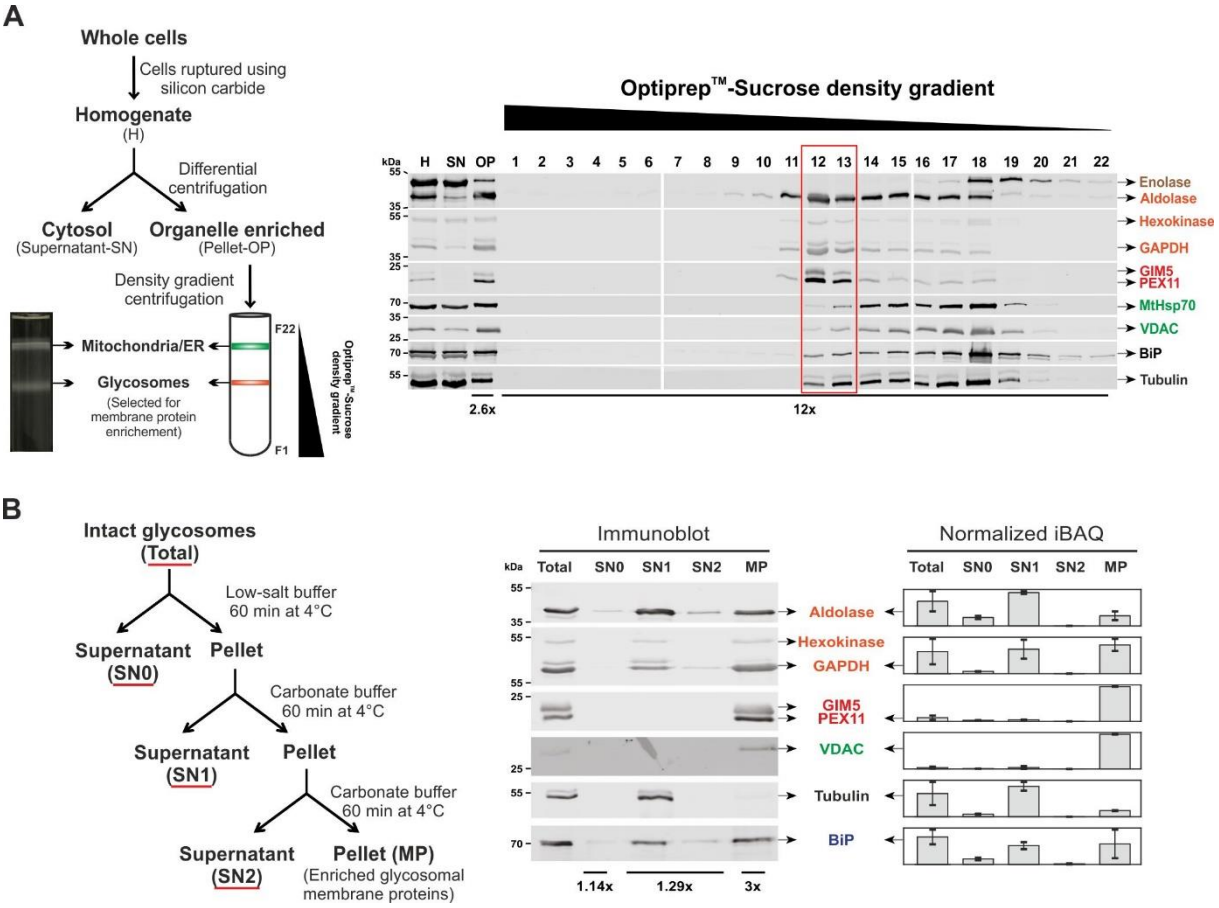


Figure 1: Glycosome isolation and membrane protein enrichment from *Trypanosoma brucei*. A) The left panel represents a schematic overview of *Trypanosoma brucei* glycosome isolation, along with an illustrative representation of band patterns in green and orange of the enriched fractions (F1-F22) of the density gradients. The glycosomes were isolated by density gradient centrifugation using OptiPrep concentrations ranging from 20% to 46% (w/v iodixanol). Fractions were collected from the bottom of the gradient, and immunoblot analysis was performed using various antibodies, as indicated. The red box highlights the enriched glycosomal fraction, showing minimal markers from mitochondria, the endoplasmic reticulum (ER), and cytoskeleton. PNS: Post-nuclear supernatant; SN: Supernatant representing cytosolic proteins; OP: Organellar pellet. Concentration factors relative to the PNS are depicted at the bottom (right panel). **B)** Scheme of the glycosomal membrane protein enrichment from the isolated intact glycosomes (left panel). Glycosomes were sequentially treated with low-salt buffer (5 mM Tris-HCl) and twice with alkaline carbonate buffer (100 mM Na_2CO_3, pH 11.5). After each centrifugation step, supernatant and pellet samples were collected for immunoblot analysis and iBAQ-based quantitative MS. Immunoblot results (middle panel) and normalized iBAQ intensities (right panel) are shown for the collected samples (indicated by the red line in the left panel) with specific subcellular marker proteins. Total: Enriched glycosomes; SN0: Supernatant post low-salt treatment; SN1: First supernatant post carbonate treatment; SN2: Second supernatant post carbonate treatment; MP: Membrane-enriched pellet; Norm iBAQ: Normalized intensity-based absolute quantification values. Error bars show standard deviations for n = 3. Concentration factors relative to the total glycosome are depicted at the bottom (middle panel). The markers are enolase (cytosolic marker in brown); aldolase, hexokinase, GAPDH (glycosomal matrix markers in orange); GIM5, PEX11 (glycosomal membrane markers in red); MtHSP70 (mitochondrial matrix marker in green); VDAC (mitochondrial membrane marker in green); tubulin (cytoskeletal marker in black) and BiP (ER marker in blue).

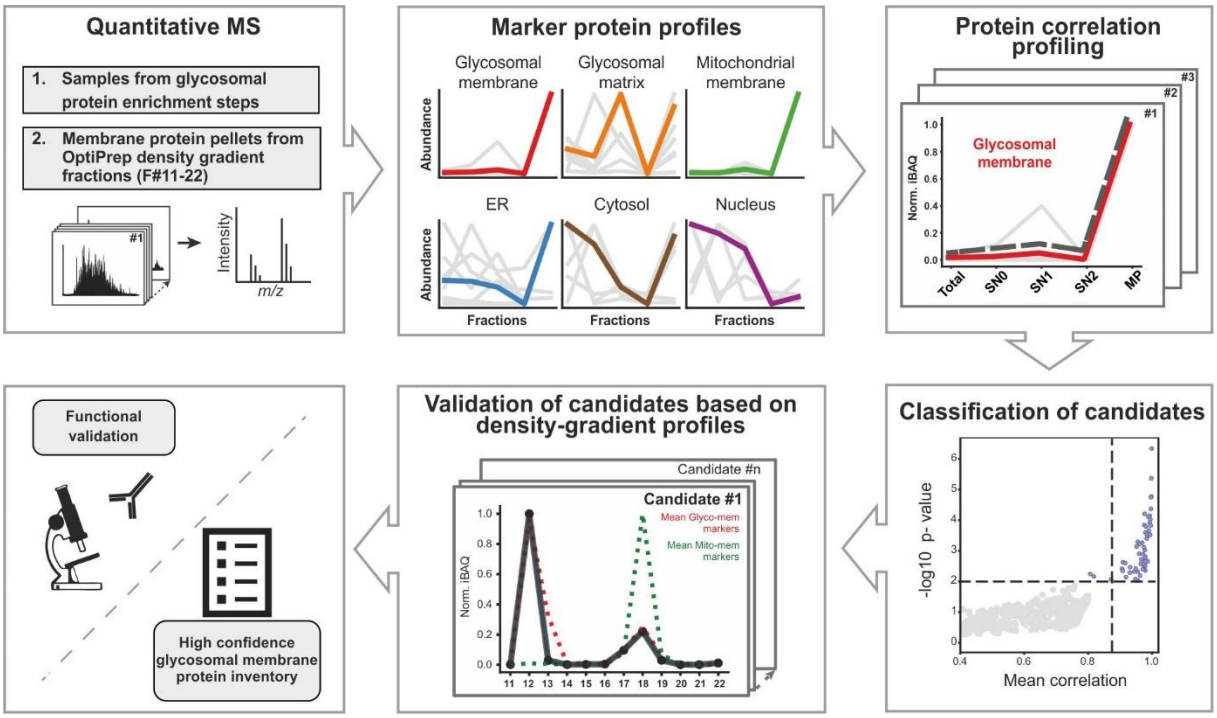


Figure 2: Establishment of high confidence glycosomal membrane protein inventory using a quantitative proteomics approach. A quantitative mass spectrometry approach was employed for two sample sets: one derived from glycosomal membrane protein enrichment steps (Fig. 1B), and the other from carbonate-resistant membrane protein pellets prepared for specific density-gradient fractions (F11-F22: Supplementary Figure 1). In the first set, protein correlation profiling (PCP) was conducted using normalized iBAQ intensities and a curated marker proteins list (glycosomal membrane in red, glycosomal matrix in orange, mitochondrial membrane in green, endoplasmic reticulum in blue, cytosol in brown, nucleus in purple). Based on the enrichment of proteins across selected fractions during the steps for preparing enriched glycosomal membrane proteins, Pearson correlation was calculated for the mean of glycosomal marker proteins for each replicate. The rank-product test was then utilized to compute the mean correlation to glycosomal markers along with p-values. For the second set, marker protein profiles were generated using normalized iBAQ intensities for PCP of glycosomal candidates, thereby defining the membrane protein inventory. Selected novel candidate glycosomal membrane proteins were validated through immunofluorescence microscopy, and their functional analysis was carried out.

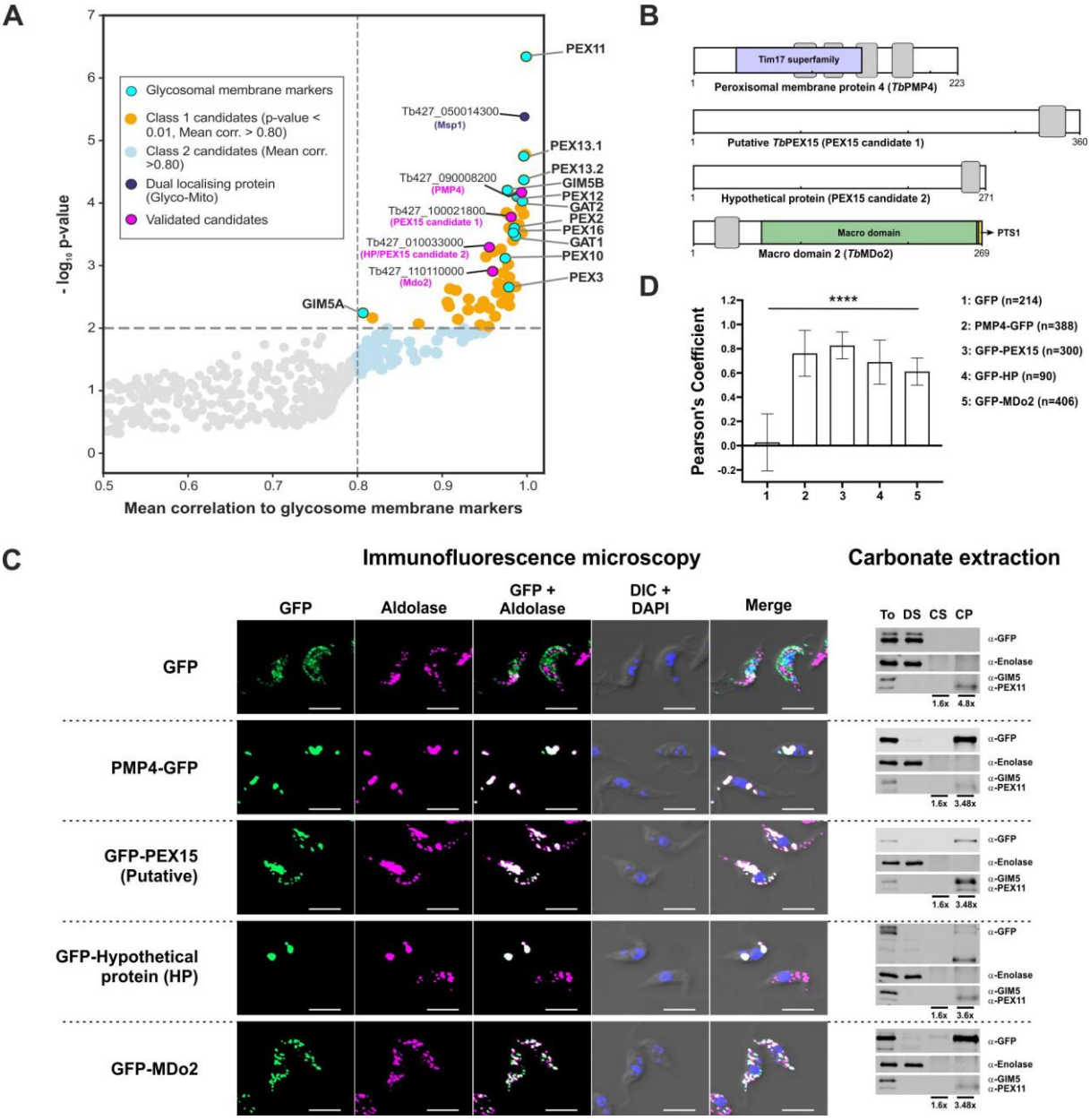


Figure 3: From inventory to validation: Identification of novel glycosomal membrane proteins. A) Identification of potential glycosomal(/peroxisomal) membrane proteins (PMP): In this analysis, a rank-product volcano plot was generated using a dataset that showcases the mean correlation of individual proteins with the known glycosomal marker proteins. The displayed proteins were quantified in a minimum of 2 out of 3 independent replicates, with p-values calculated via the rank sum method (Breitling and Herzyk 2005, Del Carratore et al. 2017). The annotation of the proteins shortlisted for the validation is based on their domain architecture. The horizontal dashed line denotes a false discovery rate of 1%, while the vertical dashed line indicates a mean correlation threshold (Mean corr.) of 80% with the glycosomal membrane markers. **B)** Schematic representation of the domain architecture of the putative *Tb*PMP4 (Tim17 superfamily domain is shown in blue), putative *Tb*PEX15, Hypothetical protein (uncharacterized glycosomal membrane protein) and *Tb*MDo2 with macrodomain (shown in green) with PTS1 signal represented in yellow. Transmembrane segments are predicted using Phobius webtool (https://www.ebi.ac.uk/Tools/pfa/phobius/). The domain architecture of the proteins was based on the InterPro Scan. **C)** Immunofluorescence super-resolution microscopy analysis of indicated GFP-tagged PMP constructs (Pmp4, putative PEX15, hypothetical protein, MDo2) show that they co-localize with the glycosomal marker aldolase (pseudocolored in magenta). DAPI stains nucleus and kinetoplast. In contrast, the negative control GFP is cytosolic, as evident from the overall diffuse cell labelling. The scale bar represents 2 µm. Same cell lines were subjected to detergent extraction with digitonin, and alkaline carbonate, which demonstrates that the GFP-tagged proteins are integral membrane proteins. For the corresponding immunoblots, shown on the right, fractions of indicated cells were analyzed using anti-GFP antibodies to detect the fusion proteins along with control proteins enolase, GIM5 and PEX11, cytosolic and glycosomal membrane proteins, respectively. To: total

lysate; DS: digitonin supernatant containing cytosolic proteins; CS: carbonate supernatant, containing organellar matrix and peripheral membrane proteins; and CP: carbonate pellet containing integral membrane proteins. Concentration factors relative to the total are depicted at the bottom. **D)** Quantification of the colocalization of GFP-tagged PMPs to glycosomes. The Pearson's coefficient of colocalization to respective organelle is shown. Error bars represent mean with standard deviations. Statistical significances were calculated by one-way ANOVA, with Dunnett's multiple comparisons test ($n \geq 90$ cells). ****, $p < 0.0001$.

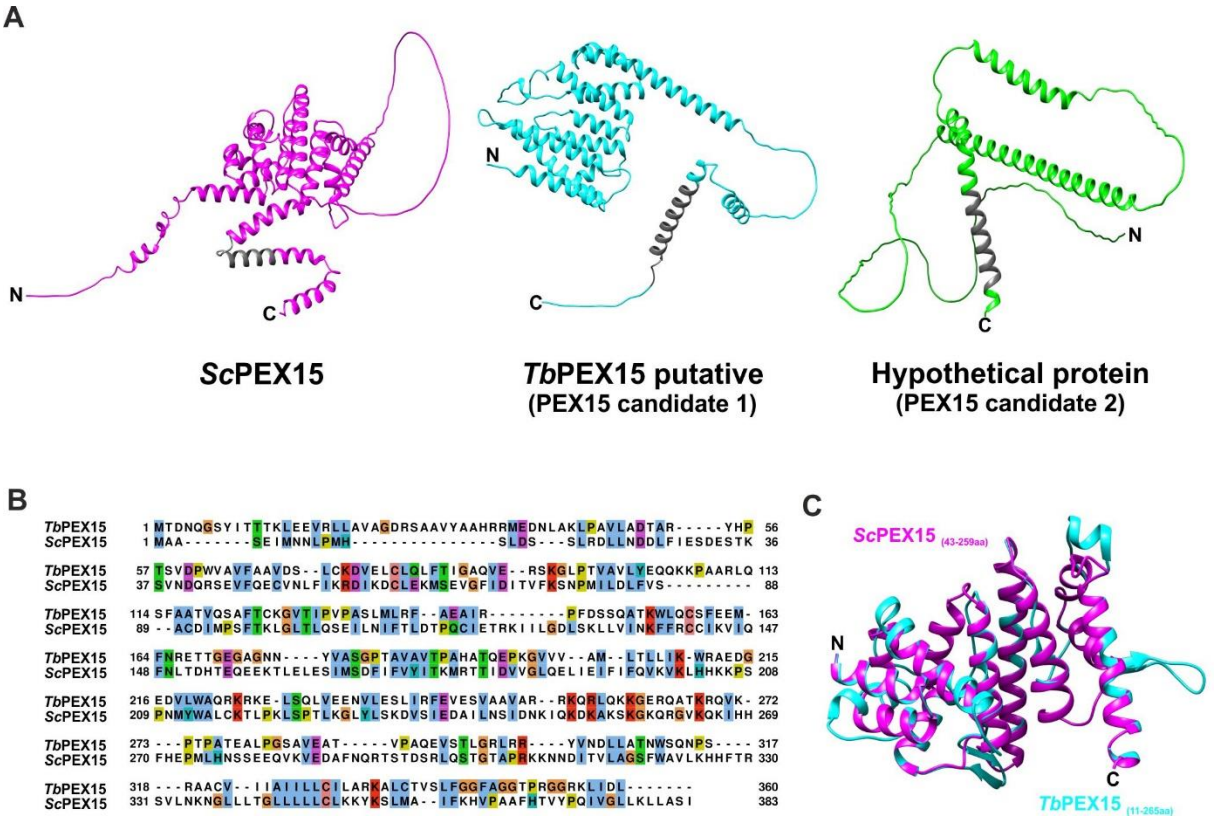


Figure 4: Identification of the trypanosomal PEX15 using a structural approach. A) AlphaFold predicted structures of yeast PEX15 (magenta, left panel), and *T. brucei* PEX15 candidates, of which candidate 1 (middle panel) shows high structural similarity with a common N-terminal helical bundle that for the yeast protein is known to interact with PEX6. A predicted C-terminal transmembrane domain that is characteristic for the tail-anchored PEX15 is marked in grey. **B)** Multiple sequence alignment of *Trypanosoma* (TriTrypDB ID: Tb427_100021800, for annotated Reference stain: Tb927.10.1860) and yeast PEX15 proteins, performed using the Clustal Omega tool, and visualized using JalView software with a ClustalX color scheme. The sequence identity and similarity of *Trypanosoma* putative PEX15 in comparison to yeast PEX15 is 11.94% and 23.05%, respectively. **C)** Structural comparison of the putative *Tb*PEX15 modelled using Phyre2 tool (represented in cyan) with the crystal structure of *Sc*PEX15 (PDB Id: 5VXV, represented in magenta) as a template.

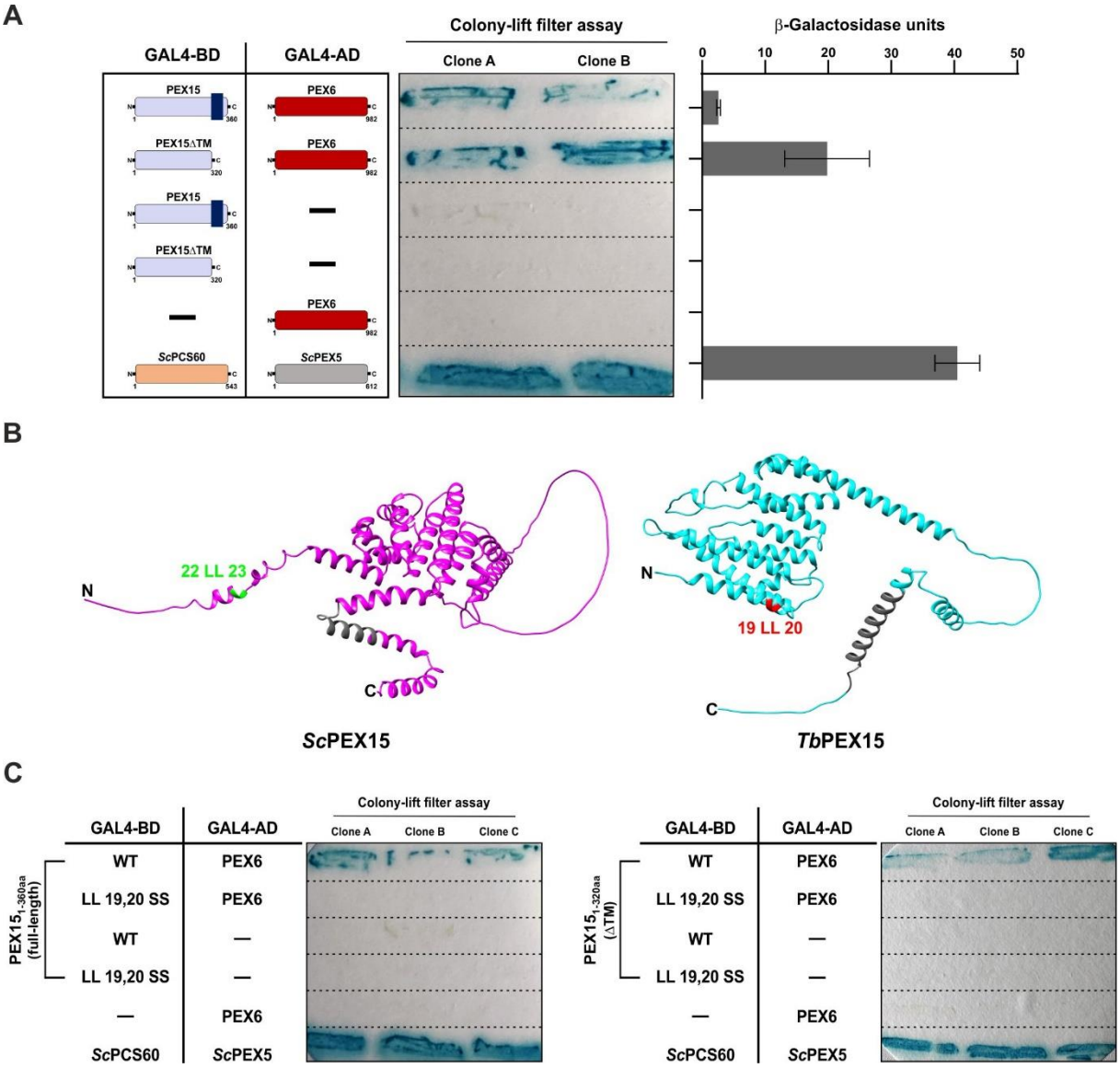


Figure 5: Validation of *Tb*PEX15: Analysis of interaction with *Tb*PEX6. A) In the Yeast Two-Hybrid Assay, full-length *Tb*PEX6 and *Tb*PEX15 (either full-length or lacking the transmembrane domain) were fused to GAL4-AD and -BD, respectively (left panel). These constructs were then co-transformed into the PCY2 yeast strain. The interaction was assessed through Colony-lift filter assays (middle panel) and liquid ONPG assays (right panel), with yeast PEX5-PCS60 as a positive control. The results indicate that both full-length *Tb*PEX15 and *Tb*PEX15 lacking the transmembrane domain (ΔTM) show a clear interaction with full-length *Tb*PEX6. The GAL4-AD fusion of *Tb*PEX6 and the various GAL4-BD fusions of *Tb*PEX15 did not result in autoactivation. The presented β-galactosidase activity units represent the average values of technical replicates from three different clones, with error bars indicating the mean and standard deviations. **B)** Mutational analysis of the *Tb*PEX15-*Tb*PEX6 interaction**.** In the yeast system, as demonstrated by Birschmann et al., it was shown that a mutation of the 22nd Leucine residue to Serine of *Sc*PEX15 leads to a loss of interaction with *Sc*PEX6 (Birschmann et al. 2003). This region contains two consecutive leucine residues located in the first N-terminal helical region of *Sc*PEX15, labelled in green (left panel). Interestingly, the same two consecutive leucine residues are observed in the first N-terminal helical region *Tb*PEX15, labelled in red (right panel). The C-terminal transmembrane domain is marked in grey. **C)** A yeast two-hybrid assay was performed to study the effect of mutation on the interaction of *Tb*PEX15 with *Tb*PEX6. Residues were selected for mutation according to the analysis above, as indicated in the scheme. Results indicate a clear interaction of *Tb*PEX15 (full length as well as ΔTM) with *Tb*PEX6, evident from the blue coloration in the colony-lift filter assay. As reported for the 22nd L mutation in *Sc*PEX15, also the LL mutation in *Tb*PEX15 (full length as well as ΔTM) did abolish the interaction with *Tb*PEX6. The GAL4-AD fusion of *Tb*PEX6 and the various GAL4-BD fusions of *Tb*PEX15 (full-length and ΔTM) WT/mutants were tested for autoactivation, and no autoactivation was observed. The assay was performed in three replicates with three different clones.

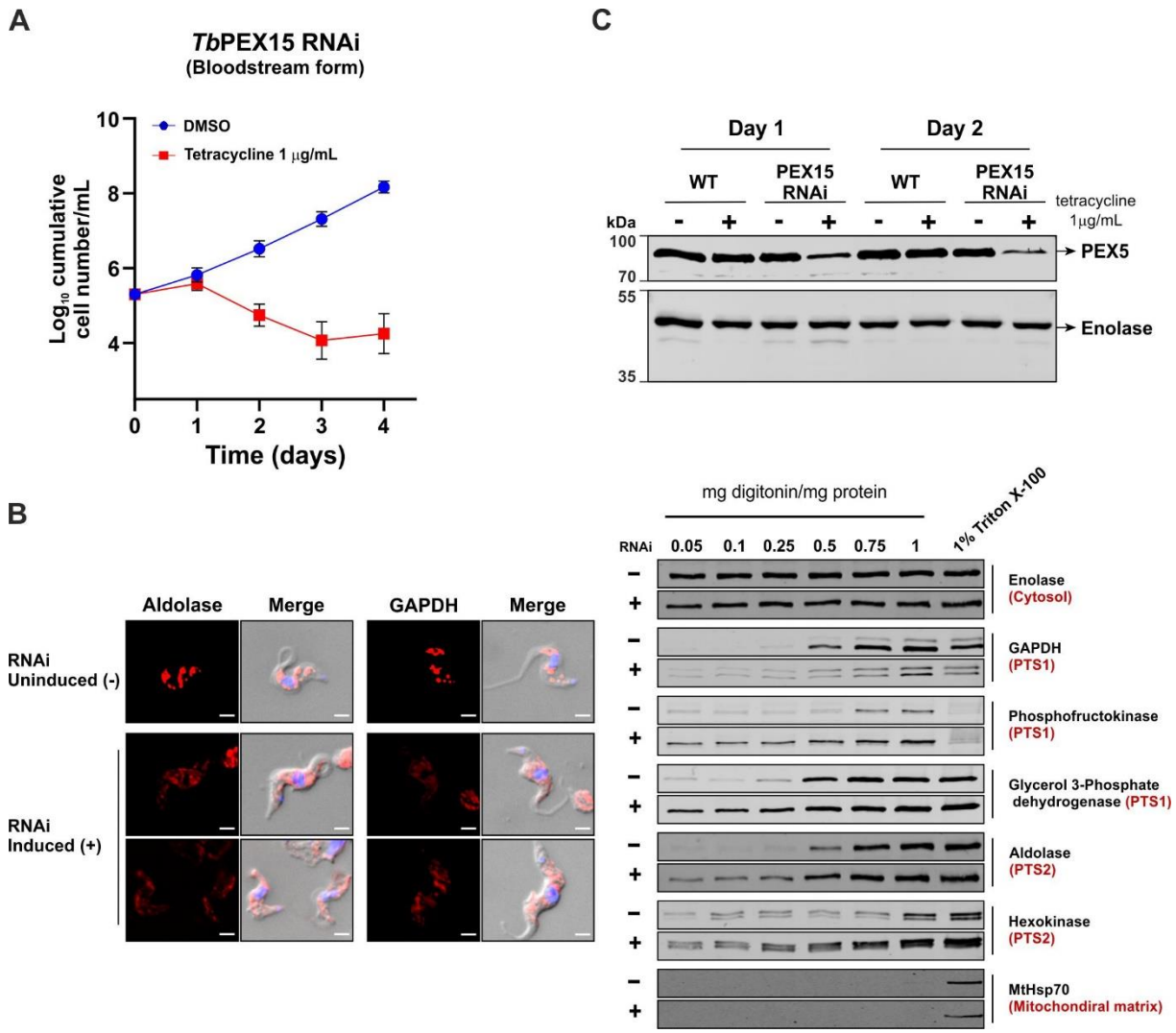
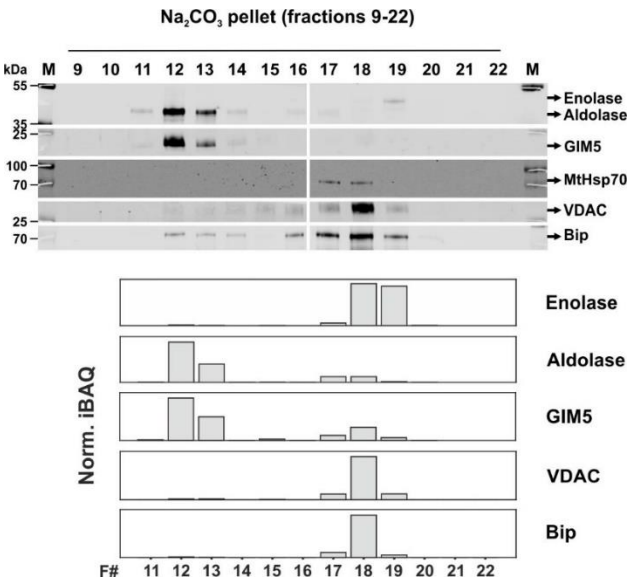
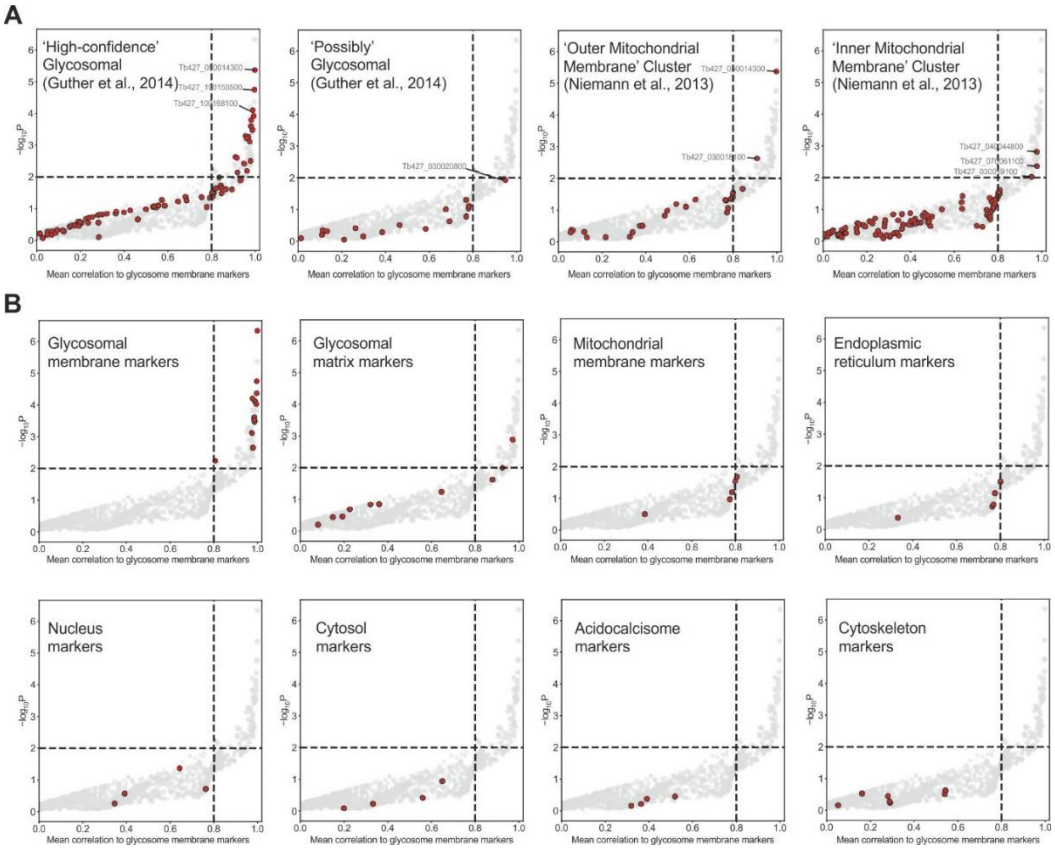


Figure 6: PEX15 is essential for *Trypanosoma* glycosome biogenesis. A) Cumulative growth curves for uninduced (DMSO) and induced (Tetracycline) PEX15 RNAi cell lines demonstrate that *Tb*PEX15 is essential for the parasite survival. The data represents the Log_{10} of cumulative cell counts plotted using GraphPad Prism. The experiment was conducted in triplicate, and error bars indicate the mean and standard deviations. **B)** PEX15 RNAi results in the mislocalization of glycosomal matrix proteins, confirmed by immunofluorescence microscopy (left panel) and biochemical fractionation of cells using digitonin (right panel). On day 2 of PEX15 RNAi, both DMSO and tetracycline treated cells were fixed with 4% formaldehyde and processed for immunofluorescence microscopy with aldolase and GAPDH. In DMSO-treated cells, aldolase and GAPDH display a punctate glycosomal pattern. However, upon PEX15 RNAi induction, partial or complete mislocalization of aldolase and GAPDH to the cytosol is seen. The scale bar represents 2 μm. Additionally, mislocalization of glycosomal matrix proteins was assessed by biochemical fractionation of cells using digitonin. On day 2 of PEX15 RNAi, both DMSO-treated and Tetracycline-induced cells were incubated with increasing digitonin concentrations. Cell suspensions were centrifuged to yield supernatants, which were further analyzed by immunoblotting using various antibodies as indicated. Cells treated with 1% Triton-X 100 represents complete release of proteins by dissolving all membranes. **C)** Degradation of glycosomal import receptor PEX5 upon PEX15 RNAi. Immunoblot analysis of total cell lysates from wildtype and PEX15 RNAi cells, treated with DMSO and Tetracycline, was performed using antibodies against PEX5 and Enolase. The analysis clearly demonstrates the instability of PEX5 (receptor for PTS1 cargo proteins and co-receptor for PTS2 proteins) upon induction of PEX15 RNAi, with significant degradation observed in both day 1 and day 2. Enolase served as the loading control.

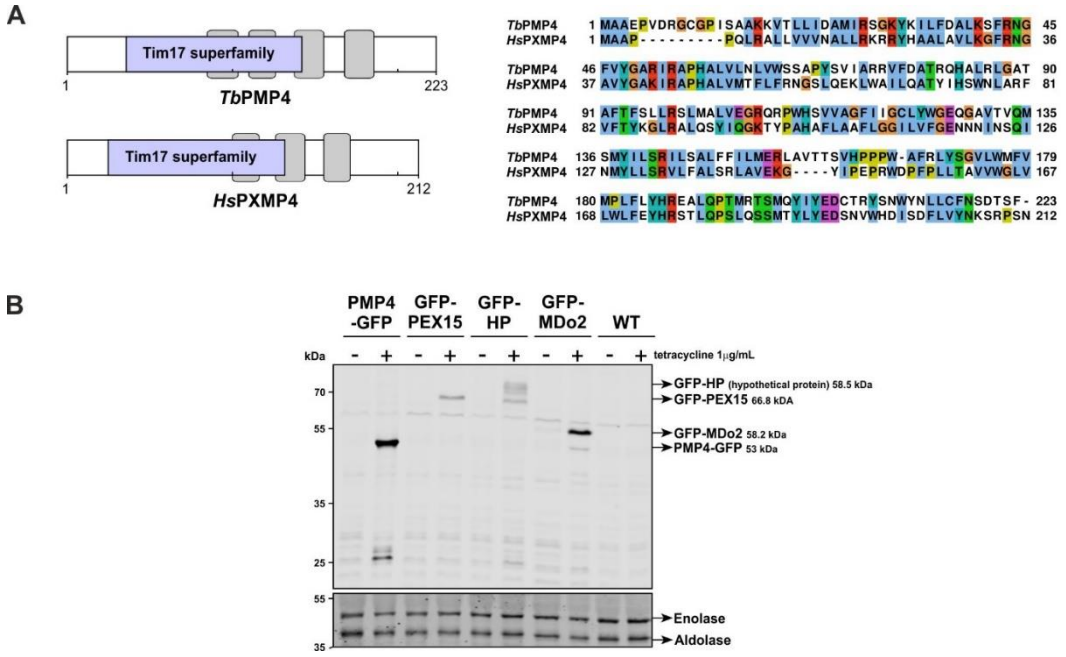
Supplementary Figures



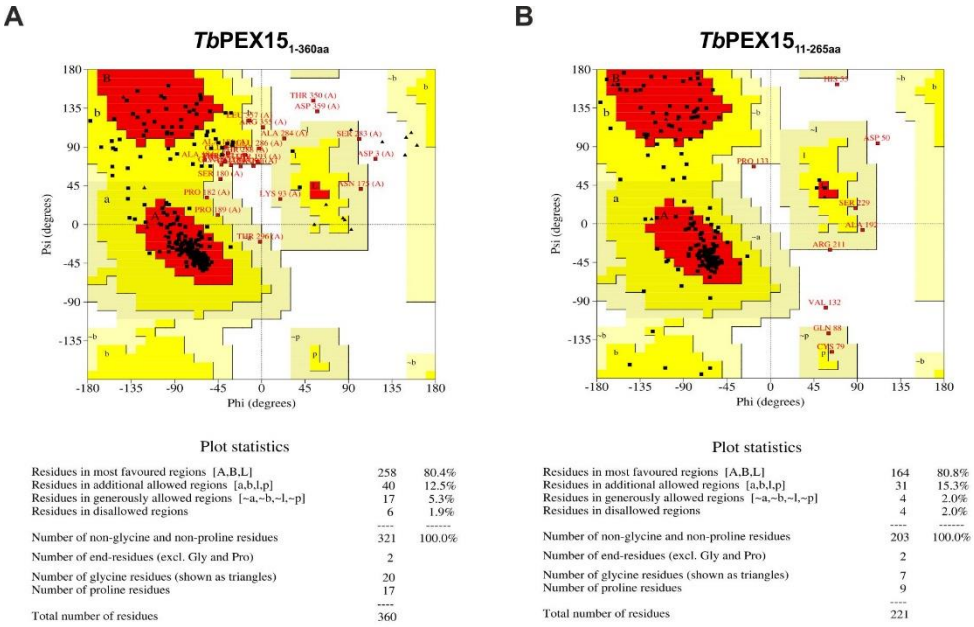
Supp. figure 1: Upper panel: Immunoblot analysis of carbonate-resistant protein samples of fractions F#9-22 (shown in Fig. 1A) using indicated antibodies (upper panel) and abundance profiles of selected organellar marker proteins across density gradient fractions F#11-22 based on normalized MS intensities (iBAQ values) (lower panel). The markers include enolase (cytosolic marker), aldolase (glycosomal matrix marker), GIM5 (glycosomal membrane marker), MtHSP70 (mitochondrial matrix marker), VDAC (mitochondrial membrane marker), and BiP (ER marker). Norm iBAQ: Normalized intensity-based absolute quantification values.



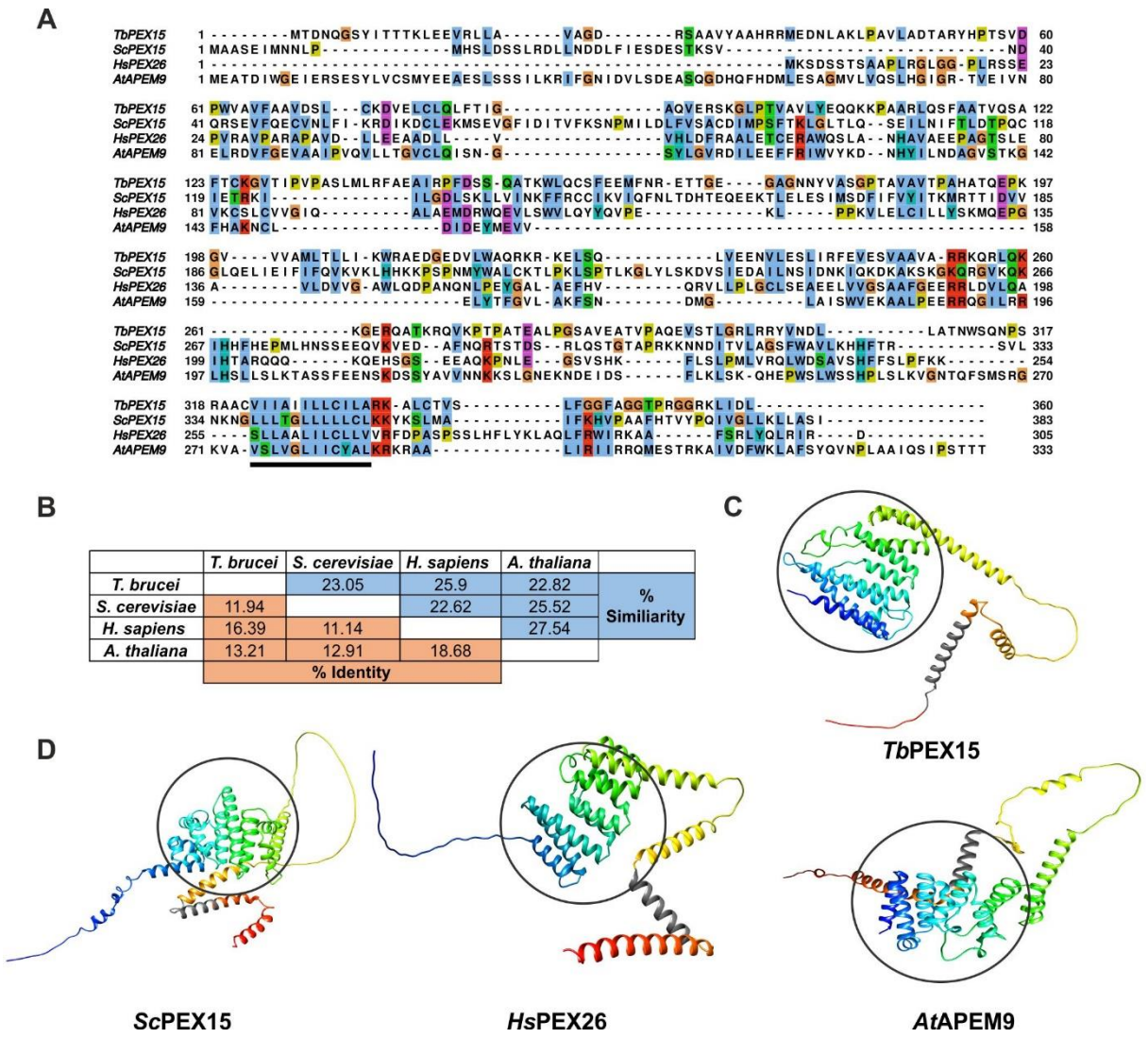
Supp. figure 2: Quality control of glycosomal membrane protein classification based on the known glycosomal and subcellular marker proteins. A) Rank-product volcano plots of proteins quantified in the dataset from PCP of membrane-enrichment steps. 'High confidence glycosomal' and 'Possibly glycosomal' categories in Güther et al., 2014 are highlighted as red spots on the left. Additionally, outer, and inner mitochondrial proteins from (Niemann et al. 2013) are indicated in right panels. **B)** Same as in (A), highlighting selected protein groups as indicated.



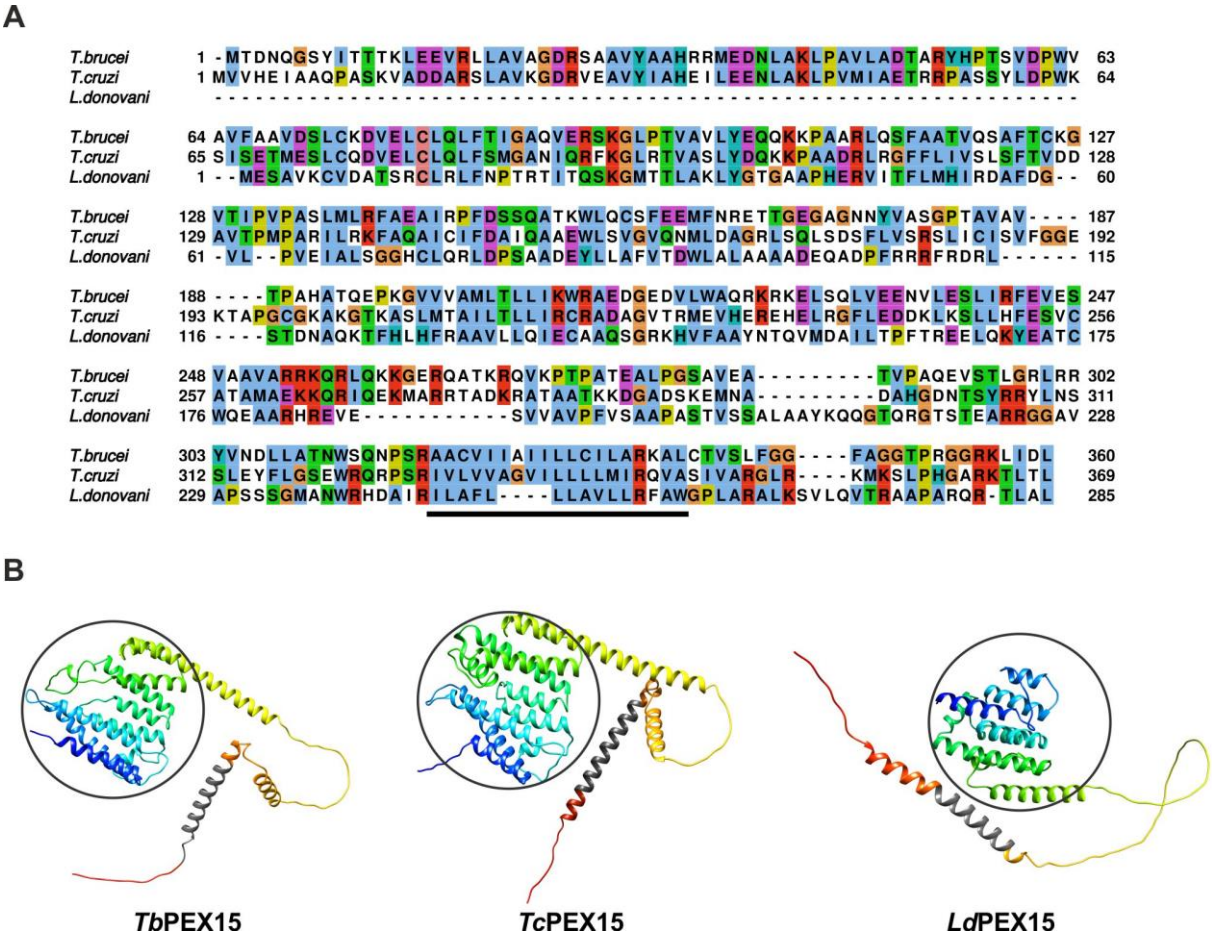
Supp. figure 3: Domain architecture of the validated glycosomal membrane proteins. A) Schematic representation of the domain architecture of the putative *Tb*PMP4 and *Hs*PXMP4 (left panel). The Tim17 superfamily domain is well conserved and is shown in blue. Multiple sequence alignment of *Tb*PMP4 (*T. brucei*) and *Hs*PXMP4 (*H. sapiens*) protein sequence (right panel). *Trypanosoma brucei* protein sequences were obtained from TriTrypDB with Accession ID Tb427_090008200 whereas human protein sequence from UniProt with Uniport ID Q9Y6I8Scheme is represented in relation to transmembrane segments predicted using Phobius webtool (https://www.ebi.ac.uk/Tools/pfa/phobius/). The domain architecture of the proteins based on the InterPro Scan. **B)** The expression of tetracycline-inducible N/C-terminally GFP-tagged PMPs in procyclic form (PCF) trypanosomes (shown in Fig. 3B) was verified by immunoblotting using monoclonal antibodies against the GFP tag (upper panel). Enolase and aldolase served as a loading control (upper panel).



Supp. figure 4: Validation of modelled structure using Ramachandran plot. A) The Ramachandran plot for *Tb*PEX15 protein (full-length) structure modelled using AlphaFold2. **B)** Likewise, the Ramachandran plot for *Tb*PEX15 protein (11-265aa), modelled using the Phyre2 tool with the crystal structure of *Sc*PEX15 (43-259aa) as a template. In both plots, ≤ 2% of residues were observed in the disallowed region (white area) of Ramachandran plot. Ramachandran plots were generated using the PROCHECK tool from the Saves server (https://saves.mbi.ucla.edu/).



Supp. figure 5: Bioinformatic analysis of PEX15/PEX26 orthologs across organisms. A) Multiple sequence alignment of *Tb*PEX15 (*T. brucei*), *Sc*PEX15 (*S. cerevisiae*), *Hs*PEX26 (*H. sapiens*) and *At*APEM9 (*Arabidopsis thaliana*) protein sequences. *Tb*PEX15 protein sequence was obtained from TriTrypDB with accession ID Tb427_100021800 whereas yeast, human and plant protein sequences were obtained from UniProt with UniProt ID Q08215 (*S. cerevisiae*), Q7Z412 (*H. sapiens*), Q8W4B2 (*Arabidopsis thaliana*). Multiple Sequence alignment was performed using the Clustal Omega tool, and the aligned sequences were visualized using JalView software with a Clustal color scheme. The conserved C-terminal domain comprises the predominant portion of the transmembrane domain across organisms (underlined). **B)** Percentage identity and similarity matrix of PEX15 counterpart across organism based on SIAS homology modelling (http://imed.med.ucm.es/Tools/sias.html). **C, D)** AlphaFold predicted structures of PEX15 (*T. brucei* and *S. cerevisiae*), PEX26 (*H. sapiens*), and APEM9 (*Arabidopsis thaliana*) show high similarity. The black circles represent the common helical bundle which is known to interact with PEX6. The predicted structures are presented with a rainbow color scheme, where the N-terminus is blue, and the C-terminus is in red color. The characteristic predicted C-terminal transmembrane domain and the putative hydrophobic membrane-spanning segment of *Hs*PEX26 (citation) are marked in grey.



Supp. figure 6: Bioinformatic analysis of PEX15 orthologs of parasites. A) Multiple sequence alignment of *Tb*PEX15 (*T. brucei*), *Tc*PEX15 (*T. cruzi*), and *Ld*PEX15 (*L. donovani*) protein sequences. Parasite protein sequences were obtained from TriTrypDB with accession IDs Tb427_100021800 (*T. brucei*), TcCLB.508059.20 (*T. cruzi*), and LdBPK_210110.1 (*L. donovani*). Multiple sequence alignment was performed using the Clustal Omega tool, and the aligned sequences were visualized using JalView software with a ClustalX color scheme. Alignment indicates that *Trypanosoma* (*Tb* and *Tc*) PEX15 have longer N-terminal extension than *Leishmania* counterparts. The conserved C-terminal domain comprises the predominant portion of the predicted transmembrane domain (underlined). **B)** AlphaFold predicted structures of PEX15 (*T. brucei*, *T. cruzi*, and *L. donovani*) show high similarity. The black circles represent the common helical bundle similar to yeast and human counterpart. The predicted structures are presented with a rainbow color scheme, where the N-terminus is blue, and the C-terminus is red in color. The predicted C-terminal transmembrane domain is marked in grey.

Table 1: Strains and Plasmids

Sl no.	Expression in	Construct	Primer pair	Restriction sites	Cloned in vector
1	*S. cerevisiae*	GAL4 BD-*Tb*PEX15$_{1\text{-}360aa}$	RE8663 – RE8664	SalI/NotI	pPC97
2	*S. cerevisiae*	GAL4 BD-*Tb*PEX15$_{1\text{-}320aa}$	RE8663 – RE8665	SalI/NotI	pPC97
3	*S. cerevisiae*	GAL4 BD-*Tb*PEX15$_{1\text{-}360aa}$ (LL 19,20 SS)	RE8755 – RE8756	Quick change PCR	pPC97
4	*S. cerevisiae*	GAL4 BD-*Tb*PEX15$_{1\text{-}320aa}$ (LL 19,20 SS)	RE8755 – RE8756	Quick change PCR	pPC97
5	*S. cerevisiae*	GAL4 AD-*Sc*PEX5	(Kerssen et al. 2006)		
6	*S. cerevisiae*	GAL4 BD-*Sc*PCS60	(Effelsberg et al. 2016)		
7	*S. cerevisiae*	GAL4 AD-*Tb*PEX6	(Mahadevan et al. 2023)		
8	*T. brucei*	*Tb*PMP4-GFP	RE7839 – RE7840	HindIII/BamHI	pGN1
9	*T. brucei*	GFP-*Tb*PEX15	RE8896 – RE8897 (vector) / RE8898 – RE8899 (insert)	In-fusion cloning	pGC1
10	*T. brucei*	GFP-*Tb*HP	RE7476 – RE7477	HindIII/BamHI	pGC1
11	*T. brucei*	GFP-*Tb*MDo2	RE7579 – RE7578	HindIII/BamHI	pGC1
12	*T. brucei*	GFP-PTS1	(Dawidowski et al. 2017)		
13	*T. brucei*	*Tb*PEX15 RNAi (Fragment 1)	RE8659 – RE8660	BamHI/MfeI	Ligate fragments 1 & 2
14	*T. brucei*	*Tb*PEX15 RNAi (Fragment 2)	RE8661 – RE8662	HindIII/MfeI	
15	*T. brucei*	*Tb*PEX15 RNAi (Stem-loop)	Ligated fragments 1 & 2	BamHI/HindIII	p2T7-177

Table 2: Oligonucleotides

Primer	Sequence 5' to 3'
RE7578	AGAGGATCCCTACAATTTGGAGGGCTGCGTGCAGT
RE7579	AAGAAAGCTTCCATGGTGTCGTACGCCTCTGCATT
RE7839	AAGACAAGCTTATGGCGGCTGAGCCCGTGGACAGG
RE7840	GACGGATCCCCAAAGGAAGTATCACTGTTAAAGCA
RE8659	AAGACGGATCCATTACGTAGCATCTGGGCCG
RE8660	AAGACCAATTGCAGAGATCGATCAGCTTC
RE8661	AAGACAAGCTTATTACGTAGCATCTGGGCCG
RE8662	AAGACCAATTGACCGTACACAGGGCTTTACG
RE8663	AAGACGTCGACGATGACAGATAATCAAGGTAGCTAT
RE8664	AACAGCGGCCGCTCAGAGATCGATCAGCTTCCT
RE8665	AAGACGCGGCCGCTTAGGCTGCACGAGATGGGTTCTG
RE8755	GTTCGCAGTTCTGCCGTCGCAGGGGATCG
RE8756	GGCAGAACTGCGAACTTCTTCAAGTTTCGTGG
RE8896	GGATCCACCGGATCTAGATAACTGAT
RE8897	TCGAGATCTGAGTCCGGACTTGTAC
RE8898	TCCGGACTCAGATCTCGAATGACAGATAATCAAGGTAGCTATATCACTACCAC
RE8899	TCTAGATCCGGTGGATCCTCAGAGATCGATCAGCTTCCTTCCG

1130 **Table 3: High-confidence potential glycosomal membrane proteins**

Sl. No	Accession number	UniProt ID(s)	Protein description	Mol. Wt. (kDa)
1	Tb427_100125500[a]	Q389G0	hypothetical protein	70.0
2	Tb427_090008200[b]	N/A	peroxisomal membrane protein 4, putative	25.4
3	Tb427_040018300	Q581S1	AAA domain containing protein, putative	27.2
4	Tb427_040034100	Q583Z3	dihydroxyacetone phosphate acyltransferase, putative	137.1
5	Tb427_110028700	Q386S1	phosphoglycerate kinase, putative	100.9
6	Tb427_100021800[ab]	N/A	hypothetical protein	39.2
7	Tb427_100035900	Q38BT8	methionine biosynthetic protein, putative	44.5
8	Tb427_080048600[a]	Q57VZ8	hypothetical protein	100.4
9	Tb427_110023000[a]	Q4FKA7	hypothetical protein	62.0
10	Tb427_010033000[ab]	Q4GY80	hypothetical protein	30.5
11	Tb427_060035900	Q583M6	hypothetical protein	20.7
12	Tb427_100090900[a]	Q38AC3	hypothetical protein	27.8
13	Tb427_080075900[a]; Tb427_040056900	N/A	hypothetical protein: NADH-ubiquinone oxidoreductase complexes I subunit, putative	42.5
14	Tb427_110110000[ab]	Q383R5	hypothetical protein	30.3
15	Tb427_040044800	Q57Y25, Q581J6	ATP-dependent zinc metalloprotease AFG3-like, putative	96.3
16	Tb427_100029100[a]	Q38BZ7	hypothetical protein	105.2
17	Tb427_110044000	Q386C9	AMP deaminase, putative	175.6
18	Tb427_100067300	Q38AX7	ATP-dependent DEAD/H RNA helicase, putative	57.6
19	Tb427_110154800[a]	N/A	hypothetical protein	20.9
20	Tb427_080072500[a]	Q57TV1	hypothetical protein	40.1
21	Tb427_070013500	Q57VU5	chaperone protein DnaJ, putative	86.6
22	Tb427_070006700	N/A	Prostaglandin E synthase, putative	38.0
23	Tb427_010022900	Q4GYQ5	40S ribosomal protein S11, putative	20.1
24	Tb427_050014000; Tb427_110072300	N/A;Q57W38,Q5BU13,Q8ITL2	hypothetical protein;NADH-dependent fumarate reductase	123.6
25	Tb427_080052800	D6XLL1, Q26687, Q9NJQ1	small GTP-binding protein Rab18	23.5
26	Tb427_050022100[a]	Q57VA0	hypothetical protein	17.7
27	Tb427_080080600[a]	Q57YV0	hypothetical protein	70.7
28	Tb427_090015900	Q38FQ3	Engulfment and cell motility domain-containing protein	41.1
29	Tb427_110103500	Q383X4	replication factor C, subunit 4, putative	37.6
30	Tb427_070061100	Q582G9	2,4-dienoyl-coa reductase-like protein	80.4
31	Tb427_090047200; Tb427_090045300	Q38EC5; Q38EC5, Q967R7	QM-like protein	24.7
32	Tb427_070045200	Q57UM8	tc40 antigen-like	97.2
33	Tb427_110066300; Tb427_030046700	Q580K6	40S ribosomal protein S15A, putative	14.6

34	Tb427_090070500; Tb427_090070400	Q38DH8	60S ribosomal subunit protein L31, putative	21.7
35	Tb427_100148400[a]	Q388J7	hypothetical protein	48.1
36	Tb427_020030300	Q586C9, Q95PL7	sedoheptulose-1,7-bisphosphatase	35.5
37	Tb427_100123000	Q389I7	40S ribosomal protein S3, putative	24.2
38	Tb427_110055300[a]	Q385R6	hypothetical protein	186.6
39	Tb427_030019100	Q57ZB7	TLP18.3, Psb32 and MOLO-1 founding proteins of phosphatase, putative	49.4

a Conserved in trypanosomatids only

b Validated in this study

References:

Acestor, N., A. K. Panigrahi, Y. Ogata, A. Anupama and K. D. Stuart (2009). "Protein composition of Trypanosoma brucei mitochondrial membranes." Proteomics **9**(24): 5497-5508.

Acosta, H., R. Burchmore, C. Naula, M. Gualdron-Lopez, E. Quintero-Troconis, A. J. Caceres, P. A. M. Michels, J. L. Concepcion and W. Quinones (2019). "Proteomic analysis of glycosomes from Trypanosoma cruzi epimastigotes." Mol Biochem Parasitol **229**: 62-74.

Agrawal, G. and S. Subramani (2016). "De novo peroxisome biogenesis: Evolving concepts and conundrums." Biochim Biophys Acta **1863**(5): 892-901.

Andersen, J. S., C. J. Wilkinson, T. Mayor, P. Mortensen, E. A. Nigg and M. Mann (2003). "Proteomic characterization of the human centrosome by protein correlation profiling." Nature **426**(6966): 570-574.

Banerjee, H., B. Knoblach and R. A. Rachubinski (2019). "The early-acting glycosome biogenic protein Pex3 is essential for trypanosome viability." Life Sci Alliance **2**(4).

Birschmann, I., A. K. Stroobants, M. van den Berg, A. Schafer, K. Rosenkranz, W. H. Kunau and H. F. Tabak (2003). "Pex15p of Saccharomyces cerevisiae provides a molecular basis for recruitment of the AAA peroxin Pex6p to peroxisomal membranes." Mol Biol Cell **14**(6): 2226-2236.

Breitling, R. and P. Herzyk (2005). "Rank-based methods as a non-parametric alternative of the T-statistic for the analysis of biological microarray data." J Bioinform Comput Biol **3**(5): 1171-1189.

Brun, R. and Schonenberger (1979). "Cultivation and in vitro cloning or procyclic culture forms of Trypanosoma brucei in a semi-defined medium. Short communication." Acta Trop **36**(3): 289-292.

Buitinck, L., G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt and G. Varoquaux (2013). API design for machine learning software: experiences from the scikit-learn project. ECML PKDD Workshop: Languages for Data Mining and Machine Learning.

Chou, S. E., V. C. Kalel and R. Erdmann (2023). "Biochemical Fractionation of Trypanosomes for the Analysis of Glycosomal Protein Import Defects." Methods Mol Biol **2643**: 445-453.

Colasante, C., F. Voncken, T. Manful, T. Ruppert, A. G. M. Tielens, J. J. van Hellemond and C. Clayton (2013). "Proteins and lipids of glycosomal membranes from Leishmania tarentolae and Trypanosoma brucei." F1000Res **2**: 27.

Cox, J. and M. Mann (2008). "MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification." Nat Biotechnol **26**(12): 1367-1372.

Cox, J., N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen and M. Mann (2011). "Andromeda: a peptide search engine integrated into the MaxQuant environment." J Proteome Res **10**(4): 1794-1805.

Dawidowski, M., L. Emmanouilidis, V. C. Kalel, K. Tripsianes, K. Schorpp, K. Hadian, M. Kaiser, P. Mäser, M. Kolonko, S. Tanghe, A. Rodriguez, W. Schliebs, R. Erdmann, M. Sattler and G. M. Popowicz (2017). "Inhibitors of PEX14 disrupt protein import into glycosomes and kill Trypanosoma parasites." Science **355**(6332): 1416-1420.

De Rycker, M., S. Wyllie, D. Horn, K. D. Read and I. H. Gilbert (2023). "Anti-trypanosomatid drug discovery: progress and challenges." Nat Rev Microbiol **21**(1): 35-50.

Del Carratore, F., A. Jankevics, R. Eisinga, T. Heskes, F. Hong and R. Breitling (2017). "RankProd 2.0: a refactored bioconductor package for detecting differentially expressed features in molecular profiling datasets." Bioinformatics **33**(17): 2774-2775.

DNDi. (2023). "*Visceral Leishmaniasis Facts* | Drugs for Neglected Diseases initiative (DNDi)." Retrieved August 31, 2023, from https://dndi.org/diseases/visceral-leishmaniasis/facts/.

Effelsberg, D., L. D. Cruz-Zaragoza, W. Schliebs and R. Erdmann (2016). "Pex9p is a new yeast peroxisomal import receptor for PTS1-containing proteins." J Cell Sci **129**(21): 4057-4066.

Erdmann, R. and W. Schliebs (2005). "Peroxisomal matrix protein import: the transient pore model." Nat Rev Mol Cell Biol **6**(9): 738-742.

Foster, L. J., C. L. de Hoog, Y. Zhang, Y. Zhang, X. Xie, V. K. Mootha and M. Mann (2006). "A mammalian organelle map by protein correlation profiling." Cell **125**(1): 187-199.

Galland, N., F. Demeure, V. Hannaert, E. Verplaetse, D. Vertommen, P. Van der Smissen, P. J. Courtoy and P. A. Michels (2007). "Characterization of the role of the receptors PEX5 and PEX7 in the import of proteins into glycosomes of Trypanosoma brucei." Biochim Biophys Acta **1773**(4): 521-535.

Gao, Y., M. L. Skowyra, P. Feng and T. A. Rapoport (2022). "Protein import into peroxisomes occurs through a nuclear pore-like phase." Science **378**(6625): eadf3971.

Gardner, B. M., D. T. Castanzo, S. Chowdhury, G. Stjepanovic, M. S. Stefely, J. H. Hurley, G. C. Lander and A. Martin (2018). "The peroxisomal AAA-ATPase Pex1/Pex6 unfolds substrates by processive threading." Nat Commun **9**(1): 135.

Gardner, B. M., D. T. Castanzo, S. Chowdhury, G. Stjepanovic, M. S. Stefely, J. H. Hurley, G. C. Lander and A. Martin (2018). "The peroxisomal AAA-ATPase Pex1/Pex6 unfolds substrates by processive threading." Nature Communications **9**(1): 135.

Gerber, M., I. Suppanz, S. Oeljeklaus, M. Niemann, S. Kaser, B. Warscheid, A. Schneider and C. E. Dewar (2023). "A Msp1-containing complex removes orphaned proteins in the mitochondrial outer membrane of T. brucei." Life Sci Alliance **6**(11): e202302004.

Gietz, R. D. and R. A. Woods (2002). "Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method." Methods Enzymol **350**: 87-96.

Goto, S., S. Mano, C. Nakamori and M. Nishimura (2011). "Arabidopsis ABERRANT PEROXISOME MORPHOLOGY9 is a peroxin that recruits the PEX1-PEX6 complex to peroxisomes." Plant Cell **23**(4): 1573-1587.

Gould, S. J., G. A. Keller, N. Hosken, J. Wilkinson and S. Subramani (1989). "A conserved tripeptide sorts proteins to peroxisomes." J Cell Biol **108**(5): 1657-1664.

Gualdron-Lopez, M., A. Brennand, L. Avilan and P. A. Michels (2013). "Translocation of solutes and proteins across the glycosomal membrane of trypanosomes; possibilities and limitations for targeting with trypanocidal drugs." Parasitology **140**(1): 1-20.

Guther, M. L., M. D. Urbaniak, A. Tavendale, A. Prescott and M. A. Ferguson (2014). "High-confidence glycosome proteome for procyclic form Trypanosoma brucei by epitope-tag organelle enrichment and SILAC proteomics." J Proteome Res **13**(6): 2796-2806.

Haanstra, J. R., B. M. Bakker and P. A. Michels (2014). "In or out? On the tightness of glycosomal compartmentalization of metabolites and enzymes in Trypanosoma brucei." Mol Biochem Parasitol **198**(1): 18-28.

Haanstra, J. R., A. van Tuijl, P. Kessler, W. Reijnders, P. A. Michels, H. V. Westerhoff, M. Parsons and B. M. Bakker (2008). "Compartmentation prevents a lethal turbo-explosion of glycolysis in trypanosomes." Proc Natl Acad Sci U S A **105**(46): 17718-17723.

Haikarainen, T. and L. Lehtio (2016). "Proximal ADP-ribose Hydrolysis in Trypanosomatids is Catalyzed by a Macrodomain." Sci Rep **6**(1): 24213.

Hettema, E. H., R. Erdmann, I. van der Klei and M. Veenhuis (2014). "Evolving models for peroxisome biogenesis." Curr Opin Cell Biol **29**: 25-30.

Hirumi, H. and K. Hirumi (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**(6): 985-989.

Huang, Q., D. Szklarczyk, M. Wang, M. Simonovic and C. von Mering (2023). "PaxDb 5.0: Curated Protein Quantification Data Suggests Adaptive Proteome Changes in Yeasts." Mol Cell Proteomics **22**(10): 100640.

Jardim, A., D. B. Hardie, J. Boitz and C. H. Borchers (2018). "Proteomic Profiling of Leishmania donovani Promastigote Subcellular Organelles." J Proteome Res **17**(3): 1194-1215.

Judy, R. M., C. J. Sheedy and B. M. Gardner (2022). "Insights into the Structure and Function of the Pex1/Pex6 AAA-ATPase in Peroxisome Homeostasis." Cells **11**(13): 2067.

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S. A. A. Kohl, A. J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A. W. Senior, K. Kavukcuoglu, P. Kohli and D. Hassabis (2021). "Highly accurate protein structure prediction with AlphaFold." Nature **596**(7873): 583-589.

Kalel, V. C., L. Emmanouilidis, M. Dawidowski, W. Schliebs, M. Sattler, G. M. Popowicz and R. Erdmann (2017). "Inhibitors of glycosomal protein import provide new leads against trypanosomiasis." Microb Cell **4**(7): 229-232.

Kalel, V. C., M. Li, S. Gaussmann, F. Delhommel, A. B. Schafer, B. Tippler, M. Jung, R. Maier, S. Oeljeklaus, W. Schliebs, B. Warscheid, M. Sattler and R. Erdmann (2019). "Evolutionary divergent PEX3 is essential for glycosome biogenesis and survival of trypanosomatid parasites." Biochim Biophys Acta Mol Cell Res **1866**(12): 118520.

Kalel, V. C., P. Maser, M. Sattler, R. Erdmann and G. M. Popowicz (2018). "Come, sweet death: targeting glycosomal protein import for antitrypanosomal drug development." Curr Opin Microbiol **46**: 116-122.

Kalel, V. C., W. Schliebs and R. Erdmann (2015). "Identification and functional characterization of Trypanosoma brucei peroxin 16." Biochim Biophys Acta **1853**(10 Pt A): 2326-2337.

Kande Betu Ku Mesu, V., W. Mutombo Kalonji, C. Bardonneau, O. Valverde Mordt, D. Ngolo Tete, S. Blesson, F. Simon, S. Delhomme, S. Bernhard, H. Mahenzi Mbembo, C. Mpia Moke, S. Lumeya Vuvu, J. Mudji E'kitiak, F. Akwaso Masa, M. Mukendi Ilunga, D. Mpoyi Muamba Nzambi, T. Mayala Malu, S. Kapongo Tshilumbwa, F. Botalema Bolengi, M. Nkieri Matsho, C. Lumbala, B. Scherrer, N. Strub-Wourgaft and A. Tarral (2021). "Oral fexinidazole for stage 1 or early stage 2 African Trypanosoma brucei gambiense trypanosomiasis: a prospective, multicentre, open-label, cohort study." Lancet Glob Health **9**(7): e999-e1008.

Kaufer, A., J. Ellis, D. Stark and J. Barratt (2017). "The evolution of trypanosomatid taxonomy." Parasit Vectors **10**(1): 287.

Kelley, L. A., S. Mezulis, C. M. Yates, M. N. Wass and M. J. Sternberg (2015). "The Phyre2 web portal for protein modeling, prediction and analysis." Nat Protoc **10**(6): 845-858.

Kerssen, D., E. Hambruch, W. Klaas, H. W. Platta, B. de Kruijff, R. Erdmann, W. H. Kunau and W. Schliebs (2006). "Membrane association of the cycling peroxisome import receptor Pex5p." J Biol Chem **281**(37): 27003-27015.

Kim, P. (2017). "Peroxisome Biogenesis: A Union between Two Organelles." Curr Biol **27**(7): R271-R274.

Kourbeli, V., E. Chontzopoulou, K. Moschovou, D. Pavlos, T. Mavromoustakos and I. P. Papanastasiou (2021). "An Overview on Target-Based Drug Design against Kinetoplastid Protozoan Infections: Human African Trypanosomiasis, Chagas Disease and Leishmaniases." Molecules **26**(15): 4629.

Krazy, H. and P. A. Michels (2006). "Identification and characterization of three peroxins--PEX6, PEX10 and PEX12--involved in glycosome biogenesis in Trypanosoma brucei." Biochim Biophys Acta **1763**(1): 6-17.

Krishna, C. K., L. Franke, R. Erdmann and V. C. Kalel (2023). "Isolation of Glycosomes from Trypanosoma brucei." Methods Mol Biol **2643**: 33-45.

Krishna, C. K., N. Schmidt, B. G. Tippler, W. Schliebs, M. Jung, K. F. Winklhofer, R. Erdmann and V. C. Kalel (2023). "Molecular basis of the glycosomal targeting of PEX11 and its mislocalization to mitochondrion in trypanosomes." Front Cell Dev Biol **11**: 1213761.

Kushwaha, V. and N. Capalash (2022). "Aminoacyl-tRNA synthetase (AARS) as an attractive drug target in neglected tropical trypanosomatid diseases-Leishmaniasis, Human African Trypanosomiasis and Chagas disease." Mol Biochem Parasitol **251**: 111510.

Laskowski, R. A., M. W. Macarthur, D. S. Moss and J. M. Thornton (1993). "Procheck - a Program to Check the Stereochemical Quality of Protein Structures." Journal of Applied Crystallography **26**(2): 283-291.

Lazarow, P. B. and Y. Fujiki (1985). "Biogenesis of peroxisomes." Annu Rev Cell Biol **1**: 489-530.

Li, C., A. Wen, B. Shen, J. Lu, Y. Huang and Y. Chang (2011). "FastCloning: a highly simplified, purification-free, sequence- and ligation-independent PCR cloning method." BMC Biotechnol **11**(1): 92.

Li, M., S. Gaussmann, B. Tippler, J. Ott, G. M. Popowicz, W. Schliebs, M. Sattler, R. Erdmann and V. C. Kalel (2021). "Novel Trypanocidal Inhibitors that Block Glycosome Biogenesis by Targeting PEX3-PEX19 Interaction." Front Cell Dev Biol **9**: 737159.

Lorenz, P., A. G. Maier, E. Baumgart, R. Erdmann and C. Clayton (1998). "Elongation and clustering of glycosomes in Trypanosoma brucei overexpressing the glycosomal Pex11p." EMBO J **17**(13): 3542-3555.

Madeira, F., M. Pearce, A. R. N. Tivey, P. Basutkar, J. Lee, O. Edbali, N. Madhusoodanan, A. Kolesnikov and R. Lopez (2022). "Search and sequence analysis tools services from EMBL-EBI in 2022." Nucleic Acids Res **50**(W1): W276-W279.

Mahadevan, L., H. Arya, W. Schliebs, R. Erdmann and V. C. Kalel (2023). "PEX1 is essential for the glycosome biogenesis and trypanosomatid parasite survival." bioRxiv: 2023.2008.2005.550684.

Matsumoto, N., S. Tamura and Y. Fujiki (2003). "The pathogenic peroxin Pex26p recruits the Pex1p-Pex6p AAA ATPase complexes to peroxisomes." Nat Cell Biol **5**(5): 454-460.

Meinecke, M., P. Bartsch and R. Wagner (2016). "Peroxisomal protein import pores." Biochim Biophys Acta **1863**(5): 821-827.

Michels, P. A., J. Moyersoen, H. Krazy, N. Galland, M. Herman and V. Hannaert (2005). "Peroxisomes, glyoxysomes and glycosomes (review)." Mol Membr Biol **22**(1-2): 133-145.

Michels, P. A. M. and M. Gualdron-Lopez (2022). "Biogenesis and metabolic homeostasis of trypanosomatid glycosomes: New insights and new questions." J Eukaryot Microbiol **69**(6): e12897.

Mirdita, M., K. Schutze, Y. Moriwaki, L. Heo, S. Ovchinnikov and M. Steinegger (2022). "ColabFold: making protein folding accessible to all." Nat Methods **19**(6): 679-682.

Morales, J., M. Hashimoto, T. A. Williams, H. Hirawake-Mogi, T. Makiuchi, A. Tsubouchi, N. Kaga, H. Taka, T. Fujimura, M. Koike, T. Mita, F. Bringaud, J. L. Concepcion, T. Hashimoto, T. M. Embley and T. Nara (2016). "Differential remodelling of peroxisome function underpins the environmental and metabolic adaptability of diplonemids and kinetoplastids." Proc Biol Sci **283**(1830): 20160520.

Nagle, A., A. Biggart, C. Be, H. Srinivas, A. Hein, D. Caridha, R. J. Sciotti, B. Pybus, M. Kreishman-Deitrick, B. Bursulaya, Y. H. Lai, M. Y. Gao, F. Liang, C. J. N. Mathison, X. Liu, V. Yeh, J. Smith, I. Lerario, Y. Xie, D. Chianelli, M. Gibney, A. Berman, Y. L. Chen, J. Jiricek, L. C. Davis, X. Liu, J. Ballard, S. Khare, F. K. Eggimann, A. Luneau, T. Groessl, M. Shapiro, W. Richmond, K. Johnson, P. J. Rudewicz, S. P. S. Rao, C. Thompson, T. Tuntland, G. Spraggon, R. J. Glynne, F. Supek, C. Wiesmann and V. Molteni (2020). "Discovery and Characterization of Clinical Candidate LXE408 as a Kinetoplastid-Selective Proteasome Inhibitor for the Treatment of Leishmaniases." J Med Chem **63**(19): 10773-10781.

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

Nuttall, J. M., A. M. Motley and E. H. Hettema (2014). "Deficiency of the exportomer components Pex1, Pex6, and Pex15 causes enhanced pexophagy in Saccharomyces cerevisiae." Autophagy **10**(5): 835-845.

Opperdoes, F. R., P. Baudhuin, I. Coppens, C. De Roe, S. W. Edwards, P. J. Weijers and O. Misset (1984). "Purification, morphometric analysis, and characterization of the glycosomes (microbodies) of the protozoan hemoflagellate Trypanosoma brucei." J Cell Biol **98**(4): 1178-1184.

Opperdoes, F. R. and P. Borst (1977). "Localization of nine glycolytic enzymes in a microbody-like organelle in Trypanosoma brucei: the glycosome." FEBS Lett **80**(2): 360-364.

Oryan, A. and M. Akbari (2016). "Worldwide risk factors in leishmaniasis." Asian Pac J Trop Med **9**(10): 925-932.

Paysan-Lafosse, T., M. Blum, S. Chuguransky, T. Grego, B. L. Pinto, G. A. Salazar, M. L. Bileschi, P. Bork, A. Bridge, L. Colwell, J. Gough, D. H. Haft, I. Letunic, A. Marchler-Bauer, H. Mi, D. A. Natale, C. A. Orengo, A. P. Pandurangan, C. Rivoire, C. J. A. Sigrist, I. Sillitoe, N. Thanki, P. D. Thomas, S. C. E. Tosatto, C. H. Wu and A. Bateman (2023). "InterPro in 2022." Nucleic Acids Res **51**(D1): D418-D427.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay (2011). "Scikit-learn: Machine Learning in Python." Journal of Machine Learning Research **12**: 2825–2830.

Perez-Molina, J. A. and I. Molina (2018). "Chagas disease." Lancet **391**(10115): 82-94.

Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin (2004). "UCSF Chimera--a visualization system for exploratory research and analysis." J Comput Chem **25**(13): 1605-1612.

Platta, H. W., W. Girzalsky and R. Erdmann (2004). "Ubiquitination of the peroxisomal import receptor Pex5p." Biochem J **384**(Pt 1): 37-45.

Platta, H. W., S. Grunau, K. Rosenkranz, W. Girzalsky and R. Erdmann (2005). "Functional role of the AAA peroxins in dislocation of the cycling PTS1 receptor back to the cytosol." Nat Cell Biol **7**(8): 817-822.

Platta, H. W., S. Hagen and R. Erdmann (2013). "The exportomer: the peroxisomal receptor export machinery." Cell Mol Life Sci **70**(8): 1393-1411.

Reguenga, C., M. E. Oliveira, A. M. Gouveia, C. Eckerskorn, C. Sa-Miranda and J. E. Azevedo (1999). "Identification of a 24 kDa intrinsic membrane protein from mammalian peroxisomes." Biochim Biophys Acta **1445**(3): 337-341.

Schwanhausser, B., D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen and M. Selbach (2011). "Global quantification of mammalian gene expression control." Nature **473**(7347): 337-342.

Stuart, K., R. Brun, S. Croft, A. Fairlamb, R. E. Gurtler, J. McKerrow, S. Reed and R. Tarleton (2008). "Kinetoplastids: related protozoan pathogens, different diseases." J Clin Invest **118**(4): 1301-1310.

Swinkels, B. W., S. J. Gould, A. G. Bodnar, R. A. Rachubinski and S. Subramani (1991). "A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase." EMBO J **10**(11): 3255-3262.

Tyanova, S., T. Temu and J. Cox (2016). "The MaxQuant computational platform for mass spectrometry-based shotgun proteomics." Nat Protoc **11**(12): 2301-2319.

Varadi, M., S. Anyango, M. Deshpande, S. Nair, C. Natassia, G. Yordanova, D. Yuan, O. Stroe, G. Wood, A. Laydon, A. Zidek, T. Green, K. Tunyasuvunakool, S. Petersen, J. Jumper, E. Clancy, R. Green, A. Vora, M. Lutfi,

M. Figurnov, A. Cowie, N. Hobbs, P. Kohli, G. Kleywegt, E. Birney, D. Hassabis and S. Velankar (2022). "AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models." Nucleic Acids Res **50**(D1): D439-D444.

Verplaetse, E., D. J. Rigden and P. A. Michels (2009). "Identification, characterization and essentiality of the unusual peroxin 13 from Trypanosoma brucei." Biochim Biophys Acta **1793**(3): 516-527.

Wang, M., C. J. Herrmann, M. Simonovic, D. Szklarczyk and C. von Mering (2015). "Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines." Proteomics **15**(18): 3163-3168.

Wheeler, R. J. (2021). "A resource for improved predictions of Trypanosoma and Leishmania protein three-dimensional structure." PLoS One **16**(11): e0259871.

WHO. (2023). "*Leishmaniasis Fact Sheet* | World Health Organization (WHO)." Retrieved August 31, 2023, from https://www.who.int/news-room/fact-sheets/detail/leishmaniasis.

Wiese, S., T. Gronemeyer, R. Ofman, M. Kunze, C. P. Grou, J. A. Almeida, M. Eisenacher, C. Stephan, H. Hayen, L. Schollenberger, T. Korosec, H. R. Waterham, W. Schliebs, R. Erdmann, J. Berger, H. E. Meyer, W. Just, J. E. Azevedo, R. J. Wanders and B. Warscheid (2007). "Proteomics characterization of mouse kidney peroxisomes by tandem mass spectrometry and protein correlation profiling." Mol Cell Proteomics **6**(12): 2045-2057.