Contents lists available at ScienceDirect

International Journal for Parasitology: Drugs and Drug Resistance
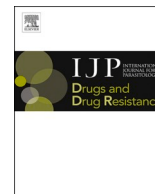
High-throughput prioritization of target proteins for development of new antileishmanial compounds

Lucas G. Azevedo [a,b,*], Ezequiel Sosa [c], Artur T.L. de Queiroz [a,b], Aldina Barral [d], Richard J. Wheeler [e], Marisa F. Nicolás [f], Leonardo P. Farias [b,d], Dario Fernández Do Porto [c], Pablo Ivan P. Ramos [a,b,*]

[a] Center for Data and Knowledge Integration for Health (CIDACS), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (Fiocruz Bahia), Salvador, Bahia, Brazil
[b] Post-graduate Program in Biotechnology and Investigative Medicine, Instituto Gonçalo Moniz, Salvador, Bahia, Brazil
[c] Universidad de Buenos Aires, Buenos Aires, Argentina
[d] Laboratório de Medicina e Saúde Pública de Precisão (MeSP2), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (Fiocruz Bahia), Salvador, Bahia, Brazil
[e] Peter Medawar Building for Pathogen Research, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
[f] Laboratório Nacional de Computação Científica, Petrópolis, Rio de Janeiro, Brazil

ARTICLE INFO

ABSTRACT

Leishmaniasis, a vector-borne disease, is caused by the infection of *Leishmania* spp., obligate intracellular protozoan parasites. Presently, human vaccines are unavailable, and the primary treatment relies heavily on systemic drugs, often presenting with suboptimal formulations and substantial toxicity, making new drugs a high priority for LMIC countries burdened by the disease, but a low priority in the agenda of most pharmaceutical companies due to unattractive profit margins. New ways to accelerate the discovery of new, or the repositioning of existing drugs, are needed. To address this challenge, our study aimed to identify potential protein targets shared among clinically-relevant *Leishmania* species. We employed a subtractive proteomics and comparative genomics approach, integrating high-throughput multi-omics data to classify these targets based on different druggability metrics. This effort resulted in the ranking of 6502 ortholog groups of protein targets across 14 pathogenic *Leishmania* species. Among the top 20 highly ranked groups, metabolic processes known to be attractive drug targets, including the ubiquitination pathway, aminoacyl-tRNA synthetases, and purine synthesis, were rediscovered. Additionally, we unveiled novel promising targets such as the nicotinate phosphoribosyltransferase enzyme and dihydrolipoamide succinyltransferases. These groups exhibited appealing druggability features, including less than 40% sequence identity to the human host proteome, predicted essentiality, structural classification as highly druggable or druggable, and expression levels above the 50th percentile in the amastigote form. The resources presented in this work also represent a comprehensive collection of integrated data regarding trypanosomatid biology.

1. Introduction

Leishmaniasis is a vector-borne disease caused by the infection with *Leishmania* spp., an obligate intracellular protozoa parasite. Fifty-one species comprise the *Leishmania* genus, of which at least 21 are pathogenic. The life cycle of these microorganisms involves infection stages within the vertebrate host (mammals) and in the midgut of the sand fly vector (Diptera: Phlebotominae) (Akhoundi et al., 2017; Bekhit et al., 2018).

Leishmaniases are classified into three main clinical forms: cutaneous leishmaniasis (CL), mucocutaneous leishmaniasis (MCL) and visceral leishmaniasis (VL). Each form is a result of the interplay among *Leishmania* and sand fly species, in addition to the state of the host immune system. Leishmaniasis is found mainly in low-income regions with

large vulnerable populations at risk of disease, being the second leading cause of death in the group of neglected tropical diseases (NTDs) (Giraud et al., 2019). Globally, the incidence of VL and CL lies between 0.2-0.4 and 0.7–1.2 million cases per year, respectively, and up to 40,000 deaths are estimated each year (Alvar et al., 2012; Mann et al., 2021).

Currently, no human vaccines against leishmaniasis are available. Treatment relies on the use of non-specific drugs that typically have an immunomodulatory activity on the innate immune system, but the limited efficacy, strong adverse effects, emerging parasite resistance, and high costs for some formulations are factors that highlight the urgent need to develop new, safer, and more effective drugs to control this NTD (Bekhit et al., 2018; Mann et al., 2021). For instance, the pentavalent antimonial, the first antileishmanial medicine discovered in the 1940s, is still used for VL and CL treatment. However, this drug class shows increasing resistance and high toxicity, sometimes leading to death (Bekhit et al., 2018). Paromomycin, an aminoglycoside antibiotic, demonstrates efficacy against CL and VL in Old World species, particularly in the Indian subcontinent. Yet, its effectiveness is limited in New World *Leishmania* species (Roatt et al., 2020).

Amphotericin B encapsulated in liposomes solves part of the problems common to antileishmanial drugs, being highly effective in VL and requiring a shorter treatment period that reduces adverse effects and dropout. Nevertheless, its high-cost limits widespread deployment in poorer regions, which are usually more burdened by the disease. The first drug to be administered orally to treat leishmaniasis was miltefosine, but its elevated cost, resistance and teratogenic potential also limited ample usage (Braga, 2019; Cruz et al., 2009). Therefore, currently, there are no affordable and effective drugs available that have minimal adverse effects and that can be used to treat all leishmaniasis manifestations, as aimed by the Drugs for Neglected Diseases initiative (DNDi) (Pécoul, 2004).

The complete pathway for drug discovery is expensive, high-risk and time-consuming (Fernández-Prada et al., 2019), comprising several steps: target discovery, identification of drug candidates, and phases 1–3 clinical trials. Unfortunately, most candidates fail in the clinical trials phase (Duelen et al., 2019). The primary approaches in drug discovery are phenotypic and target screens. Phenotypic screens identify compounds with therapeutic potential based on observed biological effects, yielding promising candidates such as the kinetoplastid proteasome inhibitor (Khare et al., 2016; Roquero et al., 2019). Combining these screens with target-based approaches increases the drug candidate pool, potentially expediting the drug discovery processes (Heilker et al., 2019). This pipeline can be optimized by applying rational computer-aided methods in the initial stages of target discovery, reducing the molecular space of candidates following to *in vitro* screenings, and increasing the probability of finding a molecule that serves as a target. *In silico* screenings can also support the repositioning of already-approved drugs, for which the safety and pharmacological profiles have been previously established, avoiding common problems in the clinical study as well as later stages of drug discovery (Chávez-Fumagalli et al., 2019; Jamal et al., 2017; Ramos et al., 2018).

By leveraging the recent availability of full-length genome sequences for various *Leishmania* species, here we identified potential protein targets common to clinically-relevant species. We applied a subtractive proteomics and comparative genomics approach that integrated high-throughput multi-omics data to classify targets based on druggability metrics. Our aim was to identify potential targets with desirable characteristics for the development of novel compounds or the repositioning of currently available ones to effectively treat this overlooked disease.

2. Materials and methods

Classifying protein targets involves identifying desirable features from a drug discovery perspective (Field et al., 2017), which includes their presence in relevant *Leishmania* species (conservation); their potential to be inhibited by drug-like compounds (structural druggability);

unwanted toxicity effects by inhibiting homologous proteins in the human host (selectivity); importance for the parasite metabolism (essentiality); and target accessibility (relevant life-stage gene expression and subcellular localization). The methods used to access each of these characteristics are addressed in the following subsections and overviewed in Fig. 1A.

2.1. Selection of clinically relevant Leishmania proteomes

Fourteen proteomes of pathogenic *Leishmania* strains were selected based on the current taxonomic classification (Akhoundi et al., 2016), composed of: two strains of *L. braziliensis*, four of *L. donovani*, one of *L. amazonensis*, one of *L. infantum*, three of *L. major*, one of *L. mexicana*, and two of *L. panamensis* (Table 1). These data were retrieved from TriTrypDB (Aslett et al., 2010) and Leish-ESP (http://leish-esp.cbm.uam.es) projects. Accessions, sources and characteristics of the genomes used in this study are shown in Table 1.

2.2. Identification of ortholog groups among clinically relevant Leishmania

The standalone version of OrthoVenn2 (Xu et al., 2019) was used for the construction of orthologous protein groups among the 14 *Leishmania* studied species. The inflation value and *E-value* parameters were set to 1.5 and 10^{-10}, respectively. Only groups with at least one protein representative of each species were kept.

Once the proteins conserved among all 14 clinically-relevant *Leishmania* species were identified, a smaller set of representative species were used for downstream analyses, based on the availability of expression and protein structure data. Thus, in subsequent prioritization stages only proteins from the following set of species were used: *L. infantum* JPCM5, *L. major* Friedlin, *L. braziliensis* M2904, and *L. donovani* HU3 (bolded accessions in Table 1).

2.3. Identification of proteins non-homologous to the host

To lessen potential cross-reactions due to structural conservation between the target proteins of the pathogen and those of the host, local BLASTp alignments were performed (Altschul, 1997). The decision to prioritize global protein similarity over active sites is aimed at accommodating both orthosteric (active sites) and allosteric pockets within the scope of druggability prediction. The conserved protein set was used as a query against a database formed by the predicted human proteome (assembly GRCh38.p13 obtained from Ensembl <http://www.ensembl.org/>). The parameters used to determine homology between sequences were *E-value* $\leq 10^{-5}$, identity ≥40% and bit score ≥100. Homologous proteins between *Leishmania* and humans were excluded from the study. For proteins that passed the previous filter, the *off_target_human* metric, ranging from 0 to 1, was created from the BLASTp identity results. This metric was calculated as: *off_target_human* $= 1 - \left(\frac{Ih}{100}\right)$, *Ih* being the best hit identity percentage. The closer to 1, the lower the sequence similarity query with the database.

2.4. Target accessibility: expression at relevant life-cycle stages

The transcriptome profiles in the amastigote life stage of the four *Leishmania* species were analyzed to infer target accessibility. The amastigote stage was aimed at because of its relevance during human infection. An RNA-seq pipeline developed in our previous works was used to analyze transcriptome data (Azevedo et al., 2020). Searches for expression datasets published on the Sequence Read Archive were performed with the following filters: "methodology" equal to RNA-seq, "species" equal to *Leishmania*, and the keywords "intracellular amastigote*" or "axenic amastigote*". After manual inspection, the datasets were downloaded locally, and read quality control was performed using
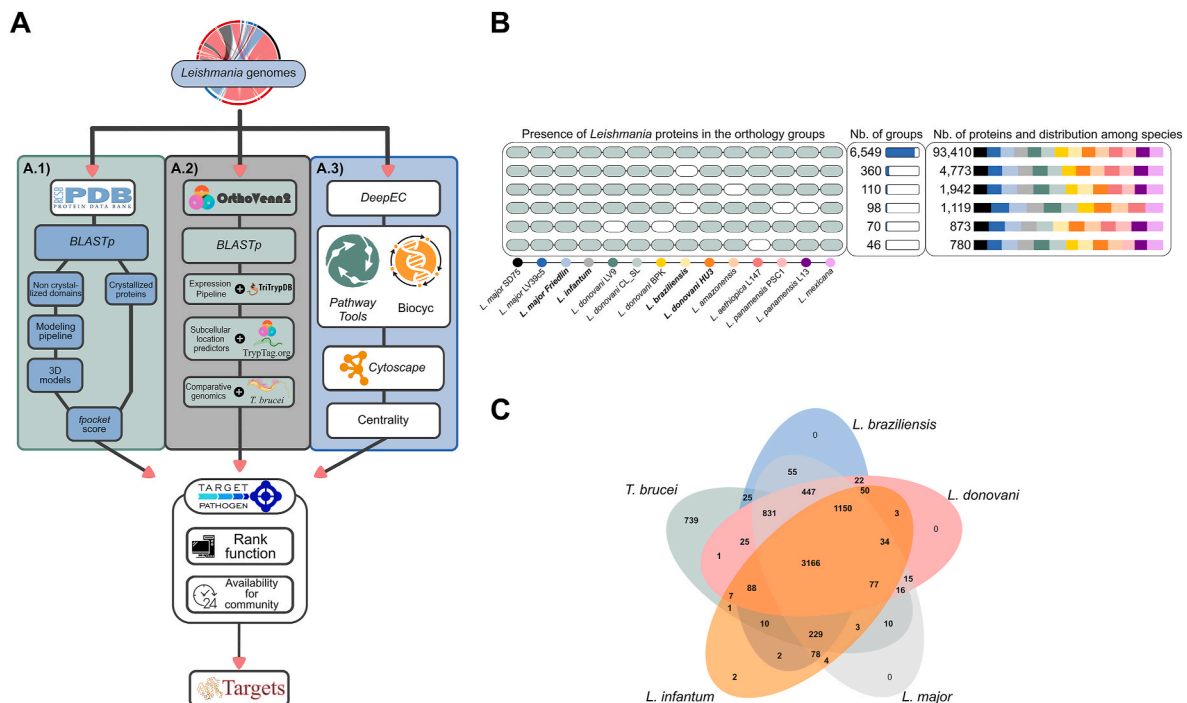
Fig. 1. A) Overview of the prioritization approach, involving (A.1) the computation of structural druggability from 3D models of Leishmania proteins, (A.2) inference of orthology groups, proteins non-homologous to the host (BLASTp), amastigote-expressed genes, and subcellular localization, along with (A.3) protein essentiality assessment. The integrated data and prioritized proteins are accessible via the Target-Pathogen web resource. B) Representation of orthology protein groups across the 14 Leishmania species. The leftmost rectangle categorizes orthology groups based on the presence or absence of proteins in each species, with colored circles representing each Leishmania species at the bottom. Each cell indicates absence (white) or presence (green) in any species. The middle rectangle displays the count of orthology groups corresponding to the species' presence in the previous section. The last rectangle presents the total number of proteins and their distribution by species, indicated by color and size. C) Venn diagram illustrating orthology groups between Leishmania post-filter expression proteins and T. brucei homologs. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Characteristics and sources of *Leishmania* spp. genomes used in this study. High-quality genome assemblies are highlighted in bold.

Leishmania species	Source	Nb. of genes	Assembly state	Genome size (Mbp)	G + C content (%)
L. aethiopica L147	GeneDB	8826	36 chromosomes, 789 supercontig	32.2	57.88%
L. amazonensis MHOM/BR/71973/M2269	GenBank	8127	2627 supercontig	29.0	59.21%
***L. braziliensis* MHOM/BR/75/M2904**	Leish-ESP	8493	36 chromosomes	32.3	58.05%
L. donovani BPK282A1	GeneDB	8135	36 chromosomes	33.0	56.37%
L. donovani LV9	Sanger	8387	36 chromosomes, 108 contig	35.4	58.39%
L. donovani CL-SL	GenBank	9757	36 chromosomes	33.5	58.77%
***L. donovani* HU3**	Leish-ESP	8595	36 chromosomes	33.0	59.77%
***L. infantum* JPCM5**	RequenaLab	8748	36 chromosomes	33.3	58.77%
***L. major* Friedlin**	GeneDB	9378	36 chromosomes	33.4	58.74%
L. major LV39c5	GeneDB	9101	36 chromosomes, 850 supercontig	32.9	57.76%
L. major SD 75.1	GeneDB	8931	36 chromosomes, 36 supercontig	31.8	58.38%
L. mexicana MHOM/GT/2001/U1103	GeneDB	8677	36 chromosomes, 554 contig	32.6	58.76%
L. panamensis MHOM/PA/94/PSC-1	GenBank	8048	36 chromosomes	31.2	55.33%
L. panamensis MHOM/COL/81/L13	GeneDB	8778	36 chromosomes, 1980 supercontig	31.8	56.13%

FastQC (Andrews, 2010). Low-quality sequences and adapters were removed using bbduk v. 38.22, with minlen parameters equal to 70% of the study read size, trimq (trimming quality using Phred) equal to 30, and the tpe and tbo flags (Bushnell et al., 2017). Each read set was mapped against the reference *Leishmania* genome (bolded rows in Table 1) using HISAT2 v. 2.1.0 using default parameters (Kim et al., 2019). Count tables were produced using FeatureCounts v. 1.6.3 (Liao et al., 2014), and normalization of expression values were performed using edgeR with the Trimmed Mean of M-values (TMM) method (Robinson et al., 2010), followed by transformation into counts per million (CPM), along with expression percentiles for each gene. To avoid biases, all datasets retrieved from RNA-seq studies were submitted to the same integrated pipeline. Transcripts with less than 1 CPM in each sample were considered as not expressed in amastigotes and removed from the study. For species that lacked transcriptomic data generated by RNA-seq, a previously analyzed microarray expression dataset was retrieved from TriTrypDB (Aslett et al., 2010). The expression values from this study were normalized to percentiles. Only genes above the 20th percentile of expression were kept.

2.5. Target accessibility: subcellular localization

Localization information is helpful to inform the development of therapeutics, including targets for therapy. We collated or predicted subcellular localization data to provide contextual information of each target but did not use this information for ranking purposes (section 2.9), given that candidate molecules can be altered, optimized or conjugated with chemical probes to reach specific compartments within the

cell.

To overcome the lack of large-scale experimental studies on the subcellular localization of *Leishmania* proteins, two levels of evidence were explored for inferring these locations: 1) Hybrid experimental/computational and 2) Strictly computational.

2.5.1. Hybrid experimental/computational localization inference

Initially, large-scale subcellular localization data for *Trypanosoma brucei brucei* TREU 927, produced by the TrypTag project (Billington et al., 2023) using nNeonGreen labeling followed by microscopy, was used to infer protein localization. Since *T. brucei* also belongs to the Trypanosomatidae family, present similar morphology and share many general characteristics with *Leishmania* representatives (Halliday et al., 2019; Tibayrenc and Ayala, 2013), this dataset was selected to inform, through comparative genomics, the subcellular location in *Leishmania* as well. Orthology groups were calculated between *Leishmania* and *T. b. brucei* TREU 927 (TriTrypDB version 44). For this task, OrthoVenn2 was used with the parameter's inflation value 1.5 and $E\text{-}value \le 10^{-5}$. Localization data from *T. brucei* proteins belonging to an ortholog group was then propagated to *Leishmania* proteins falling in the same group. The data provided by TrypTag were classified using the controlled vocabulary of the Gene Ontology (GO) (Consortium and Gene Ontology Consortium, 2004) to define the cellular location of the N- and C-terminal portions of the proteins, and there may be more than one GO term from these independent experiments. The GO terms were manually normalized to parent terms to facilitate their use in this work, as well as to make the comparison of the classification obtained using this method with those used by the "strictly computational" predictors (described in the following subsection) more direct. Information about the normalization of terms is shown in Supplementary Table 1.

2.5.2. Strictly computational localization inference

For cases in which it was not possible to obtain a classification based on the experimental data, a consensus of multiple computational evidence was calculated. For this, the following predictors were used: DeepLoc v. 1.0 (Almagro Armenteros et al., 2017), MultiLoc2 (Blum et al., 2009), WoLF PSORT (Horton et al., 2007), TargetP v. 2.0 and SignalP v. 5.0 (Almagro Armenteros et al., 2019). These tools were run at their default parameters using as input the sets of protein files in FASTA format. Each tool uses different terms, which may vary in specificity of their locations, and there may be overlaps between them. The term sets used by each tool were normalized (Supplementary Table 2) to standardize and simplify the data, making results from each tool comparable. For each protein, the normalized terms that had the highest counts were selected. In addition, manual annotations of Gene Ontology terms related to localization for *Leishmania* proteins at Uniprot (https://www.uniprot.org) were also used.

2.6. Inference of protein essentiality

Two complementary approaches were used to infer the essentiality of *Leishmania* proteins, first relying on data made available for *T. brucei*, as well as inference using an integrated whole-metabolic approach.

2.6.1. Processing of T. brucei RNA interference high-throughput data

High-throughput phenotyping data using RNA interference (RNAi) is accessible for *T. brucei* and was employed to infer essentiality (Alsford et al., 2011). Out of the 7435 *T. brucei* proteins examined across four life-cycle stages in bloodstream and insect host, 750 were associated with the loss-of-fitness phenotype in all life-cycle stages when their expression was silenced. The orthology classification from the previous methodology step (**2.5**) was then used to propagate essentiality information to the orthology group shared between *Leishmania* and the 750 proteins classified as essential for *T. b. brucei* TREU 927.

2.6.2. Reconstruction and topological analysis of a Leishmania genome-scale metabolic model

Complementarily, a genome-scale metabolic network for the four *Leishmania* species was reconstructed using a semi-automated approach followed by manual validation. The PathoLogic module of the Pathway Tools v. 18.0 (Karp et al., 2002) software was used during this step. PathoLogic takes an annotated genome in Genbank format as input and creates a Pathway/Genome Database (PGDB) containing all the predicted metabolic pathways of the organism. *Leishmania* annotation files were obtained in GFF format from TriTrypDB and Leish-ESP. The annotations were enriched with Enzyme Commission (EC) numbers using DeepEC (Ryu et al., 2019), a computational framework based on deep learning that predicts EC numbers with high precision. After semi-automatic prediction of metabolic pathways, manual curation was performed by the inclusion and removal of reactions based on evidence from the literature and the overview graph of the pathway. Finally, the genome-scale metabolic model was exported in SBML format and converted into a reaction graph (Ramos et al., 2018). We also calculated choke-point reactions (or bottlenecks), reactions that uniquely consume or produce a given substrate or product, respectively. Inhibition of choke-point reactions could lead to toxic buildup of a substrate or a lack of a product essential for the survival of the parasite. Betweenness centrality was calculated for each reaction (node) in the graph and used as a prioritization criterion. Reactions with high centrality values function as bridges to different parts of the metabolic network and represent potentially attractive activities to be targeted. Visualization and calculation of topological metrics were performed using Cytoscape v. 3.8.2 (Shannon et al., 2003).

2.7. Generation of 3D models for Leishmania proteins based on structural homology

Crystal structures for proteins from the four *Leishmania* species selected in earlier steps were sought and retrieved, in PDB format, from the Protein Data Bank (PDB [http://www.rcsb.org],). For cases in which no experimentally-determined structure was available in the PDB, homology-based models were built using a structural genomic pipeline (Radusky et al., 2014). Protein sequences were used as input for searches using the PSI-BLAST algorithm (parameters -j 3, $E\text{-}value < 10^{-5}$) against UniRef50 (UniProt at 50% redundancy). Once a position-specific score matrix (PSSM) was obtained, it was used to search for queries against PDB95 (non-redundant PDB at 95% conservation) using PSI-BLAST (parameter $E\text{-}value < 10^{-5}$). Up to five retrieved PDB models were used as a reference for homology-based reconstruction, and MODELLER (Webb and Sali, 2016) was applied to build five models per template for each structureless target protein in the PDB. A representative model was chosen based on the GA341 score (>0.7) and the maximization of the QMEAN Z-score function.

2.8. Classification of Leishmania proteins according to their druggability

The operational definition of druggability followed in this work associates the ability of a peptide to bind to a drug-like compound, ultimately resulting in its modulation in a desired manner, such as inhibiting its molecular activities (Keller et al., 2006). Thus, inferring druggability of a protein requires that its surface harbors desirable features, particularly druggable cavities. fpocket was used for predicting protein cavities and estimating druggability scores (DS) for each identified cavity (Le Guilloux et al., 2009). Cavities were classified into 4 categories following previous works (Radusky et al., 2014): non-druggable (ND; DS < 0.2), poorly druggable (PD; $0.2 \le DS < 0.5$), druggable (D; $0.5 \le DS < 0.7$) and highly druggable (HD; $0.7 \le DS < 1.0$). All proteins for which structures were obtained (either from the generated models or experimentally-determined structures) were subjected to this classification. The representative DS of a protein was defined as the maximum DS of any of its predicted cavities.

2.9. Protein prioritization

All features described in the previous sections identified for the set of *Leishmania* proteins, except subcellular localization, were integrated into the Target-Pathogen (TP) webserver and are publicly available at http://target.sbg.qb.fcen.uba.ar/ (Sosa et al., 2018).

To classify proteins in terms of their potential to serve as a target based on the generated data, a Ranking Function of Druggable Targets (RFDT) was defined:

$$RFDT(prot) = \frac{Drg + H + \frac{Ep}{100} + Chk + Ct + Ess_{Tb}}{6}$$

where *Drg* represents the DS, *H* represents the value of the *off-target_human* variable; *Ep* is the expression percentile in the amastigote stage; *Chk*, a binary variable representing whether the protein is a chokepoint; *Ct* is the normalized betweenness centrality in the reaction graph; *Ess_{Tb}*, a binary variable representing whether the *T. brucei* ortholog (if identified) was essential. The RFDT was ordered decreasingly, and the top targets were evaluated in-depth.

To characterize the top-scoring proteins annotated as hypothetical, we employed a structural homology methodology. Initially, we utilized the structural predictions of an AlphaFold-based database trained on the Discoba group to generate predicted protein structures (Wheeler, 2021). Subsequently, these predicted structures underwent a structural similarity search in the PDB database, with the usage of mmCIF, strict search criteria, and chain structure as parameters (Burley et al., 2017). The identified structures were then downloaded and aligned against the previously modeled structure of the protein using USalign (C. Zhang et al., 2022). The template modeling score (TM-score) was employed to evaluate the quality of the structural alignment. Additionally, Pfam was employed to explore associated domain families (Paysan-Lafosse et al., 2023). The generated structures were also submitted to the FoldSeek webserver (van Kempen et al., 2024) using 3Di/AA mode in order to determine structurally similar proteins deposited in UniProt50, Swiss-Prot, Proteome, CATH50, MGnify, PDB100, and GMGCL databases.

3. Results

3.1. Constructing a data-rich compendium of Leishmania proteins

3.1.1. Pathogenic conservation

To identify conserved proteins among the 14 proteomes from pathogenic *Leishmania*, orthology groups were established and annotated using OrthoVenn2 (Xu et al., 2019). From the initial set consisting of 118124 proteins for these pathogens, 93410 reached the minimum thresholds to fall into 'core' orthology groups, yielding 6549 groups that had at least one protein of each studied species (Fig. 1B). This results in 79% of the proteins of all *Leishmania* species being orthologous to each other, in agreement with previous studies (Bansal and Meyer, 2002; Dávila López et al., 2010). The remaining 684 orthology groups lacked representatives of at least one species and were excluded from further analyses. Once the core orthology groups were established, proteins from the four species chosen as references in this work, namely *L. infantum* JPCM5, *L. major* Friedlin, *L. braziliensis* M2904 and *L. donovani* HU3 were further advanced into the prioritization pipeline.

3.1.2. Filtering off-target proteins

Proteins were filtered for similarity with the human proteome to avoid potential adverse effects on the host. Forty-six orthology groups with homologs in humans were removed from the 6549 groups obtained from the previous step, resulting in 6503 groups for further analysis.

3.1.3. Expression in relevant life-cycle stages and subcellular locations

Four studies were selected from the Sequence Read Archive (SRA/ NCBI) for gene expression analysis using RNA-seq technology. These studies included samples of intracellular and axenic amastigote forms from four *Leishmania* species: *L. major* Friedlin (Fernandes et al., 2016) (BioProject PRJNA290995), *L. donovani* evaluated in two studies (Cuypers et al., 2017) (BioProject PRJNA375925) and (BioProject PRJNA2915321), *L. braziliensis* M2903 (BioProject PRJNA494068). Expression data for *L. infantum* were obtained using microarray technology (Rochette et al., 2009) (BioProject PRJNA110243). Metadata for all RNA-seq studies are presented in Supplementary Table 3. After normalization and transformation of expression data, 2270 non-unique proteins were removed, advancing 6502 orthology groups for further analysis.

By analyzing the subcellular localization of *Leishmania* proteins, we were able to enrich the annotations of their positions within the subcellular compartment, shedding light on their organization and functional roles within the cell. Orthology groups were constructed between *Leishmania* proteins and the *T. b. brucei* TREU 927 proteome (composed of 11203 proteins). Of these, 157 groups were composed of proteins from *T. b. brucei* TREU 927 without a defined location by the TrypTag data, which were removed from the analysis, yielding 4427 groups. The subcellular location annotations of *T. b. brucei* TREU 927 proteins were propagated to *Leishmania* proteins in the same group to determine their subcellular localization (Fig. 2A). The consensus of subcellular prediction tools revealed that 16.3% (3884) of the *Leishmania* proteins were classified as "Cell Periphery/Secreted," with the most frequently observed term being "Nucleus", accounting for 34.7% of the total. These results are consistent with findings from previous studies in eukaryotes (Barylyuk et al., 2020; Fagerberg et al., 2011) (Fig. 2B).

3.1.4. Protein categorization according to their essentiality

Data for *T. brucei* protein phenotyping was used as a reference to determine essentiality of *Leishmania* proteins (Alsford et al., 2011). 4591 orthology groups containing at least one protein from *Leishmania* species and one of *T. brucei* were disclosed (Fig. 1C). This approach led to the identification of 1399 *Leishmania* proteins (377 groups) that were defined as essential in our study.

Metabolic networks of *Leishmania* species were reconstructed and analyzed as graphs to calculate topological metrics. Table 2 shows the number of enzymes and pathways found, while Fig. 2C displays the metabolic pathway of *L. infantum* as a graph. 3683 proteins were classified as choke points, with each species varying in the number of annotated enzymes, possibly because of different sequencing technologies and annotation tools used. The Pathway/Genome Database (PGDB) files, allowing reuse of this model, can be found in the Data Availability section.

3.1.5. Classification of proteins according to their druggability

Structural data from the PDB and homology-based models were used to identify potential protein cavities and their corresponding DS for *Leishmania* species. As expected, most proteins did not have crystallized structures deposited in the PDB, with the vast majority of structures obtained from the modeling pipeline (Table 3). By means of comparison, the DS of all ligand-bound structures present in PDB95 (non-redundant subset) were calculated. The majority of *Leishmania* proteins were classified as highly druggable, with an enrichment of DS in their predicted pockets compared to the PDB95 subset (Fig. 2D).

3.1.6. Target prioritization

In this study, 6502 orthology groups of *Leishmania* proteins were identified and ranked based on their potential to serve as a target. The ranking was determined by considering various factors such as conservation, absence of homologs in the human genome, and expression in the pathogenic form of the parasite. The data were used as variables in the RFDT function to calculate scores for each protein.

All 23897 proteins that passed through the filter steps were mapped to their orthology groups. The top 20 orthology groups with highest
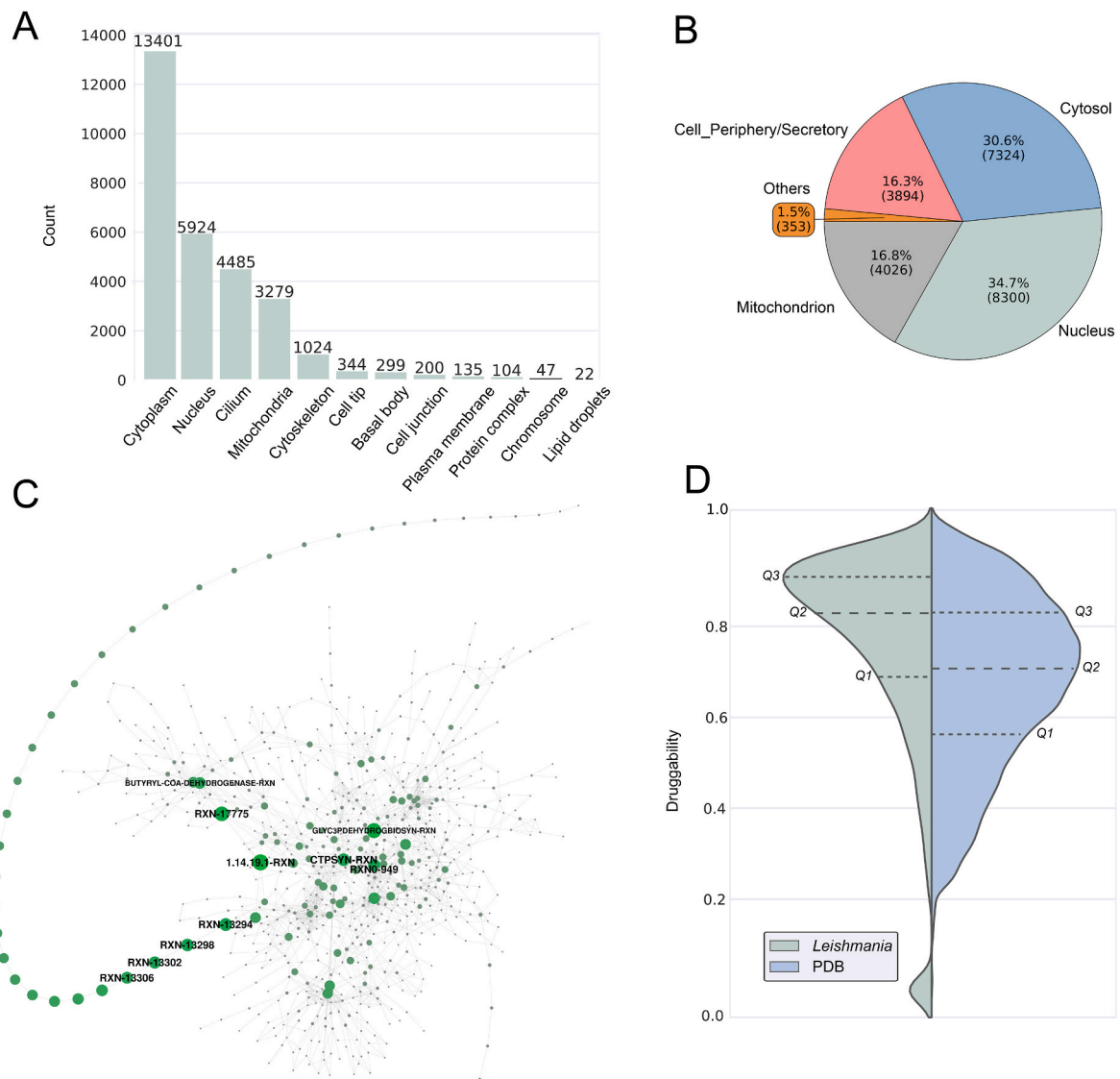
Fig. 2. A) Bar graph illustrating the count of terms associated with each protein that passed the filtered steps, normalized by method 1 (experimental + computational). *Nlm refers to the term intracellular non-membrane-bounded organelle. B) Frequency of subcellular location terms associated with each protein that passed the filtered steps, determined by the consensus of strictly computational tools. C) Metabolic network of *L. infantum* JPCM5. Nodes represent pathway reactions, with edges between node pairs indicating that the product of one reaction is used as a substrate for the next reaction. Node size is proportional to the betweenness centrality. The 10 most central reactions are labeled in figure. D) Violin chart displaying druggability values. The curve for *Leishmania* represents the druggability score distribution of all structural models for the four studied *Leishmania* species, while the curve for PDB represents the druggability score distribution of all proteins bound to a compound in the PDB. Quartiles are represented by: Q1 (1st quartile), Q2 (2nd quartile, median), and Q3 (3rd quartile).".

Table 2

Metabolic reconstruction metrics.

Metabolic reconstruction metrics	*L. infantum*	*L. major*	*L. donovani*	*L. braziliensis*
Pathways count	137	140	104	103
Enzymes count	1152	1174	845	843

Table 3

Metadata from structural models for Leishmania proteins generated in this work, using the methodology of reconstruction by structural homology.

Genomes	Nb. of crystal structures available	Predicted structural models
L. braziliensis M2904	117	4289
L. donovani HU3	260	4438
L. infantum JPCM5	107	4682
L. major Friedlin	121	6535

average RFDT are presented in Table 4 along with the biological process and most frequent protein product associated. In Supplementary Table 4, the information regarding the top 100 clusters as determined by RFDT are presented, providing comprehensive details about the constituent proteins. For ortholog groups, the average of RFDT scores were calculated. The complete set of proteins that compose the top 20 orthology groups present attractive druggability features: they have less than 40% sequence identity to the human host proteome; were predicted as essential; the majority were structurally classified as highly druggable

or druggable; and are expressed above the 50th percentile of expression in the amastigote form.

Beyond the top-ranked proteins, we extended our investigation to include proteins annotated as hypothetical, which have remained poorly characterized in the existing literature. Our aim was to gather additional

Table 4
Top 20 orthology clusters with highest RFDT, representing groups of proteins with high potential to be targets for the search of new or repositioning drugs for the treatment of pathogenic *Leishmania* spp.

Aggregated rank[a]	Cluster ID	Process	Common product[b]	Presence in the 4 selected species[c]			
				L. major	*L. infantum*	*L. braziliensis*	*L. donovani*
1	cluster3320	Protein ubiquitination	Ubiquitin-activating enzyme E1	✓	✓	✓	✓
2	cluster6048	Ubiquitin-dependent protein catabolic process	Ubiquitin-conjugating enzyme E2	✓	✓	✓	✓
3	cluster5827	Seryl-tRNA aminoacylation	Seryl-tRNA synthetase	✓	✓	✓	✓
4	cluster3210	Pentose-phosphate shunt	6-phosphogluconate dehydrogenase decarboxylating	✓	✓	✓	✓
5	cluster332	Asparaginyl-tRNA aminoacylation	Asparaginyl-tRNA synthetase		✓	✓	✓
6	cluster945	Purine nucleotide biosynthetic process	Adenylosuccinate lyase	✓	✓	✓	
7	cluster3744	Tricarboxylic acid cycle	2-oxoglutarate dehydrogenase E2 component dihydrolipoamide succinyltransferase	✓	✓	✓	✓
8	cluster61	Prolyl-tRNA aminoacylation	Prolyl-tRNA synthetase	✓	✓	✓	✓
9	cluster5049	Phosphatidylethanolamine biosynthetic	Choline kinase	✓		✓	✓
10	cluster5561	Protein ubiquitination	Ubiquitin-activating enzyme E1	✓	✓	✓	✓
11	cluster3806	Absence of annotation	Cytochrome *c* oxidase subunit VI	✓	✓	✓	
12	cluster5273	Nicotinate nucleotide biosynthetic process	Nicotinatephosphoribosyltransferase	✓	✓	✓	✓
13	cluster4732	Ubiquitin-dependent ERAD pathway	Ubiquitin-conjugating enzyme E2		✓	✓	✓
14	cluster3098	Protein K63-linked ubiquitination	Ubiquitin-conjugating enzyme E2	✓	✓	✓	✓
15	cluster1388	Protein ubiquitination	Ubiquitin-conjugating enzyme E2	✓	✓	✓	✓
16	cluster1475	Respiratory chain complex IV assembly	Cytochrome *c* oxidase assembly factor	✓	✓	✓	✓
17	cluster1599	Protein neddylation	Ubiquitin-activating enzyme	✓		✓	✓
18	cluster4417	Ubiquitin-dependent protein catabolic process	Ubiquitin-conjugating enzyme E2	✓	✓	✓	✓
19	cluster5524	Protein ubiquitination	Ubiquitin-conjugating enzyme	✓	✓	✓	✓
20	cluster4611	Unsaturated fatty acid biosynthetic process	Fatty-acid desaturase	✓	✓	✓	✓

[a] , an average of the RFDT score of the proteins that formed the orthology cluster.
[b] , the protein product most frequently listed in the cluster.
[c] , the presence of proteins representing the four species of *Leishmania* studied present in the clusters that passed the filters defined in the methodology.

information that might position these proteins as potential targets. Among the top 100 clusters (Supplementary Table 5), we identified three clusters comprising hypothetical proteins, ranked 25th, 84th, and 91st, respectively. To augment our understanding, we employed Pfam for exploring associated domain families, executed structural predictions, conducted homology assessments, and carried out functional annotations utilizing AlphaFold, US-align and PDB data. These analyses were further enriched with insights from *Leishmania* proteomics dataset (Burley et al., 2017; Paysan-Lafosse et al., 2023; Wheeler, 2021; C. Zhang et al., 2022).

4. Discussion

To accelerate the identification of target proteins for the development of novel antileishmanial drugs, extensive metadata related to pathogenic *Leishmania* species was mined, generated, processed, linked, and classified. The result was the creation of a digital collection, specifically designed for the scientific community, including chemists, with the aim of advancing knowledge in the field of neglected diseases and drug development against *Leishmania* parasites.

Among the top 20 highly ranked candidates, nine were found to be associated with metabolic processes and common products in the ubiquitination pathway (Table 4). Additionally, three were associated with tRNA aminoacylation, two with electron transport chain, as well as pentose-phosphate pathway (PPP) and purine synthesis. These processes are frequently recognized as targets in the literature and some have already been the focus of proposed formulations or ongoing studies for combating infections (Bijlmakers, 2020; Boitz et al., 2013; Fidalgo and Gille, 2011; Jakkula et al., 2021; Luque-Ortega and Rivas, 2007; Rubio Gomez and Ibba, 2020). This finding supports the effectiveness of *in silico* prioritization in predicting biologically meaningful target proteins. Next, we detail some of these biological processes, as well as propose new targets disclosed by the integrative methodology.

4.1. Ubiquitination pathway

Ubiquitination is a post-translational protein modification process in which one or more ubiquitin molecules (Ub) are covalently attached (tagged) to a substrate protein. It is an essential mechanism of eukaryotes that regulates a variety of cellular processes and plays an important role in several human diseases (Popovic et al., 2014). The best-known function is the quality control of defective proteins by the proteasome degradation system (John Mayer et al., 2008; Karin and Ben-Neriah, 2000; Zinngrebe et al., 2014). Three enzymes participate in the ubiquitination pathway: ubiquitin activating enzyme (E1 or UBE), ubiquitin conjugating enzyme (E2), and ubiquitin ligase (E3). *Leishmania* possess a complete repertoire of orthologous genes for the synthesis of E1, E2, and E3 enzymes, making this pathway an attractive target for anti-leishmanial drug development, especially considering its essentiality in other trypanosomatids (Alsford et al., 2011; Bijlmakers, 2020; Burge et al., 2020; Gupta et al., 2018). Compound TAK-243 is the first inhibitor of human UBA1, targeted for therapeutic treatment of some types of cancer (Barghout et al., 2019; Borgo et al., 2022; X. Zhang et al., 2022). The repositioning of TAK-243, with few modifications in structure, was suggested for the inhibition of UBA1 orthologs in *T. brucei, T. cruzi* and *Leishmania* (Boer and Bijlmakers, 2019).

4.2. tRNA aminoacylation

Another class of proteins that were ranked in the top 20 are aminoacyl-tRNA synthetases (aaRSs). These widely distributed enzymes are responsible for catalyzing the first step of protein synthesis, the binding of tRNA with its cognate amino acid, guided by the tRNA triplet anticodon (Pang et al., 2014; Rubio Gomez and Ibba, 2020). Recent research shows that aaRSs play a role in several cellular processes beyond translation, including antiviral immunity, inflammation and gene regulation (Katsyv et al., 2016; Mirande, 2017; Yakobov et al., 2018). Processes related to protein biosynthesis are one of the primary targets in the development of antimicrobial drugs, including the ribosomal production of proteins (Wilson, 2014), which places aaRSs among highly attractive targets. Marketed drugs, such as borrelidin (which inhibits threonine-tRNA synthetase), muropicin (a broad-spectrum antibiotic), and halofuginone (an antiprotozoal), have their mechanisms of action based on the inhibition of aaRSs activity through different effector mechanisms (Fang et al., 2015; Kwon et al., 2019; Rubio Gomez and Ibba, 2020). In *Leishmania*, aaRSs have been identified as a consistent target for drug development, due to their essentiality and the presence of pockets, with differences in the structure and number of catalytic domains compared to human aaRSs (Kelly et al., 2020; Kushwaha and Capalash, 2022; Pham et al., 2014).

Clusters assigned to aaRSs appeared at 3 positions in the top 20. These results support the validity of the methodology in identifying potential target proteins in the parasite and highlight the significance of diverse types of aaRSs as drug targets. To mitigate the possibility of resistance to aaRSs inhibitors, it is recommended to develop a compound or a combination of compounds that target various types of aaRSs. This multi-target approach can be achieved by utilizing the unique structures of *Leishmania*'s aaRSs (Kushwaha and Capalash, 2022).

4.3. Purine salvage

Leishmania are purine auxotrophs and rely on the host and vector's *de novo* pathways to obtain the purine ring, which is essential for their survival and growth. Furthermore, the parasite possesses the purine salvage pathway, facilitating the recycling and incorporation of acquired nucleosides, nucleotides, and nucleic acids into its purine pool. Given these characteristics, the purine salvage pathway has been focused by numerous studies and proposed as a potential drug target (Ansari et al., 2016; Boitz et al., 2012; Carter et al., 2008).

The enzyme adenylosuccinate lyase (ASL), which comprises the proteins at the 6th position, catalyzes the final step in the conversion of inosine monophosphate (IMP) into Adenosine monophosphate (AMP). It functions as a bottleneck in the purine salvage pathway of the parasite (Boitz et al., 2013; Tsai et al., 2007). ASL is a well-established protein target in the literature, and studies have shown that null mutants of ASL in *L. donovani* exhibit a restricted growth phenotype and a reduced capacity to infect murine macrophages (Boitz et al., 2013; Bora and Jha, 2020; Galina et al., 2017).

4.4. Pentose phosphate pathway (PPP)

6-phosphogluconate dehydrogenase (6PGD) proteins, which are ranked in the 4th position (Table 4), belong to a well-known class of enzymes. 6PGD plays a crucial role in the initial stages of the PPP, facilitating the catalytic conversion of 6-phosphogluconate (6 PG) into ribulose-5-phosphate, accompanied by the transformation of NADP + into NADPH. The former product plays an important role in preserving the redox balance, which is pivotal for intracellular pathogens like *Leishmania*, particularly in the amastigote stage. As a result, both 6PGD and the pathway as a whole have garnered attention as promising candidates for targeted drug development (Berneburg et al., 2023; Dhumal et al., 2022; Stincone et al., 2015; Van Assche et al., 2011). 6PGD has been demonstrated to be essential in *T. brucei* and various eukaryotes, establishing it as a viable drug target in numerous organisms (Alsford et al., 2011; Jakkula et al., 2021; Kerkhoven et al., 2013).

Despite its presence in humans, the crystal structure of 6PGD in *T. brucei* (Tb6PGD) revealed a distinct active site configuration (Phillips et al., 1998). The *Leishmania* 6PGD counterpart exhibits specificities that differentiate its structure and functionality. Notably, it lacks Gly-Ser-rich sequences at its C-terminal. This disparity stands as a significant argument favoring 6PGD as an antileishmanial drug target. In Tb6PGD, various inhibitor classes, including carbohydrate analogs, reaction intermediates, non-carbohydrate analogs, and triphenylmethane compounds, have shown notable specificity (Hanau et al., 2004). Moreover, there is the potential for selective inhibition through gold (I)-containing compounds, such as the approved drug auranofin (Berneburg et al., 2023; Jakkula et al., 2021).

4.5. Electron transport chain (ETC)

The ETC stands as one of the most extensively studied biochemical processes, playing a central role in converting proton electrochemical gradients into cellular energy, ultimately producing ATP. This process is of utmost significance in sustaining eukaryotic life, as it is responsible for generating energy that drives a variety of essential cell processes (Cogliati et al., 2021; Zhao et al., 2019).

At the core of this intricate system, cytochrome *c* oxidase (COX), also known as complex IV, serves as the terminal electron acceptor in the ETC, shuttling electrons to oxygen, leading to formation of water. COX is a multi-enzymatic complex composed of numerous subunits and assembly factors (Vercellino and Sazanov, 2021; Watson and McStay, 2020). COX biogenesis is a process that relies on several post-translational modifications, maturation, translocation and other finely tuned processes that are orchestrated by the assembly factors (Timón-Gómez et al., 2018). Although the complex structure and assembly are highly conserved across eukaryotes, several reported differences offer potential avenues for exploration in the pursuit of potential targets (Mansilla et al., 2018; Sen and Majumder, 2008; Watson and McStay, 2020).

In our results, we have identified two distinct clusters (positions 7 and 9 in Table 4) corresponding to a subunit and an assembly factor, respectively, of COX. In general, ETC complexes are often considered potential drug targets in intracellular parasites due to differences in their formation, composition, and functioning compared to human mechanisms (Fidalgo and Gille, 2011; Fisher et al., 2020; Sen and Majumder,

2008). In particular, miltefosine is a COX inhibitor, while other enzymes involved in the oxidative phosphorylation pathway have also been studied as targets, such as cyclobenzaprine (Lima et al., 2022).

4.6. Novel targets

In the previous sections, we discussed metabolic functions that have been previously proposed as druggable for other organisms, and now emerge as plausible targets for pathogenic *Leishmania* as well. However, some groups in our top ranked proteins have never been proposed as druggable, and thus represent a set of potentially novel targets that could be further explored in the trypanosomatid drug discovery pipelines considering the biological evidence presented herein.

The nicotinate phosphoribosyltransferase enzyme (NaPRT) catalyzes the conversion of nicotinic acid (Na) into nicotinic acid mononucleotide (NaMN), representing the initial step in the Preiss-Handler pathway for NAD^+ production, a molecule that participates in various biological processes, including electron transport, redox reactions, and cellular signaling (Gossmann et al., 2012). In the case of *Leishmania*, NAD^+ is an auxotrophic essential metabolite, being synthesized from precursors such as Na obtained from the host environment through the Preiss-Handler pathway, the exclusive route to synthesize NAD^+ in *Leishmania* (Gazanion et al., 2011). The cluster in the 8th position is composed of putative NaPRT enzymes found in *Leishmania* species. Although many enzymes involved in the production of NAD^+ have been identified as potential targets, NaPRT has received less attention despite its important role in NAD^+ synthesis (Gazanion et al., 2011; Ortiz-Joya et al., 2022; Singhal and Cheng, 2019). One possible reason for this is that NaPRT is conserved across eukaryotes, including humans, which may lead to unwanted cross-reactions. Therefore, it is imperative to conduct a thorough characterization of NaPRT, elucidating its functioning and coenzyme affinity, to ascertain its potential as a specific drug target.

The tricarboxylic acid cycle, also recognized as the Krebs cycle or citric acid cycle, plays a fundamental role in the energy metabolism of all oxidative organisms. Additionally, the TCA cycle is associated with the regulation of chromatin modifications, DNA methylation, and post-translational modifications of proteins (Akram, 2014; Martínez-Reyes and Chandel, 2020). However, a lesser-known role of the TCA cycle is its involvement in intracellular ROS homeostasis. This control is achieved through the regulation of key enzymes, such as the inhibition of α-ketoglutarate dehydrogenase (KGDH) complex resulting in the accumulation of α-ketoglutarate (KG), which acts as an antioxidant (Mailloux et al., 2007).

The cluster in the 4th position comprises dihydrolipoamide succinyltransferases (DLST), key components of the structural core of KGDH responsible for converting α-ketoglutarate into succinyl-CoA (Ranjan and Dubey, 2023). Remarkably, the DLST protein in *T. brucei* serves a dual function. In addition to its role in the TCA cycle, it participates in the tripartite attachment complex, a molecular structure crucial for kDNA segregation during cell division (Amodeo et al., 2021; Sykes and Hajduk, 2012). Insufficient knowledge is available concerning DLST function besides TCA cycle in *Leishmania*, and further characterization is necessary.

4.7. Clusters of interest among the top 100

The polyamine pathway is a well-established drug target for trypanosomatids, exerting significant influence on the cell growth, differentiation, cellular redox reactions and parasite proliferation (Singh et al., 2013). S-adenosylmethionine decarboxylase (AdoMetDc) plays a crucial role in this pathway (Supplementary Table 5), catalyzing the formation of decarboxylated S-adenosylmethionine. This compound acts as an aminopropyl group donor for the synthesis of spermidine from putrescine, a central metabolite in this pathway (Mishra et al., 2015). Knockout studies in *Leishmania* demonstrate that the deletion of AdoMetDc reduces infectivity compared to wild-type parasites (Roberts et al., 2002). Inhibitors for *Leishmania* AdoMetDc have been identified, such as 5-(((Z)-4-amino-2-butenyl)methylamino)-5-deoxyadenosine and methylglyoxalbis (guanylhydrazone) (Carter et al., 2022).

Leishmania are auxotrophic for purines, depending on the acquisition of their precursors from their host, making this metabolic pathway a crucial target for drug development. The enzyme 5′-methylthioadenosine phosphorylase (MTAP) plays a role in the recycling and salvaging of purine (Supplementary Table 5). MTAP cleaves 5′-methylthioadenosine into adenine and 5′-methylthioribose-1-phosphate, which are incorporated into the salvage pathways of purine (Abid et al., 2017; Bacchi et al., 1991). Moreover, MTAP is intricately linked with the methionine cycle and polyamine synthesis pathways. The accumulation of 5′-methylthioadenosine, a byproduct of the polyamine pathway, serves as an inhibitory factor, disrupting the normal progression of polyamine synthesis (Hofer, 2023). MTAP has emerged as a promising target for chemotherapy, resulting in the creation of specialized transition-state analogs (Abid et al., 2017; Sufrin et al., 1991).

Dihydrolipoamide dehydrogenase (DLD) orthologs also presented with desirable druggability features during our prioritization (Supplementary Table 5). DLD is a mitochondrial enzyme essential for energy metabolism across eukaryotes (Babady et al., 2007), exerting an antioxidant action through the scavenge of nitric oxide and reducing ubiquinone to ubiquinol (Xia et al., 2001). Mitochondrial DLD was found differentially abundant in drug-resistant *L. infantum* strains, suggesting its association with energy metabolism in resistant parasites (Vincent et al., 2015). Compounds aimed at modulating the function of DLD could be used as a potentiator of miltefosine and antimony effects.

Phosphoribosyl pyrophosphate (PRPP) is an essential metabolite for pyrimidine synthesis, and also participates in production of NAD, NADP, histidine and tryptophan, as well as the salvage pathway and the *de novo* synthesis of purines (Li et al., 2007). Phosphoribosyl pyrophosphate synthetase (Supplementary Table 5) catalyzes the synthesis of PRPP in *Leishmania* (Eriksen et al., 2000), and although these pathways have been recognized as druggable (J et al., 2021), this enzyme has not been characterized as a target in the literature. Our results suggest its potential to act as a drug target, with some limitations such as the presence of a human ortholog (albeit with low sequence identity, <30%) and the presence of possible paralogs in the *Leishmania* genome (in *L. major*, LmjF.08.1130 and LmjF.36.5390). Expression data analyzed here showed high levels of expression in the amastigote stage, the presence of druggable cavities, and a role as a choke point. Together, these results reveal a possibility that these groups of orthologs are proteins with a biologically relevant function that could be further explored.

4.8. Hypothetical proteins

All the *L. major* protein representatives from the 3 clusters annotated as hypothetical proteins (ranked 25th, 84th and 91st) were found to have an abundance greater than the 50th percentile in the promastigote form, as indicated by a quantitative proteomics study conducted by Polanco et al. (2022), potentially countering the hypothetical nature of these proteins.

The top 84 cluster exhibited a low TM-score (<2.7), indicating a limited similarity to the human augmin complex and the *S. cerevisiae* spliceosome (C complex). On the other hand, the top 25 and 91 clusters did not yield any structural similarity. While these outcomes provide evidence towards the bona fide nature of these proteins, as supported by proteomics data, additional investigations are imperative to elucidate their functions in *Leishmania*.

4.9. Limitations and strengths of this work

This work presents several previously recognized and new potential drug targets that could be further developed in the context of trypanosomatid drug discovery, although there are also several limitations and

challenges for these analyses. First, *Leishmania* biology presents with paucity of experimental data as compared to other pathogens, and only in recent years large-scale genomic resources have become available. To overcome this limitation, we relied upon extensive comparative genomics between *Leishmania* and other related parasites for which large-scale data, such as essentiality and localization, were already available, with the underlying rationale that biological characteristics and properties shared between these organisms reflect conservation among their genes and functions. When no experimental data were available, we relied upon computational tools to predict properties such as protein localization, structure, and essentiality, which also presents with their underlying challenges, specifically those that were trained using data outside the range of trypanosomatids (Sunter et al., 2023). For instance, caution is required in interpreting the essentiality data from *T. brucei*, considering the extrapolation of findings between distinct life stages and species. However, this strategy served as an alternative to address the limitations associated with the lack of high-throughput essentiality data for *Leishmania*. We applied a consensus-based approach to minimize the effects of biases in individual tools that perform protein localization, as well as employed different computational strategies to infer metabolic essentiality, coupled with manual pathway annotation. In addition to serving as a computational repository of target candidates for the scientific community, the resources presented in this work also represents a compendium of general knowledge of trypanosomatid biology.

Data Availability statement

All data used in this study are publicly available in the following web resources: Ensembl (http://www.ensembl.org), TriTrypDB (http://tritrypdb.org/tritrypdb/), Leish-ESP (http://leish-esp.cbm.uam.es), Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra), TrypTag (http://tryptag.org). This analysis can also be reproduced and customized through the web interface of Target-Pathogen (http://target.sbg.qb.fcen.uba.ar/patho/). The metabolic network described in section 2.6 is available in the Zenodo repository (https://zenodo.org/doi/10.5281/zenodo.10804883).

Funding

CRediT authorship contribution statement

Lucas G. Azevedo: Data curation, Investigation, Methodology, Software, Writing – original draft, Writing – review & editing. **Ezequiel Sosa:** Methodology, Software, Data curation. **Artur T.L. de Queiroz:** Resources, Writing – review & editing. **Aldina Barral:** Resources, Writing – review & editing. **Richard J. Wheeler:** Data curation, Resources, Writing – review & editing. **Marisa F. Nicolás:** Writing – original draft, Writing – review & editing. **Leonardo P. Farias:** Writing – original draft, Writing – review & editing. **Dario Fernández Do Porto:** Investigation, Methodology, Supervision, Writing – original draft, Writing – review & editing. **Pablo Ivan P. Ramos:** Funding acquisition, Investigation, Methodology, Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare no conflicts of interest to disclose. This research was conducted in an impartial and unbiased manner, and no external influences or competing interests have affected the design, execution, or reporting of this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpddr.2024.100538.

References

Abid, H., Harigua-Souiai, E., Mejri, T., Barhoumi, M., Guizani, I., 2017. Leishmania infantum 5′-Methylthioadenosine Phosphorylase presents relevant structural divergence to constitute a potential drug target. BMC Struct. Biol. 17, 9.

Akhoundi, M., Downing, T., Votýpka, J., Kuhls, K., Lukeš, J., Cannet, A., Ravel, C., Marty, P., Delaunay, P., Kasbari, M., Granouillac, B., Gradoni, L., Sereno, D., 2017. Leishmania infections: molecular targets and diagnosis. Mol. Aspect. Med. 57, 1–29.

Akhoundi, M., Kuhls, K., Cannet, A., Votýpka, J., Marty, P., Delaunay, P., Sereno, D., 2016. A historical overview of the classification, evolution, and dispersion of leishmania parasites and sandflies. PLoS Neglected Trop. Dis. 10, e0004349.

Akram, M., 2014. Citric acid cycle and role of its intermediates in metabolism. Cell Biochem. Biophys. 68, 475–478.

Almagro Armenteros, J.J., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, G., Elofsson, A., Nielsen, H., 2019. Detecting sequence signals in targeting peptides using deep learning. Life Sci. Alliance 2. https://doi.org/10.26508/lsa.201900429.

Almagro Armenteros, J.J., Sønderby, C.K., Sønderby, S.K., Nielsen, H., Winther, O., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33, 3387–3395.

Alsford, S., Turner, D.J., Obado, S.O., Sanchez-Flores, A., Glover, L., Berriman, M., Hertz-Fowler, C., Horn, D., 2011. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. Genome Res. 21, 915–924.

Altschul, S., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. https://doi.org/10.1093/nar/25.17.3389.

Alvar, J., Vélez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., den Boer, M., WHO Leishmaniasis Control Team, 2012. Leishmaniasis worldwide and global estimates of its incidence. PLoS One 7, e35671.

Amodeo, S., Kalichava, A., Fradera-Sola, A., Bertiaux-Lequoy, E., Guichard, P., Butter, F., Ochsenreiter, T., 2021. Characterization of the novel mitochondrial genome segregation factor TAP110 in. J. Cell Sci. 134 https://doi.org/10.1242/jcs.254300.

Andrews, S., 2010. FastQC: a Quality Control Tool for High Throughput Sequence Data.

Ansari, M.Y., Equbal, A., Dikhit, M.R., Mansuri, R., Rana, S., Ali, V., Sahoo, G.C., Das, P., 2016. Establishment of correlation between in-silico and in-vitro test analysis against Leishmania HGPRT to inhibitors. Int. J. Biol. Macromol. 83, 78–96.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B.P., Carrington, M., Depledge, D.P., Fischer, S., Gajria, B., Gao, X., Gardner, M.J., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Logan, F.J., Miller, J.A., Mitra, S., Myler, P.J., Nayak, V., Pennington, C., Phan, I., Pinney, D.F., Ramasamy, G., Rogers, M.B., Roos, D.S., Ross, C., Sivam, D., Smith, D.F., Srinivasamoorthy, G., Stoeckert Jr., C.J., Subramanian, S., Thibodeau, R., Tivey, A., Treatman, C., Velarde, G., Wang, H., 2010. TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res. 38, D457–D462.

Azevedo, L.G., de Queiroz, A.T.L., Barral, A., Santos, L.A., Ramos, P.I.P., 2020. Proteins involved in the biosynthesis of lipophosphoglycan in Leishmania: a comparative genomic and evolutionary analysis. Parasites Vectors 13, 44.

Babady, N.E., Pang, Y.-P., Elpeleg, O., Isaya, G., 2007. Cryptic proteolytic activity of dihydrolipoamide dehydrogenase. Proc. Natl. Acad. Sci. U.S.A. 104, 6158–6163.

Bacchi, C.J., Sufrin, J.R., Nathan, H.C., Spiess, A.J., Hannan, T., Garofalo, J., Alecia, K., Katz, L., Yarlett, N., 1991. 5′-Alkyl-substituted analogs of 5′-methylthioadenosine as trypanocides. Antimicrob. Agents Chemother. 35, 1315–1320.

Bansal, A.K., Meyer, T.E., 2002. Evolutionary analysis by whole-genome comparisons. J. Bacteriol. https://doi.org/10.1128/jb.184.8.2260-2272.2002.

Barghout, S.H., Patel, P.S., Wang, X., Xu, G.W., Kavanagh, S., Halgas, O., Zarabi, S.F., Gronda, M., Hurren, R., Jeyaraju, D.V., MacLean, N., Brennan, S., Hyer, M.L., Berger, A., Traore, T., Milhollen, M., Smith, A.C., Minden, M.D., Pai, E.F., Hakem, R., Schimmer, A.D., 2019. Preclinical evaluation of the selective small-molecule UBA1 inhibitor, TAK-243, in acute myeloid leukemia. Leukemia 33, 37–51.

Barylyuk, K., Koreny, L., Ke, H., Butterworth, S., Crook, O.M., Lassadi, I., Gupta, V., Tromer, E., Mourier, T., Stevens, T.J., Breckels, L.M., Pain, A., Lilley, K.S., Waller, R. F., 2020. A comprehensive subcellular atlas of the toxoplasma proteome via hyperLOPIT provides spatial context for protein functions. Cell Host Microbe 28, 752–766.e9.

Bekhit, A.A., El-Agroudy, E., Helmy, A., Ibrahim, T.M., Shavandi, A., Bekhit, A.E.-D.A., 2018. Leishmania treatment and prevention: natural and synthesized drugs. Eur. J. Med. Chem. 160, 229–244.

Berneburg, I., Stumpf, M., Velten, A.-S., Rahlfs, S., Przyborski, J., Becker, K., Fritz-Wolf, K., 2023. Structure of leishmania donovani 6-phosphogluconate dehydrogenase and inhibition by phosphine gold(I) complexes: a potential approach to leishmaniasis treatment. Int. J. Mol. Sci. 24, 8615.

Bijlmakers, M.-J., 2020. Ubiquitination and the proteasome as drug targets in trypanosomatid diseases. Front. Chem. 8, 630888.

Billington, K., Halliday, C., Madden, R., Dyer, P., Barker, A.R., Moreira-Leite, F.F., Carrington, M., Vaughan, S., Hertz-Fowler, C., Dean, S., Sunter, J.D., Wheeler, R.J., Gull, K., 2023. Genome-wide subcellular protein map for the flagellate parasite Trypanosoma brucei. Nat Microbiol 8, 533–547.

Blum, T., Briesemeister, S., Kohlbacher, O., 2009. MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. BMC Bioinf. 10, 274.

Boer, D.R., Bijlmakers, M.-J., 2019. Differential inhibition of human and trypanosome ubiquitin E1S by TAK-243 offers possibilities for parasite selective inhibitors. Sci. Rep. 9, 16195.

Boitz, J.M., Strasser, R., Yates, P.A., Jardim, A., Ullman, B., 2013. Adenylosuccinate synthetase and adenylosuccinate lyase deficiencies trigger growth and infectivity deficits in Leishmania donovani. J. Biol. Chem. 288, 8977–8990.

Boitz, J.M., Ullman, B., Jardim, A., Carter, N.S., 2012. Purine salvage in Leishmania: complex or simple by design? Trends Parasitol. 28, 345–352.

Bora, N., Jha, A.N., 2020. Metabolic pathway analysis identifying target against leishmaniasis - a kinetic modeling approach. Front. Genet. 11, 179.

Borgo, C., D'Amore, C., Capurro, V., Tomati, V., Sondo, E., Cresta, F., Castellani, C., Pedemonte, N., Salvi, M., 2022. Targeting the E1 ubiquitin-activating enzyme (UBA1) improves elexacaftor/tezacaftor/ivacaftor efficacy towards F508del and rare misfolded CFTR mutants. Cell. Mol. Life Sci. 79, 192.

Braga, S.S., 2019. Multi-target drugs active against leishmaniasis: a paradigm of drug repurposing. Eur. J. Med. Chem. 183, 111660.

Burge, R.J., Damianou, A., Wilkinson, A.J., Rodenko, B., Mottram, J.C., 2020. Leishmania differentiation requires ubiquitin conjugation mediated by a UBC2-UEV1 E2 complex. PLoS Pathog. 16, e1008784.

Burley, S.K., Berman, H.M., Kleywegt, G.J., Markley, J.L., Nakamura, H., Velankar, S., 2017. Protein Data Bank (PDB): the single global macromolecular structure archive. Methods Mol. Biol. 1607, 627–641.

Bushnell, B., Rood, J., Singer, E., 2017. BBMerge - accurate paired shotgun read merging via overlap. PLoS One 12, e0185056.

Carter, N.S., Kawasaki, Y., Nahata, S.S., Elikaee, S., Rajab, S., Salam, L., Alabdulal, M.Y., Broessel, K.K., Foroghi, F., Abbas, A., Poormohamadian, R., Roberts, S.C., 2022. Polyamine metabolism in parasites: a promising therapeutic target. Med. Sci. 10 https://doi.org/10.3390/medsci10020024.

Carter, N.S., Yates, P., Arendt, C.S., Boitz, J.M., Ullman, B., 2008. Purine and pyrimidine metabolism in Leishmania. Adv. Exp. Med. Biol. 625, 141–154.

Chávez-Fumagalli, M.A., Lage, D.P., Tavares, G.S.V., Mendonça, D.V.C., Dias, D.S., Ribeiro, P.A.F., Ludolf, F., Costa, L.E., Coelho, V.T.S., Coelho, E.A.F., 2019. In silico Leishmania proteome mining applied to identify drug target potential to be used to treat against visceral and tegumentary leishmaniasis. J. Mol. Graph. Model. 87, 89–97.

Cogliati, S., Cabrera-Alarcón, J.L., Enriquez, J.A., 2021. Regulation and functional role of the electron transport chain supercomplexes. Biochem. Soc. Trans. 49, 2655–2668.

Consortium, G.O., Gene Ontology Consortium, 2004. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res. https://doi.org/10.1093/nar/gkh036.

Cruz, A.K., de Toledo, J.S., Falade, M., Terrão, M.C., Kamchonwongpaisan, S., Kyle, D.E., Uthaipibull, C., 2009. Current treatment and drug discovery against Leishmania spp. and Plasmodium spp.: a review. Curr. Drug Targets 10, 178–192.

Cuypers, B., Domagalska, M.A., Meysman, P., Muylder, G. de, Vanaerschot, M., Imamura, H., Dumetz, F., Verdonckt, T.W., Myler, P.J., Ramasamy, G., Laukens, K., Dujardin, J.-C., 2017. Multiplexed spliced-leader sequencing: a high-throughput, selective method for RNA-seq in trypanosomatids. Sci. Rep. 7, 3725.

Dávila López, M., Martínez Guerra, J.J., Samuelsson, T., 2010. Analysis of gene order conservation in eukaryotes identifies transcriptionally and functionally linked genes. PLoS One 5, e10654.

Dhumal, T.T., Kumar, R., Paul, A., Roy, P.K., Garg, P., Singh, S., 2022. Molecular explorations of the Leishmania donovani 6-phosphogluconolactonase enzyme, a key player in the pentose phosphate pathway. Biochimie 202, 212–225.

Duelen, R., Corvelyn, M., Tortorella, I., Leonardi, L., Chai, Y.C., Sampaolesi, M., 2019. Medicinal Biotechnology for Disease Modeling, Clinical Therapy, and Drug Discovery and Development. Introduction to Biotech Entrepreneurship: From Idea to Business. https://doi.org/10.1007/978-3-030-22141-6_5.

Eriksen, T.A., Kadziola, A., Bentsen, A.K., Harlow, K.W., Larsen, S., 2000. Structural basis for the function of Bacillus subtilis phosphoribosyl-pyrophosphate synthetase. Nat. Struct. Biol. 7, 303–308.

Fagerberg, L., Stadler, C., Skogs, M., Hjelmare, M., Jonasson, K., Wiking, M., Abergh, A., Uhlén, M., Lundberg, E., 2011. Mapping the subcellular protein distribution in three human cell lines. J. Proteome Res. 10, 3766–3777.

Fang, P., Yu, X., Jeong, S.J., Mirando, A., Chen, K., Chen, X., Kim, S., Francklyn, C.S., Guo, M., 2015. Structural basis for full-spectrum inhibition of translational functions on a tRNA synthetase. Nat. Commun. 6, 6402.

Fernandes, M.C., Dillon, L.A.L., Belew, A.T., Bravo, H.C., Mosser, D.M., El-Sayed, N.M., 2016. Dual transcriptome profiling of leishmania-infected human macrophages reveals distinct reprogramming signatures. mBio 7. https://doi.org/10.1128/mBio.00027-16.

Fernández-Prada, C., Douanne, N., Minguez-Menendez, A., Pena, J., Tunes, L.G., Pires, D.E.V., Monte-Neto, R.L., 2019. Repurposed molecules: a new hope in tackling neglected infectious diseases. In: Silico Drug Design. https://doi.org/10.1016/b978-0-12-816125-8.00005-5.

Fidalgo, L.M., Gille, L., 2011. Mitochondria and trypanosomatids: targets and drugs. Pharm. Res. (N. Y.) 28, 2758–2770.

Field, M.C., Horn, D., Fairlamb, A.H., Ferguson, M.A.J., Gray, D.W., Read, K.D., De Rycker, M., Torrie, L.S., Wyatt, P.G., Wyllie, S., Gilbert, I.H., 2017. Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. Nat. Rev. Microbiol. https://doi.org/10.1038/nrmicro.2016.193.

Fisher, N., Meunier, B., Biagini, G.A., 2020. The cytochrome bc complex as an antipathogenic target. FEBS Lett. 594, 2935–2952.

Galina, L., Dalberto, P.F., Borges Martinelli, L.K., Roth, C.D., Michel Pinto, A.F., Villela, A.D., Bizarro, C.V., Machado, P., Saraiva Macedo Timmers, L.F., Norberto de Souza, O., Marcelino de Carvalho Filho, E., Basso, L.A., Santos, D.S., 2017. Biochemical, thermodynamic and structural studies of recombinant homotetrameric adenylosuccinate lyase from *Leishmania braziliensis*. RSC Adv. 7, 54347–54360.

Gazanion, E., Garcia, D., Silvestre, R., Gérard, C., Guichou, J.F., Labesse, G., Seveno, M., Cordeiro-Da-Silva, A., Ouaissi, A., Sereno, D., Vergnes, B., 2011. The Leishmania nicotinamidase is essential for NAD+ production and parasite proliferation. Mol. Microbiol. 82, 21–38.

Giraud, E., Martin, O., Yakob, L., Rogers, M., 2019. Quantifying metacyclic promastigotes from individual sandfly bites reveals the efficiency of vector transmission. Commun. Biol. 2, 84.

Gossmann, T.I., Ziegler, M., Puntervoll, P., de Figueiredo, L.F., Schuster, S., Heiland, I., 2012. NAD(+) biosynthesis and salvage–a phylogenetic perspective. FEBS J. 279, 3355–3363.

Gupta, I., Aggarwal, S., Singh, K., Yadav, A., Khan, S., 2018. Ubiquitin Proteasome pathway proteins as potential drug targets in parasite Trypanosoma cruzi. Sci. Rep. 8, 8399.

Halliday, C., Billington, K., Wang, Z., Madden, R., Dean, S., Sunter, J.D., Wheeler, R.J., 2019. Cellular landmarks of Trypanosoma brucei and Leishmania mexicana. Mol. Biochem. Parasitol. 230, 24–36.

Hanau, S., Rinaldi, E., Dallocchio, F., Gilbert, I.H., Dardonville, C., Adams, M.J., Gover, S., Barrett, M.P., 2004. 6-phosphogluconate dehydrogenase: a target for drugs in African trypanosomes. Curr. Med. Chem. 11, 2639–2650.

Heilker, R., Lessel, U., Bischoff, D., 2019. The power of combining phenotypic and target-focused drug discovery. Drug Discov. Today 24, 526–532.

Hofer, A., 2023. Targeting the nucleotide metabolism of Trypanosoma brucei and other trypanosomatids. FEMS Microbiol. Rev. 47 https://doi.org/10.1093/femsre/fuad020.

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., Nakai, K., 2007. WoLF PSORT: protein localization predictor. Nucleic Acids Res. 35, W585–W587.

Jakkula, P., Narsimulu, B., Qureshi, I.A., 2021. Biochemical and structural insights into 6-phosphogluconate dehydrogenase from Leishmania donovani. Appl. Microbiol. Biotechnol. 105, 5471–5489.

Jamal, S.B., Hassan, S.S., Tiwari, S., Viana, M.V., Benevides, L. de J., Ullah, A., Turjanski, A.G., Barh, D., Ghosh, P., Costa, D.A., Silva, A., Röttger, R., Baumbach, J., Azevedo, V.A.C., 2017. An integrative in-silico approach for therapeutic target identification in the human pathogen Corynebacterium diphtheriae. PLoS One 12, e0186401.

J, B., Brindha, J., Balamurali, M.M., Chanda, K., 2021. An overview on the therapeutics of neglected infectious diseases—leishmaniasis and chagas diseases. Front. Chem. https://doi.org/10.3389/fchem.2021.622286.

John Mayer, R., Ciechanover, A.J., Rechsteiner, M., 2008. The Ubiquitin-Proteasome System and Disease. John Wiley & Sons.

Karin, M., Ben-Neriah, Y., 2000. Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. Annu. Rev. Immunol. 18, 621–663.

Karp, P.D., Paley, S., Romero, P., 2002. The pathway tools software. Bioinformatics. https://doi.org/10.1093/bioinformatics/18.suppl_1.s225.

Katsyv, I., Wang, M., Song, W.M., Zhou, X., Zhao, Y., Park, S., Zhu, J., Zhang, B., Irie, H. Y., 2016. EPRS is a critical regulator of cell proliferation and estrogen signaling in ER + breast cancer. Oncotarget 7, 69592–69605.

Keller, T.H., Pichota, A., Yin, Z., 2006. A practical view of "druggability.". Curr. Opin. Chem. Biol. https://doi.org/10.1016/j.cbpa.2006.06.014.

Kelly, P., Hadi-Nezhad, F., Liu, D.Y., Lawrence, T.J., Linington, R.G., Ibba, M., Ardell, D. H., 2020. Targeting tRNA-synthetase interactions towards novel therapeutic discovery against eukaryotic pathogens. PLoS Neglected Trop. Dis. 14, e0007983.

Kerkhoven, E.J., Achcar, F., Alibu, V.P., Burchmore, R.J., Gilbert, I.H., Trybiło, M., Driessen, N.N., Gilbert, D., Breitling, R., Bakker, B.M., Barrett, M.P., 2013. Handling uncertainty in dynamic models: the pentose phosphate pathway in Trypanosoma brucei. PLoS Comput. Biol. 9, e1003371.

Khare, S., Nagle, A.S., Biggart, A., Lai, Y.H., Liang, F., Davis, L.C., Barnes, S.W., Mathison, C.J.N., Myburgh, E., Gao, M.-Y., Gillespie, J.R., Liu, X., Tan, J.L., Stinson, M., Rivera, I.C., Ballard, J., Yeh, V., Groessl, T., Federe, G., Koh, H.X.Y., Venable, J.D., Bursulaya, B., Shapiro, M., Mishra, P.K., Spraggon, G., Brock, A., Mottram, J.C., Buckner, F.S., Rao, S.P.S., Wen, B.G., Walker, J.R., Tuntland, T., Molteni, V., Glynne, R.J., Supek, F., 2016. Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness. Nature 537, 229–233.

Kim, D., Paggi, J.M., Park, C., Bennett, C., Salzberg, S.L., 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37, 907–915.

Kushwaha, V., Capalash, N., 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.

Kwon, N.H., Fox, P.L., Kim, S., 2019. Aminoacyl-tRNA synthetases as therapeutic targets. Nat. Rev. Drug Discov. 18, 629–650.

Le Guilloux, V., Schmidtke, P., Tuffery, P., 2009. Fpocket: an open source platform for ligand pocket detection. BMC Bioinf. 10, 168.

Liao, Y., Smyth, G.K., Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930.

Lima, M.L., Abengózar, M.A., Torres-Santos, E.C., Borborema, S.E.T., Godzien, J., López-Gonzálvez, Á., Barbas, C., Rivas, L., Tempone, A.G., 2022. Energy metabolism as a target for cyclobenzaprine: a drug candidate against Visceral Leishmaniasis. Bioorg. Chem. 127, 106009.

Li, S., Lu, Y., Peng, B., Ding, J., 2007. Crystal structure of human phosphoribosylpyrophosphate synthetase 1 reveals a novel allosteric site. Biochem. J. 401, 39–47.

Luque-Ortega, J.R., Rivas, L., 2007. Miltefosine (hexadecylphosphocholine) inhibits cytochrome c oxidase in Leishmania donovani promastigotes. Antimicrob. Agents Chemother. 51, 1327–1332.

Mailloux, R.J., Bériault, R., Lemire, J., Singh, R., Chénier, D.R., Hamel, R.D., Appanna, V. D., 2007. The tricarboxylic acid cycle, an ancient metabolic network with a novel twist. PLoS One 2, e690.

Mann, S., Frasca, K., Scherrer, S., Henao-Martínez, A.F., Newman, S., Ramanan, P., Suarez, J.A., 2021. A review of leishmaniasis: current knowledge and future directions. Curr Trop Med Rep 8, 121–132.

Mansilla, N., Racca, S., Gras, D.E., Gonzalez, D.H., Welchen, E., 2018. The complexity of mitochondrial complex IV: an update of cytochrome c oxidase biogenesis in plants. Int. J. Mol. Sci. 19 https://doi.org/10.3390/ijms19030662.

Martínez-Reyes, I., Chandel, N.S., 2020. Mitochondrial TCA cycle metabolites control physiology and disease. Nat. Commun. 11, 102.

Mirande, M., 2017. The aminoacyl-tRNA synthetase complex. Subcell. Biochem. 83, 505–522.

Mishra, A.K., Agnihotri, P., Srivastava, V.K., Pratap, J.V., 2015. Novel protein-protein interaction between spermidine synthase and S-adenosylmethionine decarboxylase from Leishmania donovani. Biochem. Biophys. Res. Commun. 456, 637–642.

Ortiz-Joya, L.J., Contreras Rodríguez, L.E., Ochoa, R., Ramírez Hernández, M.H., 2022. And study of an exclusive insertion in the nicotinamide/nicotinate mononucleotide adenylyltransferase from. Heliyon 8, e12203.

Pang, Y.L.J., Poruri, K., Martinis, S.A., 2014. tRNA synthetase: tRNA aminoacylation and beyond. Wiley Interdiscip. Rev. RNA 5, 461–480.

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

Pécoul, B., 2004. New drugs for neglected diseases: from pipeline to patients. PLoS Med. 1, e6.

Pham, J.S., Dawson, K.L., Jackson, K.E., Lim, E.E., Pasaje, C.F.A., Turner, K.E.C., Ralph, S.A., 2014. Aminoacyl-tRNA synthetases as drug targets in eukaryotic parasites. Int. J. Parasitol. Drugs Drug Resist 4, 1–13.

Phillips, C., Dohnalek, J., Gover, S., Barrett, M.P., Adams, M.J., 1998. A 2.8 Å resolution structure of 6-phosphogluconate dehydrogenase from the protozoan parasite Trypanosoma brucei : comparison with the sheep enzyme accounts for differences in activity with coenzyme and substrate analogues 1 1Edited by R. Huber. J. Mol. Biol. 282, 667–681.

Polanco, G., Scott, N.E., Lye, L.F., Beverley, S.M., 2022. Expanded proteomic survey of the human parasite leishmania major focusing on changes in null mutants of the golgi GDP-mannose/fucose/arabinopyranose transporter and of the mitochondrial fucosyltransferase. Microbiol. Spectr. 10, e0305222.

Popovic, D., Vucic, D., Dikic, I., 2014. Ubiquitination in disease pathogenesis and treatment. Nat. Med. 20, 1242–1253.

Radusky, L., Defelipe, L.A., Lanzarotti, E., Luque, J., Barril, X., Marti, M.A., Turjanski, A. G., 2014. TuberQ: a Mycobacterium tuberculosis protein druggability database. Database 2014, bau035.

Ramos, P.I.P., Fernández Do Porto, D., Lanzarotti, E., Sosa, E.J., Burguener, G., Pardo, A. M., Klein, C.C., Sagot, M.-F., de Vasconcelos, A.T.R., Gales, A.C., Marti, M., Turjanski, A.G., Nicolás, M.F., 2018. An integrative, multi-omics approach towards the prioritization of Klebsiella pneumoniae drug targets. Sci. Rep. 8, 10755.

Ranjan, P., Dubey, V.K., 2023. Krebs cycle enzymes for targeted therapeutics and immunotherapy for anti-leishmanial drug development using: pathways, potential targets, and future perspectives. Life Sci. 322, 121314.

Roatt, B.M., de Oliveira Cardoso, J.M., De Brito, R.C.F., Coura-Vital, W., de Oliveira Aguiar-Soares, R.D., Reis, A.B., 2020. Recent advances and new strategies on leishmaniasis treatment. Appl. Microbiol. Biotechnol. 104, 8965–8977.

Roberts, S.C., Scott, J., Gasteier, J.E., Jiang, Y., Brooks, B., Jardim, A., Carter, N.S., Heby, O., Ullman, B., 2002. S-adenosylmethionine decarboxylase from Leishmania donovani. Molecular, genetic, and biochemical characterization of null mutants and overproducers. J. Biol. Chem. 277, 5902–5909.

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140.

Rochette, A., Raymond, F., Corbeil, J., Ouellette, M., Papadopoulou, B., 2009. Whole-genome comparative RNA expression profiling of axenic and intracellular amastigote forms of Leishmania infantum. Mol. Biochem. Parasitol. 165, 32–47.

Roquero, I., Cantizani, J., Cotillo, I., Manzano, M.P., Kessler, A., Martín, J.J., McNamara, C.W., 2019. Novel chemical starting points for drug discovery in leishmaniasis and Chagas disease. Int. J. Parasitol. Drugs Drug Resist. 10, 58–68.

Rubio Gomez, M.A., Ibba, M., 2020. Aminoacyl-tRNA synthetases. RNA 26, 910–936.

Ryu, J.Y., Kim, H.U., Lee, S.Y., 2019. Deep learning enables high-quality and high-throughput prediction of enzyme commission numbers. Proc. Natl. Acad. Sci. U.S.A. 116, 13996–14001.

Sen, N., Majumder, H.K., 2008. Mitochondrion of protozoan parasite emerges as potent therapeutic target: exciting drugs are on the horizon. Curr. Pharmaceut. Des. 14, 839–846.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.

Singhal, A., Cheng, C.Y., 2019. Host NAD+ metabolism and infections: therapeutic implications. Int. Immunol. 31, 59–67.

Singh, S.P., Agnihotri, P., Pratap, J.V., 2013. Characterization of a novel putative S-adenosylmethionine decarboxylase-like protein from leishmania donovani. PLoS One 8, e65912.

Sosa, E.J., Burguener, G., Lanzarotti, E., Defelipe, L., Radusky, L., Pardo, A.M., Marti, M., Turjanski, A.G., Fernández Do Porto, D., 2018. Target-Pathogen: a structural bioinformatic approach to prioritize drug targets in pathogens. Nucleic Acids Res. 46, D413–D418.

Stincone, A., Prigione, A., Cramer, T., Wamelink, M.M.C., Campbell, K., Cheung, E., Olin-Sandoval, V., Grüning, N.-M., Krüger, A., Tauqeer Alam, M., Keller, M.A., Breitenbach, M., Brindle, K.M., Rabinowitz, J.D., Ralser, M., 2015. The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. Biol. Rev. Camb. Phil. Soc. 90, 927–963.

Sufrin, J.R., Spiess, A.J., Kramer, D.L., Libby, P.R., Miller, J.T., Bernacki, R.J., Lee, Y.H., Borchardt, R.T., Porter, C.W., 1991. Targeting 5'-deoxy-5'-(methylthio)adenosine phosphorylase by 5'-haloalkyl analogues of 5'-deoxy-5'-(methylthio)adenosine. J. Med. Chem. 34, 2600–2606.

Sunter, J.D., Dean, S., Wheeler, R.J., 2023. TrypTag.org: from images to discoveries using genome-wide protein localisation in Trypanosoma brucei. Trends Parasitol. 39, 328–331.

Sykes, S.E., Hajduk, S.L., 2012. Dual functions of α-ketoglutarate dehydrogenase E2 in the Krebs cycle and mitochondrial DNA inheritance in trypanosoma brucei. Eukaryot. Cell. https://doi.org/10.1128/ec.00269-12.

Tibayrenc, M., Ayala, F.J., 2013. How clonal are trypanosoma and leishmania? Trends Parasitol. https://doi.org/10.1016/j.pt.2013.03.007.

Timón-Gómez, A., Nývltová, E., Abriata, L.A., Vila, A.J., Hosler, J., Barrientos, A., 2018. Mitochondrial cytochrome c oxidase biogenesis: recent developments. Semin. Cell Dev. Biol. 76, 163–178.

Tsai, M., Koo, J., Yip, P., Colman, R.F., Segall, M.L., Howell, P.L., 2007. Substrate and product complexes of Escherichia coli adenylosuccinate lyase provide new insights into the enzymatic mechanism. J. Mol. Biol. 370, 541–554.

Van Assche, T., Deschacht, M., da Luz, R.A.I., Maes, L., Cos, P., 2011. Leishmania-macrophage interactions: insights into the redox biology. Free Radic. Biol. Med. 51, 337–351.

van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., Steinegger, M., 2024. Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. 42, 243–246.

Vercellino, I., Sazanov, L.A., 2021. The assembly, regulation and function of the mitochondrial respiratory chain. Nat. Rev. Mol. Cell Biol. 23, 141–161.

Vincent, I., Racine, G., Légaré, D., Ouellette, M., 2015. Mitochondrial proteomics of antimony and miltefosine resistant leishmania infantum. Proteomes. https://doi.org/10.3390/proteomes3040328.

Watson, S.A., McStay, G.P., 2020. Functions of cytochrome oxidase assembly factors. Int. J. Mol. Sci. 21 https://doi.org/10.3390/ijms21197254.

Webb, B., Sali, A., 2016. Comparative protein structure modeling using MODELLER. Curr. Protoc. Protein Sci. 86, 2.9.1–2.9.37.

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

Wilson, D.N., 2014. Ribosome-targeting antibiotics and mechanisms of bacterial resistance. Nat. Rev. Microbiol. 12, 35–48.

Xia, L., Björnstedt, M., Nordman, T., Eriksson, L.C., Olsson, J.M., 2001. Reduction of ubiquinone by lipoamide dehydrogenase. An antioxidant regenerating pathway. Eur. J. Biochem. 268, 1486–1490.

Xu, L., Dong, Z., Fang, L., Luo, Y., Wei, Z., Guo, H., Zhang, G., Gu, Y.Q., Coleman-Derr, D., Xia, Q., Wang, Y., 2019. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res. 47, W52–W58.

Yakobov, N., Debard, S., Fischer, F., Senger, B., Becker, H.D., 2018. Cytosolic aminoacyl-tRNA synthetases: unanticipated relocations for unexpected functions. Biochim. Biophys. Acta Gene Regul. Mech. 1861, 387–400.

Zhang, C., Shine, M., Pyle, A.M., Zhang, Y., 2022. US-align: universal structure alignments of proteins, nucleic acids, and macromolecular complexes. Nat. Methods 19, 1109–1115.

Zhang, X., Wu, R., Tian, C., Wang, W., Zhou, L., Guo, T., Yu, J., Wu, C., Shen, Y., Liu, X., Yu, R., 2022. GRP78 blockade overcomes intrinsic resistance to UBA1 inhibitor TAK-243 in glioblastoma. Cell Death Dis. 8, 133.

Zhao, R.-Z., Jiang, S., Zhang, L., Yu, Z.-B., 2019. Mitochondrial electron transport chain, ROS generation and uncoupling. submitted for publication Int. J. Mol. Med. 44, 3–15.

Zinngrebe, J., Montinaro, A., Peltzer, N., Walczak, H., 2014. Ubiquitin in the immune system. EMBO Rep. 15, 28–45.

Glossary

6PGD: 6-Phosphogluconate Dehydrogenase
KG: α-Ketoglutarate
ASL: adenylosuccinate lyase
AMP: Adenosine Monophosphate
CPM: Counts per Million
COX: Cytochrome *c* Oxidase
D: Druggable
DLD: Dihydrolipoamide Dehydrogenase
DS: Druggability Scores
EC: Enzyme Commission
E1: Ubiquitin Activating Enzyme
E2: Ubiquitin Conjugating Enzyme
E3: Ubiquitin Ligase
GO: Gene Ontology
HD: Highly Druggable

IMP: Inosine Monophosphate
KGDH: ɑ-Ketoglutarate Dehydrogenase Complex
NAD+: Nicotinamide Adenine Dinucleotide
NaPRT: Nicotinate Phosphoribosyltransferase
NTDs: Neglected Tropical Diseases
PDB: Protein Data Bank
Pfam: Protein Families Database
PD: Poorly Druggable
PRPP: Phosphoribosyl Pyrophosphate
rRNA: Ribosomal RNA

RFDT: Ranking Function of Druggable Targets
RNAi: RNA Interference
snoRNAs -: Small Nucleolar RNAs
SRA/NCBI -: Sequence Read Archive
TCA: Tricarboxylic Acid Cycle
TM-score: Template Modeling score
TP: Target-Pathogen
UBA1: Ubiquitin Activating Enzyme 1
Ub: Ubiquitin Molecules
VL: Visceral Leishmaniasis