Expanding the genetic toolbox of *Leishmania tarentolae* — novel techniques for the analysis of mitochondrial protein import in kinetoplastid parasites

Vom Fachbereich Chemie der Technischen Universität Kaiserslautern zur Verleihung des akademischen Grades „Doktor der Naturwissenschaften“ genehmigte Dissertation

DE - 386

TECHNISCHE UNIVERSITÄT KAISERSLAUTERN

vorgelegt von

M. Sc. Gino Lucas Turra

geboren in Rosario, Argentinien

Betreuer: Prof. Dr. Marcel Deponte

Kaiserslautern, 22.10.2020

Printed and/or published with the support of the German Academic Exchange Service

A mis abuelos

Affidavit

I hereby declare, that the experiments for the presented work are the result of my own work. I have not used any other sources or aids than those indicated in my thesis. The thesis in its current or similar version has not been submitted anywhere else.

Kaiserslautern, 01.09.2020

Gino Lucas Turra

Oral examination:	22.10.2020
Referees:	Prof. Dr. Marcel Deponte
	Prof. Dr. Nicole Frankenberg-Dinkel
	Prof. Dr. Jörg Fahrer

Table of contents

Acknowledgements

I would like to express my gratitude:

To Prof. Dr. Marcel Deponte, for opening the doors of his lab and giving me the opportunity to come to Germany. Thank you for your patience and the freedom I felt on the bench. Thank you to be open and encourage scientific discussions. You have been part of a journey where I grew as a person and as a professional.

To Prof. Dr. Nicole Frankenberg-Dinkel and Prof. Dr. Jörg Fahrer for accepting to examine this thesis.

To Frederik Sommer and Timo Mühlhaus, for the discussions and support in the mass spectrometry experiments.

To Carolina Andrade, Jessica Kehrer and Simone Eggert for the support and collaboration with the microscope. Special thanks to Jessica Kehrer for providing and helping with materials.

To Dr Eva Gluenz and Tom Beneke for the support with the CRISPR/Cas9 experiments.

To Bruce Morgan and his lab, for the help with the fluorescent measurements.

To Dr. Thomas Barends for the collaboration in the crystallization trials of Erv.

To Marcel Deponte, Robin Schumann, Sophie Möhring, Fabian Geissel and Carolina Andrade for proofreading this thesis. Thank you very much for your corrections and comments on this work.

To Jordi Vilurbina Pérez and Ogaga Daniel Ogbodu for the collaboration and support in this work.

To all my lab colleagues, Sandra Specht, Linda Liedgens, Verena Staudacher, Cletus Wezena, Kristina Feld, Johannes Krafczyk, Robin Schumann, Fabian Geissel, Luzia Schneider, Sophie Möhring and Eileen Bischoff. Thank you for always keeping a great sense of humour, for sharing the bad moments, but also the great ones. Thank you for the incredible experience I had in this country from which I take countless good memories. Thank you for the scientific discussions, the support and suggestions on many of the experiments in this work.

A todos mis amigos y amigas de Argentina: “Que afortunado soy de tener algo de lo que cuesta tanto despedirse”. Gracias por hacerme sentir, que el tiempo solo nos hizo más viejos.

A Facundo Merciadri, gracias por estar tan cerca, incluso por venir a visitarme en Kaiserslatuern.

A Romina Villagran, Lorindo Laici e la loro famiglia. Grazie per il vostro immenso supporto nei miei primi passi in Europa.

A mi familia, por apoyarme en vivir tan lejos, por acompañarme en mis sueños y creer siempre en mí.

A la hermosa alma que encontré durante este viaje, Carolina Andrade. Cruzar semejante océano no hubiera tenido sentido y nada de esto podría haber sido posible sin vos.

List of figures and tables

Figures

Tables

List of abbreviations and symbols

5-FC	5-fluorocytosine
APP	4-aminopyrazolo[3,4-d]pyrimidine
APRT	adenine phosphoribosyltransferase
APS	ammonium persulfate
At	*Arabidopsis thaliana*
BHI	brain heart infusion
BLA	ampicillin resistance cassette
BLAST	blasticidin resistance cassette
BLE	bleomycin resistance cassette
BSA	bovine serum albumin
Cas9	CRISPR-associated protein 9
CRISPR	clustered regularly interspaced short palindromic repeats
CRISPRi	CRISPR interference
C-terminal / C-termini	carboxy terminal
DAPI	4',6-diamidino-2-phenylindole
ddH_2O	double deionised water
DNA	deoxyribonucleic acid
DNAase	deoxyribonuclease
dNTPs	deoxyribonucleoside triphosphates
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediamine tetra acetic acid
FAD	flavin adenine dinucleotide
FBS	fetal bovine serum
GOI	gene of interest
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HYGRO	hygromycin resistance cassette
Lt	*Leishmania tarentolae*
MEM	minimal essential medium of Eagle
NEM	*N*-ethylmaleimide
NEO	G418 (Geneticin) resistance
N-terminal / N-termini	amino terminal
OD	optical density
ORF	open reading frame
p	parental strain (refers to TB007 transfectants)
PAM	protospacer adjacent motif
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Pf	*Plasmodium falciparum*
PFA	paraformaldehyde
PURO	puromycin resistance cassette
Rn	*Rattus norvegicus*
RNA	ribonucleic acid

RNAi	RNA interference
RT	room temperature
Sc	*Saccharomyces cerevisiae*
sgRNA	single guide RNA
SILAC	stable isotope labelling by amino acids in cell culture
spp.	*species pluralis*
T. brucei	*Trypanosoma brucei*
T. cruzi	*Trypanosoma cruzi*
T7 RNAP	T7 RNA Polymerase
Taq	*Thermus aquaticus polymerase*
TCA	trichloroacetic acid
TEMED	tetramethyl ethylenediamine
TK	thymidine kinase
UTR	untranslated region
wt	wild-type
yFCU	fusion protein yeast cytosine deaminase and uridyl phosphoribosyl transferase
β-ME	β-mercaptoethanol

CHAPTER

INTRODUCTION 1

1.1 The eukaryotic diversity of protists and its relevance to human health

Protists are single-celled eukaryotes that can be found in ecosystems over the whole planet [1]. They comprise a vast variety of species in different forms and functions and represent an important reservoir of eukaryotic phylogenetic diversity with species in each of the five supergroups (Figure 1.1). Protozoans are paraphyletic and were initially grouped based on their nutritional requirements [2, 3]. Protists are excellent models to study eukaryotic evolution. It is thought that some of them are extremely ancient, representing an early branch in the tree of life, not very distinct from the endosymbiotic ancestor that acquired the mitochondrion [4, 5]. Although the vast majority are free-living organisms, many adopted parasitic lifestyles. They are most known for being the aetiological agents of important and serious human diseases worldwide [1, 6-8]. Parasitic diseases caused by protists include: Leishmaniasis caused by *Leishmania spp*., Chagas disease, and sleeping sickness caused by *Trypanosoma spp.,* malaria caused by *Plasmodium spp*., toxoplasmosis caused by *Toxoplasma gondii* and diarrheal illness caused by *Entamoeba*, *Cryptosporidium*, and *Giardia.*

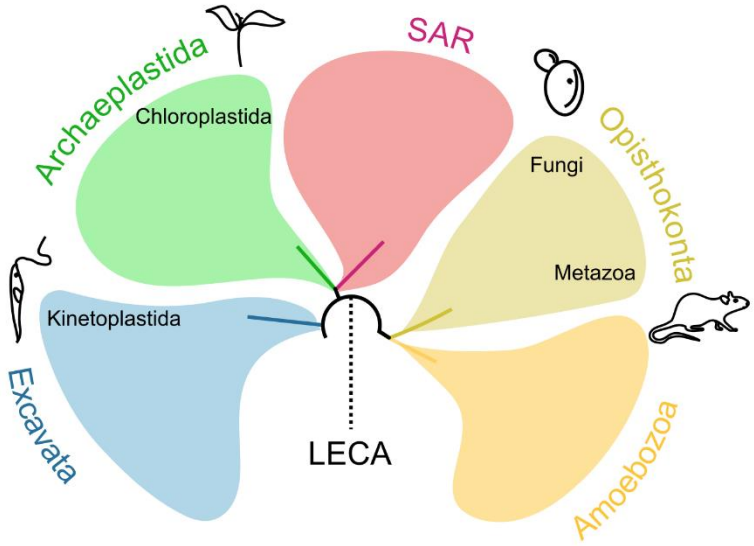


Figure 1.1 The phylogenetic tree of the Eukaryotes.
Simplified representation of the tree of life highlighting the five eukaryotic supergroups and its relationship with common model organisms. Kinetoplastida belong to the supergroup excavates. In contrast, canonical models such as yeast and animals belong to the supergroup opisthokonts. SAR, stramenopiles, alveolates, and rhizaria. LECA, last eukaryotic common ancestor. The figure was modified using information from [2, 9].

1.2 Kinetoplastid parasites: a new insight into the eukaryotic cell

Model organisms are the forefront of cell biology research. In eukaryotes, most of the basic molecular mechanisms in cell biology have been extensively studied in fungi (e.g. *Saccharomyces cerevisiae, Schizosaccharomyces pombe* and *Neurospora crassa),* and animals (e.g. *Caenorhabditis elegans, Drosophila melanogaster, Mus musculus* and Human cells). Except for plants, all these organisms represent the same supergroup of eukaryotes (a clade termed Opisthokonta) (Figure 1.1) [2]. Therefore, our understanding of the incredible diversity in the eukaryotic tree is rather limited. This is especially relevant in processes that might not be as similar or conserved as we initially thought.

An interesting group of protists distinct from opisthokonts are in Kinetoplastida (Figure 1.1). Kinetoplastid parasites would probably have received much less attention if some of them did not infect humans. Members of this group are also known to infect other vertebrates, arthropods, molluscs, other protozoans, and plants [4, 10]. Kinetoplastid parasites include the human-relevant genus of *Leishmania spp.* and *Trypanosoma spp*. Species belonging to them served as model organisms to understand their pathogenicity and fundamental aspects of their cell biology [11]. The defining identification key of all Kinetoplastids is the presence of thousands of concatenated circular DNAs within the mitochondrion known as kinetoplast [4,

12]. Another distinctive mark is their relative uniformity in the cell shape, all bearing a flagellum for locomotion that is essential in some developmental stages [13, 14]. Kinetoplastids can provide important lessons with "atypical" examples in basic biology. Exciting differences with canonical models include: the lack of transcriptional control with the absence of promoters [15, 16], unique post-transcriptional editing of RNAs [17], glycosomes with a compartmentalized glycolytic pathway [18], a flagellar pocket as a site of membrane trafficking [19], and a mitochondrial protein import machinery that differs from opisthokonts [20-22].

1.2.1 *Leishmania spp.*

More than 50 *Leishmania* species have been described to date, 31 species are known to infect mammals of which 20 are known to affect humans, causing Leishmaniasis [13, 23]. Leishmaniasis is a worldwide spectrum of human diseases present in all continents except Oceania, with serious medical and socio-economic consequences [24]. The parasite's life cycle requires a female sandfly and a mammalian host [25, 26]. In these two hosts, the parasite presents two distinct morphological developmental stages known as promastigote and amastigote. The life cycle begins when an infected sandfly injects the infective stage form promastigote into a human host during a blood meal. Here, the promastigote is phagocytized by a macrophage where it develops inside into a non-motile form called amastigote. Inside the macrophage, the amastigote form divides and multiplies and infects other macrophages. During this time, the host might become symptomatic and present with Leishmaniasis. Macrophage-infected amastigotes are ingested when a sandfly takes a blood meal from an infected host. In the insect's digestive tract, the parasite will develop into a flagellated extracellular form – the promastigote.

Is worth mentioning that the promastigote stage (in axenic cultures) is the form that is most used in genetic manipulations in *Leishmania* [27, 28], including this work.

1.2.2 *Leishmania tarentolae* as a model organism

L. tarentolae was first isolated in 1921 from an infected *Tarentola mauritanica* gecko. Today, it is probably the most studied member of the *Leishmania* subgenera that infect reptiles (*Sauroleishmania*) [29]. Although its genome has a high resemblance to the human-pathogenic species in the genus *Leishmania*, it is unable to infect humans [30]. For this reason, *L. tarentolae* is an attractive model organism. Other advantages of this species include their growth requirements and their fast growth kinetics, dividing every 5-6 h at 27 °C [29]. Cell cultures are maintained in normal incubators on agitation dispensing the need for expensive CO_2 supply and sera [29, 31]. To date, the genomes of more than 15 species of *Leishmania* [29] including *L. tarentolae* [32, 33] are available. This represents the introduction of *L. tarentolae* into the post-genomic era, becoming an established genetic model organism (Table 1.1). The use of *L. tarentolae* as a model provided great advances in Kinetoplastida research, such as the description of the RNA editing process and the kinetoplast mitochondrial DNA [17, 34], studies in DNA amplification [35], biochemistry, and classic genetics [36-38], gene transcription [39] and protein translation [40]. Furthermore, it was developed and engineered as a eukaryotic host for the heterologous production of recombinant proteins [41-43], with the advantage to produce certain glycoproteins [44]. *L. tarentolae* is also used by our group as a model to study mitochondrial protein import [20], work that will be described in more detail below.

Table 1.1 Genetic features of *L. tarentolae*

Genomic size	35.6 Mb
Chromosomes	Diploid*, 36 (2n*=72)
Numbers of protein-encoding-genes	8201
GC content	57.41%
Percentage shared with human-pathogenic species	90%
Genome sequenced in	2011, resequenced in 2020

**Leishmania* are well known for their aneuploidy, where the number of copies for a chromosome can vary. Although not examined in *L. tarentolae* it was shown that chromosomes can be found in the monosomic, disomic or trisomic state [45, 46]. Table based on the genomic information reported in Raymond *et al.* and Goto *et al.* [32, 33].

1.3 Mitochondrial protein import

The endosymbiotic theory explains the origin of the eukaryotic cell by describing the origin and evolution of the mitochondria and plastids [47]. The ancestral mitochondrion was once a free-living α-proteobacterium that was engulfed by an archaeal host and became an endosymbiont [48]. This defining event in evolution also provoked two remarkable phenomena: the migration of the genetic information from the α-proteobacterium to the nucleus of the host cell and the necessity of a machinery to import and sort the nuclear-encoded proteins needed for mitochondrial function [48, 49]. Thus, after translation in the cytosol, around 1500 proteins in the human model (representing 99% of the mitochondrial proteome) need to be imported and correctly located into the four compartments that compose the mitochondria [50]. These compartments comprise the outer (OMM) and inner mitochondrial membrane (IMM), and two enclosed spaces, the intermembrane space (IMS) between OMM and IMM and the matrix, enclosed by the IMM (Figure 1.2).

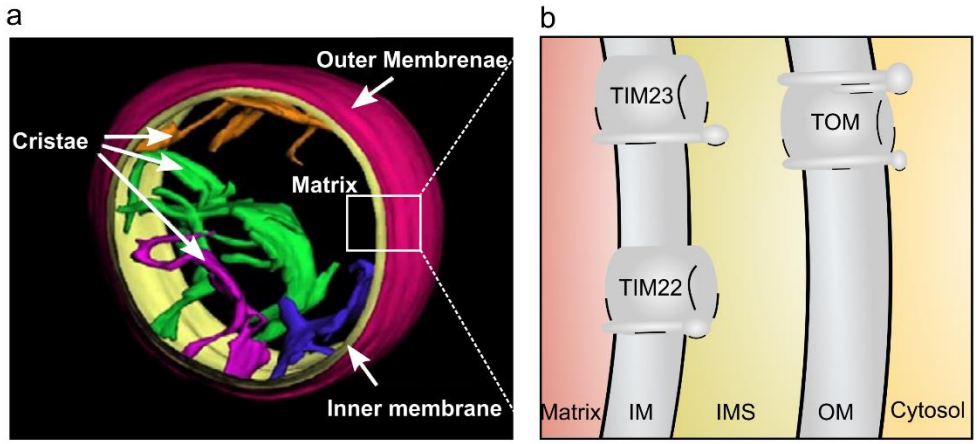


Figure 1.2 Mitochondrial compartments.
(a) Reconstructed tridimensional model by cryo-electron tomography of an intact mitochondrion from rat liver (Image taken and modified from [51]. Mitochondria are delimited by a double membrane system. Note the reduced spaced formed between the two membranes in comparison to the mitochondrial matrix. **(b)** The four different compartments formed by the double-membrane system. Major translocases complexes are indicated. IM, inner mitochondrial membrane, IMS, intermembrane space, OM, outer mitochondrial membrane.

In contrast, only 13 polypeptides in the human model remained encoded in the mitochondrial DNA. The functions of these proteins and their complexes are essential for the cell [52]. Mitochondria generate ATP by oxidative phosphorylation and are responsible for fatty acid oxidation and biosynthesis of cofactors and amino acids. They also contain the machinery for the biogenesis of Fe/S clusters, which are essential for proteins localized in the cytosol and

the mitochondria. The organelle is also involved in the cell cycle via the programmed cell death pathway [53].

The import of the polypeptides into the four different compartments is mediated by transmembrane protein complexes termed translocases. Almost all proteins targeted to the mitochondria enter through the translocase of the outer membrane (TOM) [54]. In many eukaryotes, there are two translocases located in the inner membrane (TIM23 and TIM22). The former is involved in the translocation of all nuclear encoded proteins located in the matrix, inner membrane, and some IMS proteins, whereas the latter is responsible for the integration of selected inner membrane proteins.

1.3.1 Protein import into the intermembrane space (IMS) *via* the MIA pathway

The space comprised between the mitochondrial membranes has a width similar to the membrane bilayer [55, 56]. In this compact compartment reside and function many components essential for mitochondria [57, 58]. A complete IMS proteome is still missing, and detailed knowledge of their functions remain unclear [59]. However, it is known that IMS proteins mediate the biogenesis of components of the electron-transport chain, the transport of proteins, metal ions, and metabolites [57, 58, 60, 61]. Furthermore, IMS proteins are also involved in apoptotic cell death by the release of cytochrome *c* to the cytosol, the biosynthesis of porphyrins and conversion of nucleotides [62-64].

In opisthokonts, there are three established import pathways to reach the IMS. One involves a mitochondrial targeting sequence (MTS) and the release of single transmembrane proteins by TIM23 into the inner membrane. The single transmembrane proteins are sometimes coupled to the release of IMS proteins by limited proteolysis. The second way involves the interaction of the target substrates with binding sites in the IMS [55]. However, the majority of the soluble proteins in the IMS are imported in a disulphide relay system via the mitochondrial import assembly (MIA) pathway [54, 59, 65, 66]. The clients of this pathway are typically low mass polypeptides with very simple structures [57, 58]. Their import does not require an MTS sequence but normally requires in opisthokonts the presence of conserved twin cysteine CX3C or CX9C-motifs and the catalysis of two essential enzymes: Mia40 and the flavoprotein Erv1 (Figure 1.3) [58, 60].

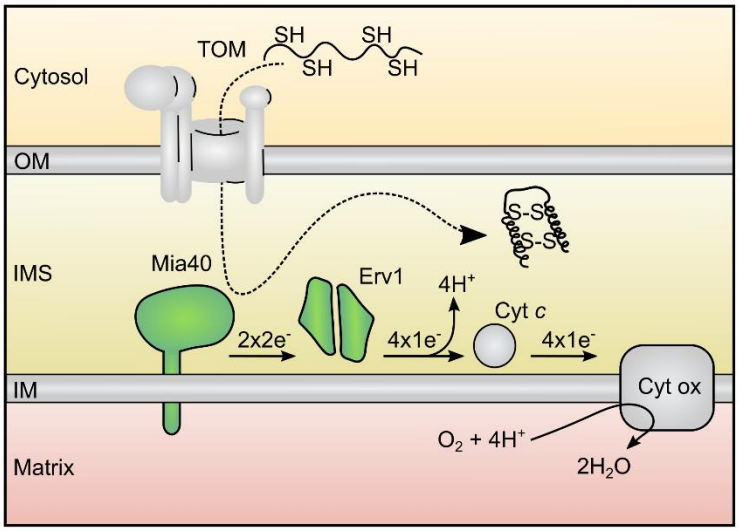


Figure 1.3 The Mia pathway in opisthokonts.
Model depicting the import and oxidation of small cysteine-rich substrates into the IMS. Proteins are synthesised in the cytosol in the reduced state and translocated through the TOM complex. Substrates are recognized and oxidized by Mia40. This oxidation drives folding and stabilize the proteins in its native state. To re-oxidize Mia40, the electrons flow in a disulphide relay system from Mia40 to the flavoprotein Erv1. The latter passes its electrons to cytochrome *c* and cytochrome *c* oxidase to oxygen, which is reduced to water. IM, inner mitochondrial membrane, IMS, intermembrane space, OM, outer mitochondrial membrane.

After translation in the cytosol, these small cysteine-rich proteins are translocated through the TOM complex in the reduced state and are recognized by Mia40 [67]. The recognition is mediated by a hydrophobic pocket, that serves as a receptor and couples the formation of a mixed disulphide bond with a CPC-motif. Both structural features are present in Mia40. The transfer of the disulphide bond that then occurs, follows the principle of a disulphide exchange reaction [68]. The oxidation plays a crucial mechanistic role. The disulphide bond formation triggers the folding of the substrates and after achieving their native conformation, the proteins become trapped in the otherwise diffusible nature of the IMS [54, 69]. In this context, Mia40 serves as a substrate receptor that introduces disulphide bonds. To oxidize another substrate, the CPC-motif of Mia40 needs to be re-oxidized. This reaction is carried out by the flavoprotein Erv1 which recycles Mia40 into its oxidized form and ultimately transfers the electrons to cytochrome *c* or oxygen. Therefore, the electrons sink into a disulphide relay from reduced substrates to the respiratory chain. This pathway has been extensively studied in opisthokonts, constituting probably the most supported experimental model. Although some features of this system are functionally conserved [20], the mechanism here described is not conserved throughout eukaryotes, as mentioned in the following section [21, 22, 70].

1.3.2 The non-conserved MIA pathway in Kinetoplastida

The mitochondrial protein import machineries were first considered to be conserved among eukaryotes [71]. However, this idea was heavily challenged in the last years. New reports of genome sequences from diverse eukaryotes distinct from opisthokonts made *in silico* comparisons of the machineries possible. The comparisons were later supported by experimental evidence showing remarkable differences between the import pathway constituents [20, 22, 70, 72-76]. Therefore, it is now clear that during the evolution of the eukaryotes different machineries were selected, offering alternative solutions to import and sort polypeptides into the mitochondria. In the case of the disulphide relay system in the IMS of Kinetoplastida, the only component found is the sulfhydryl oxidoreductase Erv (Figure 1.4) [20, 70, 72, 73].

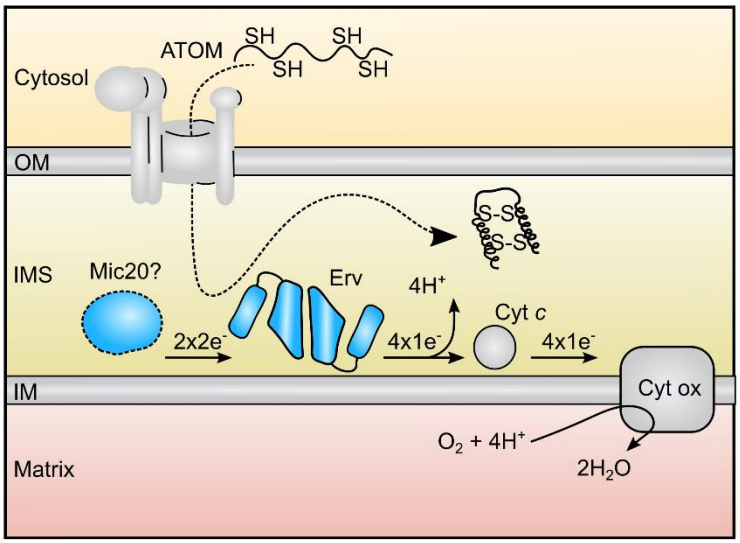


Figure 1.4 The incomplete Mia pathway in kinetoplastid parasites.
Model depicting the current knowledge of the import and oxidation of substrates into the IMS. A Mia40 homologue is missing and the only component found is the sulfhydryl oxidase Erv. Recently, a thioredoxin-like protein Mic20 was proposed to be the Mia40 homologue replacement. The existence of this so-far not characterized component suggests two distinct models: the sulfhydryl oxidase Erv can recognize and oxidize substrates alone or it is responsible for re-oxidizing Mic20 or another unknown protein. IM, inner mitochondrial membrane, IMS, intermembrane space, OM, outer mitochondrial membrane.

A similar situation exists in unicellular organisms belonging to alveolata (apicomplexan and ciliates) and heterokonts (supergroup SAR) [57, 70, 77]. How this system can operate without Mia40 in these organisms is a matter of intense research. The most studied model where no Mia40 was described belongs to Kinetoplastida. Indeed, the protein machineries present in these organisms are probably the best described after yeast and mammals [21]. Recently, a

thioredoxin-like protein termed Mic20 was proposed to replace Mia40 from opisthokonts [78]. Whether this or another protein is the true replacement of Mia40 and can shuttle electrons from substrates to Erv remains to be described [9, 21, 79].

1.4 The sulfhydryl oxidoreductase Erv

The physiological relevance of the sulfhydryl oxidoreductase Erv is known since the identification of its essentiality in yeast [80]. The phenotype of its temperature-sensitive mutant gave its actual name: essential for respiration and vegetative growth (Erv).

Erv is responsible for driving the disulphide bond formation in the IMS [81]. The structure consists of a four-helix bundle that binds a flavin cofactor non covalently to form a highly conserved core (Figure 1.5).

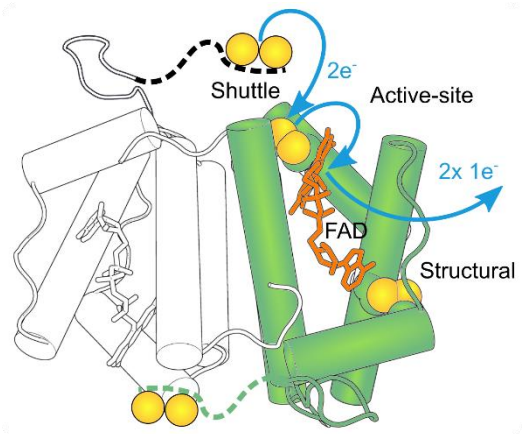


Figure 1.5 Structure diagram of the sulfhydryl oxidoreductase Erv.
The bound flavin cofactor is shown in stick representation (orange). Disulphide pairs are highlighted in sphere representation (yellow). Blue arrows indicate the direction of electron transfer. A five-helix fold binds the cofactor FAD. The shuttle pair in its oxidized form is responsible for oxidizing disulphide bonds in a disulphide exchange reaction. This redox-active pair is normally located in a flexible region of the protein that is accessible to the solvent. To engage another oxidation round, the redox pair must oxidize the active-site disulphide bond located proximal to the FAD cofactor (from the other subunit of the homodimer). The flavin then donates one electron to cytochrome *c* or oxygen. The *Rattus norvegicus* Erv structure [82] was used to generate the model (PDB: 1OQC). The figure was based on a review and study reported previously [72, 83].

The protein contains two redox-active cysteine pairs termed active-site and shuttle cysteine pairs. These redox-active cysteines combined mediate the electron transfer in an internal disulphide exchange reaction from the shuttle cysteine pair, to the active-site disulphide present in the core. This proximal active-site disulphide then reduces the cofactor FAD. The involvement of Erv in the disulphide relay could be considered a redox adaptor, switching from a two-electron transfer from thiol groups to the non-thiol-electron acceptor flavin [57,

84, 85]. This mechanism facilitates the re-oxidation of Erv, in two separate one-electron transfers to molecular oxygen (as sulfhydryl oxidase) or to cytochrome *c* (as a cytochrome *c* reductase).

1.4.1 The diversity of sulfhydryl oxidoreductases Erv

It is generally accepted that the reaction catalysed by Erv is conserved in eukaryotes. However, some structural arrangements of Erv are not conserved and the enzyme presents structural diversification (Figure 1.6) [57, 72, 83, 86]. Despite that the FAD domain is well conserved, Erv-homologues differ in the localization of the flexible arm that contains the shuttle pair. This redox-active disulphide can be located at the N- terminus (as in opisthokonts and apicomplexans), or at the C-termini (as in Kinetoplastida and plants). The length of the flexible arm can also differ. It can be a short polypeptide extension, as in the majority of the Erv-homologues [84, 87, 88], or it can be a long unusual domain only found in Kinetoplastida [72]. This long C-terminal tail is termed KISS (Kinetoplastida-specific second) domain of Erv [72]. Erv enzymes also have different shuttle pair motifs. In opisthokonts, this motif is characterized by CX2C shuttle pair in which the first X is normally an arginine or sometimes a lysine [87]. In plants and apicomplexan this motif exists in the CX4C form [72, 88]. In Kinetoplastida the shuttle pair has a CX3C-motif in which the last X is normally a tyrosine (Figure 1.6b). The amino acids that separate both thiols are termed intervening sequences. It was proposed that they likely have an impact on the redox potential and the rate of the disulphide exchange [83]. This could be explained by how the thiolate anion that is formed during the disulphide exchange reaction is stabilized [89-91]. Although this was investigated in active disulphide pairs from thioredoxins, the influence of the intervening sequences in the shuttle pair of Erv-homologues remains to be studied.

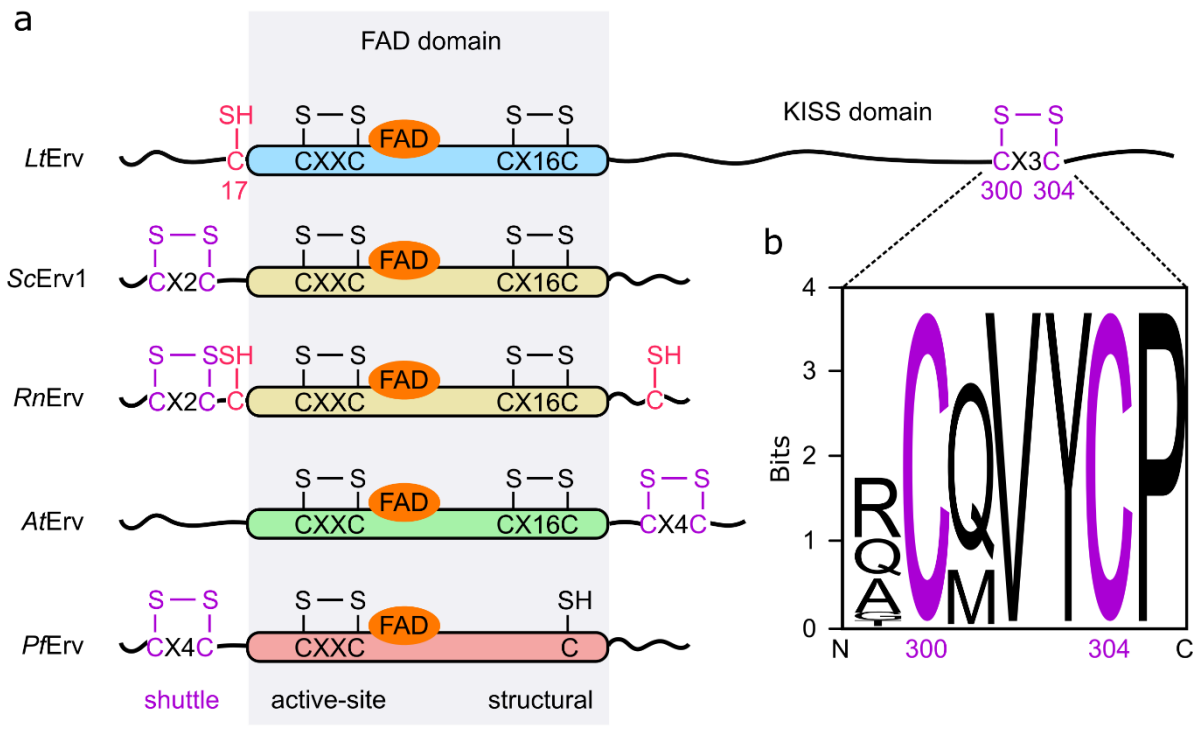


Figure 1.6 Structural diversity of Erv.
(a) Distinct domain and shuttle pair organization of five species representing four eukaryotic supergroups. *Lesihamina tarentolae* (*Lt*Erv), *Saccharomyces cerevisiae* (*Sc*Erv1), *Rattus norvegicus* (*Rn*Erv*), Arabidopsis thaliana* (*At*Erv) *Plasmodium falciparum* (*Pf*Erv*).* All proteins share a conserved FAD domain that contains a proximal or active site cysteine pair and a structural disulphide bond. A notable exception for the latter can be found in *Pf*Erv. Erv homologues also display the distal or shuttle cysteine pair (violet) at either the N or C termini. Cysteine residues located at the N and C-terminal tails of *Rn*ERv (magenta) serves in this structure to form a head-to-tail disulphide blond connecting both subunits in the homodimer. One of these cysteines is conserved among Kinetoplastida and is located at the N termini (Cysteine 17 in *Lt*Erv). Note that there is no clear correlation between the structural organization of Erv homologues and the type of MIA pathway. **(b)** Sequence logo analysis of 23 Erv homologues in Kinetoplastids. The cysteine pair is highly conserved as well as two positions comprising the intervening sequences 302 and 303. Panel a was based on previous reports [72, 86]. Panel **(b)** was generated with WebLogo [92]. Erv homologue sequences were retrieved from TriTrypDB.

1.5 Genetic manipulation of Kinetoplastida and the CRISPR/Cas9 editing revolution

The development of transfection techniques and selectable markers set the first steps to obtain mutants in kinetoplastid parasites [93]. The understanding of the unique post-transcriptional mechanism in Kinetoplastida was essential to obtain plasmid transfectants. This knowledge was useful to generate cassettes that expressed the gene of interest and a selectable marker. The generation of expression cassettes was also used to generate the first knockout mutants [94, 95]. These experiments usually consisted of the transfection of a targeting cassette with sequences homologues to the target site and a resistance cassette to kill the wild type cells. Because of the low efficiency of the transfection (or probability of

integration), the generation of these mutants would have been impractical without the availability of selectable markers [93].

Despite the availability of these sets of techniques since the 1990s, and the complete genome sequence of many kinetoplastid parasites over a decade ago, editions at the chromosomal level were a challenging endeavor [96]. To integrate a targeting cassette into the chromosome by homologous recombination, a double-strand break must occur serendipitously close to the target locus. The chromosomal stability of the cell depends on this error to happen at low frequency. This could explain the difficulty to select for this rare genetic event. This difficulty created a great interest in the development of DNA nucleases that can induce a double-strand break at the locus of interest (e.g. zinc-finger and TALEN nucleases)[97]. Thus, the report of an efficient and programmable nuclease by the CRISPR/Cas9 system created a revolution, not only in the parasitology field but across all life sciences [98-101].

The CRISPR (clustered regularly interspaced short palindromic repeat)–Cas system was initially discovered as an RNA-mediated adaptive defence mechanism in bacteria and archaea [102]. The ability of an RNA-guided endonuclease CRISPR-associated (Cas) protein to introduce double-strand breaks and silence foreign DNA was used to generate the CRISPR/Cas9 technology [103]. This technology relies on two components, the endonuclease Cas9 and a single guide RNA (sgRNA). The Cas9 can generate a double-strand break at the target locus by means of a specific sgRNA. The sgRNA is able to “guide” the endonuclease Cas9 to the target locus by Watson–Crick pairing of 20 nucleotides with the genomic DNA (Figure 1.7) [104]. Therefore, Cas9 and sgRNA form a ribonucleoprotein complex that can localize the complementary genomic sequence of the sgRNA. The formation of this complex and the generation of the double-strand break only happens if the 20 nucleotides complementary to the sequence are flanked by a protospacer-adjacent motif (PAM) [104]. The CRISPR/Cas9 technology, therefore, provides a programmable endonuclease that can target and/or cleave a sequence of DNA by simply supplying an artificially designed guide RNA [99]. To construct the sgRNA it is only necessary to add twenty nucleotides at the 5’ end of a universal Cas9-binding gRNA sequence [105].

After the Cas9-induced double-strand break 3 nucleotides upstream of the PAM site, a donor DNA or targeting cassette is provided containing homology regions to induce homology-

directed repair and a selectable marker to select the mutant. The specific design of this cassette will depend on the edition desired.

The first CRISPR/Cas9 report in Kinetoplastida came shortly after the implementation of the system in 2014 and it was done in *Trypanosome cruzi* [106]. The implementation of the system in *Leishmania* followed with report strategies in *L. donovani* [107] and *L. major* [108]. The latter reports consisted of plasmid-based strategies where Cas9 and sgRNA were expressed from vectors. Both elements, sgRNA and donor DNAs must be customized for each target gene of interest by cloning steps. These previous methods were highly improved by a cloning-free, PCR based CRISPR/Cas9 strategy reported by Beneke *et al.* [109]. In the course of this PhD project, this strategy was applied successfully in *Leishmania mexicana*, *L. major*, *L. donovani, Trypanosoma brucei* and *T. cruzi* [109-111]. The system relies on the simple and rapid generation of sgRNA and donor DNA cassettes by the extension of two DNA primers and PCR using modular antibiotic resistance cassettes respectively. Both elements are co-transfected in a line expressing Cas9 and a T7 RNA polymerase. The latter serves to transcribe the sgRNA-encoding DNA *in vivo* by means of T7 RNA promoter in the cassette (for details about the method see section 2.5.8).

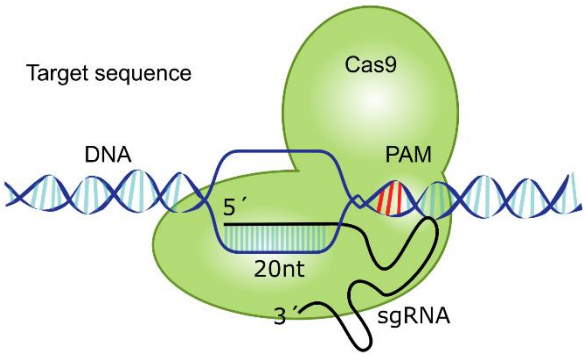


Figure 1.7 CRISPR/Cas9 elements.
Shematic of the ribonucleoprotein complex depicting the elements necessary to induce a double-strand break at the target sequence. A customized sgRNA must be complementary (20 nucleotides) to the target sequence and be immediately adjacent to a PAM NGG motif.

1.6 Techniques for the study of essential genes

The study of genes that cannot be replaced or mutated by homologous recombination, and are therefore essential, carries a technical challenge [28]. In yeast, many of the fundamental aspects of cell biology that were discovered involved essential genes and the generation of conditional mutants [112]. Conditional mutants are invaluable for the study of essential genes. They allow the cell to survive under certain conditions and reveal the mutant phenotype after external induction. In yeast, the typical conditional mutants are temperature-sensitive strains [113]. However, the generation of this type of mutation has not been described in Kinetoplastida. In *Leishmania*, techniques to generate conditional mutants by gene expression regulation or protein abundance are scarce and often inefficient. For example, RNAi is restricted to some species of *Leishmania,* as detailed in the following section. Instead, other solutions have been tested, including the use of tetracycline-inducible promoters [114], drug-regulated protein degradation by tagging a destabilisation domain [115, 116], and plasmid shuffling using a negative marker [27, 117, 118].

1.6.1 The lack of an RNA interference (RNAi) system and *glmS* riboswitch

Leishmania lacks conventional promoters and genes are transcribed as polycistronic mRNA precursors that are later maturated by trans-splicing [15]. This feature imposes a constraint on the design of knockdown systems. For example, the use of a tetracycline-inducible promoter or CRISPRi (where a mutant Cas9 blocks transcription) at the chromosomal level would be impractical. The generation of such mutants could generate multiple knockdown effects in genes within the polycistron. In this context, an inducible knockdown system should act at the post-transcriptional mRNA level. One of these techniques is RNA interference (RNAi), but it is not feasible in all the genus *Leishmania* [119], apparently because of a non-functional RNAi machinery [120]. Despite this system being functional in *L. braziliensis* [121], this technique cannot be used in the rest of the genus including *L. tarentolae.* Engineering a functional system in RNAi-deficient *Leishmania* by introducing missing RNAi components as shown in yeast [122], has been discarded for technical reasons and has not been attempted to date. An alternative knockdown system at the post-transcriptional level is the use of RNA riboswitches. An example that was then developed as a molecular tool is the *glmS* riboswitch. This sequence was first described in *Bacilus subtilis*, in the UTR region of the *glmS* gene. This gene encodes for glucosamine-6-phosphate synthase (glmS)

an enzyme that catalyses the biosynthesis of fructose-6-phosphate and glutamine to glucosamine-6-phosphate (GlcN6P) [123, 124]. The riboswitch sequence was shown to control the levels of the *glmS* transcript by means of a catalytic RNA that folds into a structure that self-cleaves upon binding of glucosamine-6-phosphate (GlcN6P) [125]. The generation of chimeric mRNA, in which the *glmS* riboswitch-coding sequence is inserted into the 3' UTR of a gene, was shown to be sufficient to down-regulate the expression of the targeted gene by the external addition of glucosamine (GlcN) [126]. This tool was successfully applied in yeast [126], *Plasmodium falciparum* [127, 128] and, more recently, *Trypanosoma brucei and T. cruzi* [129, 130].

1.7 Aim of the study

This thesis aimed to test three novel genetic tools in the model organism *Leishmania tarentolae.* These techniques included:

- The CRISPR/Cas9 system to generate knockout and knockin mutants.
- The *glmS* riboswitch to generate knockdown mutants.
- A negative selection system to perform shuffling experiments.

Some of these techniques were further used to study mitochondrial protein import using *L. tarentolae*. We centred our attention on the sulfhydryl oxidoreductase Erv and its unique structural properties compared to opisthokonts. *Leishmania* Erv structure-function relationships were previously studied *in vitro* and in yeast [72, 131]. *In silico* analysis revealed the unusual presence of a long C-terminal tail forming the KISS domain [72]. Complementation studies in yeast reported by Specht *et al.* showed that *Leishmania* Erv cannot complement its yeast homologue unless cysteine 17 is mutated [131]. An Erv interactome in Kinetoplastida was also missing. This experiment could provide essential evidence to support or dismiss the presence of a Mia40 homologue in Kinetoplastida. Taking this into account, the specific objectives of this part of the thesis can be summarized in:

- Does the KISS domain play an essential role in kinetoplastid parasites?
- Is cysteine 17 essential for survival?
- Which protein/s are interacting with Erv *via* disulphide bonds?
- Is Erv an essential gene in *L. tarentolae*?.

CHAPTER

MATERIALS & METHODS 2

2.1 Materials

2.1.1 Equipment

Table 2.1 List of laboratory equipment

Equipment	Manufacturer
Analytical balance	Kern
Benchtop centrifuge Mikro 220R	Hettich
Benchtop incubator for L. tarentolae IPP 500	Memmert
Benchtop incubator shaker MaxQ4450	Thermo Scientific
Benchtop refrigerated centrifuge Rotina 380 R	Hettich
Bright-field microscope AXIO	Zeiss
CCD digital gel imaging system Fusion SL	Vilbert Lourmat
Cell counting chamber Neubauer improved	Marien Field
Floor refrigerated centrifuge j2-21, rotor JA-10, JA-20	Beckman
Horizontal agarose minigel system	Biostep
Incubator shaker innova 44R	New Brunswick
Inverted bright-field microscope AE30	Motic
Laminar flow hood for L. tarentolae Safe 2020	Thermo Scientific
Laminar flow hood, E. coli MSC-Advantage	Thermo Scientific
Magnetic stirer-hot plate Hei-Standard	Heidolph

Equipment	Manufacturer
Metal block thermostat MBT 250	Kleinfeld Labortechnik
Micropipette PIPETMAN (p2,P10, p20,p50,p200,p1000)	Gilson
Microplate spectrophotometer CLARIOstar	BMG
Mini-PROTEAN electrophoresis cell	Biorad
NanoDrop 1000 Spectrophotometer	Thermo Scientific
Patform shaker duomax 1030	Heidolph
Personal-sized incubator	VWR
pH-meter PB-11	Sartorius
Pipette Controller Accu-jet pro	Brand
Power Supply Mini 300V	Major Science
Semi-Dry PerfectBlue system	VWR
Thermal Cycler Mastercycler Nexus	Eppendorf
Thermomixer shaker incubator block	Eppendorf
Transfection Device Nucleofector II/2b	Lonza
Ultrasonic laboratory homogenizer Sonopuls HD 2070	Bandelin electronic
Vertical autoclave V-65	Systec
Vortex-mixer Genie 2	Heidolph
Water bath F12-ED	JULABO
Wet Tank Mini-Trans blot cell system	BioRad

2.1.2 Disposables

Table 2.2 List of disposables

Disposables	Source
6-well plates	Greiner
Cell culture flask T-25 with filter cap for L. tarentolae	Greiner
Centrifuge tubes (15 and 50 mL)	Greiner
Cryovials Cryos 2 mL	Greiner
Durapore sterile filters PVDF (0.22 µm)	Merck
Electroporation Cuvettes plus 2 mm gap	BTX
Falcon 96 well black/clear flat bottom plate	Corning

Disposables	Source
Inoculation loop 10 µL	Simport
Micro reaction tubes (1.5 and 2 mL)	Sarstedt
Parafilm	Bemis
PCR reaction tubes	Sarstedt
Petri dishes 85/13 mm	Waldeck
Poly-Prep chromatography columns	Biorad
PVDF membrane	Merck
Scalpel	Braun
Serological pipettes (1, 5, 10 and 25 mL)	Sarstedt
Spin chromatography columns	Biorad
Stericup sterile vacuum filter PVDF (0.22 µm)	Merck
Sterile syringes	BD Plastipak
SuperFrost microscope slides	Buddeberg
Whatman paper	GE Healthcare

2.1.3 Chemicals

Table 2.3 List of chemicals

Reagents	Source
Acetone	Honeywell
Acrylamide/Bis Solution, 37.5:1 (30 % w/v)	Serva
Adenine sulphate	Sigma Aldrich
Adenosine	Sigma Aldrich
Agar	Carl Roth
Agarose	Serva
Aluminium sulphate	Carl Roth
Ammonium persulfate (APS)	Sigma Aldrich
Ammonium sulphate	Merck
APP (4-Aminopyrazolo[3,4-d]pyrimidine)	Sigma Aldrich
β-mercaptoethanol	Sigma Aldrich
Boric Acid	Sigma Aldrich
Bradford reagent	Bio-Rad

Reagents	Source
Brain heart infusion (BHI)	BD
Bromophenol blue	Waldeck
BSA	Serva
Calcium chloride	Merck
Cobalt(II) chloride	Carl Rorth
cOmplete EDTA-free protease Inhibitor	Roche
Coomassie Brilliant Blue G 250	AppliChem
D-(+)-Galactose	Carl Roth
D-(+)-Glucosamine hydrochloride	Sigma Aldrich
D-(+)-Glucose	Merck
Deoxyribonucleoside triphosphate (dNTPs)	Thermo Scientific
Diamide	Sigma Aldrich
Dithiothreitol (DTT)	Sigma Aldrich
DMSO	Sigma Aldrich
DNA ladder (100 bp, 1kb)	New England Biolabs
DNase I	Roche
Ethanol (absolute)	VWR
Ethylenediaminetetraacetic acid (EDTA)	Sigma Aldrich
Fetal bovine serum (FBS heat inactivated)	Gibco
Fetal bovine serum dialysed (dFBS)	Gibco
Flavin adenine dinucleotide (FAD)	TCEP
Folic acid	Sigma Aldrich
Gel loading dye, purple 6×	New England Biolabs
Gel Red dye, 6X	Biotium
Glycine	Merck
Hemin chloride	Calbiochem
HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid]	Merck
Hydrochloric acid	VWR
Hypoxanthine	c.c.pro
Imidazole	Merck
Iron(III) chloride	Sigma Aldrich
Isopropanol	Merck
L-Arginine monohydrochloride	Sigma Aldrich

20 Materials and Methods

Reagents	Source
L-Arginine-13C6 hydrochloride	Sigma Aldrich
L-Glutamine	Sigma Aldrich
L-Lysine hydrochloride	Sigma Aldrich
L-Lysine-6-13C dihydrochloride	Sigma Aldrich
Luria Bertani (LB) agar	Carl Roth
Luria Bertani (LB) medium	Carl Roth
M199 medium without L-arginine and L-lysine	Caisson Laboratories
Magnesium sulphate	Merck
Manganese (II) chloride	Merck
Manganese(II) chloride	Merck
MEM medium	Gibco
Methanol	Merck
Milk powered, skimmed blotting grade	Carl Roth
MOPS (3-(N-morpholino)propane sulfonic acid)	Gerbu Biotechnik
N-ethylmaleimide (NEM)	Sigma Aldrich
Ni-NTA-Agarose	QIAGEN
Niquel (II) chloride	Sigma Aldrich
Phosphoric acid	Sigma Aldrich
PonceauS	Serva
Potassium chloride	AppliChem
Potassium dihydrogen phosphate	Merck
Potassium monohydrogen phosphate	Merck
Prestained protein ladder	Thermo Scientific
Sodium acetate	Sigma Aldrich
Sodium chloride	Sigma Aldrich
Sodium dihydrogen phosphate	Merck
Sodium dodecyl sulphate (SDS)	Serva
Sodium hydroxide	Carl Roth
Sodium molybdate	Strem chemicals
Strep-Tactin Sepharose	IBA
Sodium monohydrogen phosphate	Merck
TCEP (tris(2-carboxyethyl)phosphine	Sigma Aldrich
Tetramethyl ethylenediamine (TEMED)	Serva

Reagents	Source
Trichloroacetic acid	Merck
TRIS (tris(hydroxymethyl)aminomethane)	Carl Roth
Triton X-100	Merck
Tryptone	BD
Tween-20	Sigma Aldrich
Yeast extract	BD
Zinc sulphate	Merck
Sodium Selenite	Sigma Aldrich

2.1.4 Selectable markers

Table 2.4 List of antibiotics

Drug	Stock	Working concentration	Supplier
Ampicilin	100 mg/mL in 50% ethanol	100 µg/mL	AppliChem
Blasticidin S	10 mg/mL in HEPES buffer	10 µg/mL	Invivogen
Chorampenicol	34 mg/mL in 100% ethanol	34 µg/mL	AppliChem
G418 disulphate	50 mg/mL in ddH_2O	0.1 mg/mL	AppliChem
Hygromycin	50 mg/mL in ddH_2O	100 µg/mL	Carol Roth
Kanamycin	50 mg/mL in ddH_2O	50 µg/mL	AppliChem
Puromycin	10 mg/mL in in HEPES buffer	20 µg/mL	Invivogen

2.1.5 Enzymes

Table 2.5 List of enzymes

Enzyme	Manufacturer
Antarctic phosphatase	New England Biolabs
Pfu DNA polymerase	Promega
Phusion DNA polymerase	New England Biolabs
Restriction enzymes	New England Biolabs
T4 DNA ligase	New England Biolabs
Taq DNA polymerase	New England Biolabs

2.1.6 Kits

Table 2.6 List of kits

Kit	Manufacturer
Cytiva Western Blotting Detection Reagents	GE Healthcare
Expand High Fidelity PCR System	Roche
QIAamp DNA Blood Mini Kit	QIAGEN
QIAprep Spin Midiprep Kit	QIAGEN
QIAprep Spin Miniprep Kit	QIAGEN
ROTI Black P Silver staining	Carl Roth
Wizard SV Gel and PCR Clean-Up System	Promega

2.1.7 Software and bioinformatics

Table 2.7 Software and bioinformatics tools

Software/ Bioinformatics tools	Developer
Benchling Life Sciences R&D Cloud (2020)	https://www.benchling.com/
Cas-OFFinder	http://www.rgenome.net/cas-offinder/
EndNote X7	Thomson Reuters
ImageJ	NIH
Inkscape 1.0 (2020)	https://inkscape.org/
MARS	BMG Labtech
MitoFates	http://mitf.cbrc.jp/MitoFates/cgi-bin/top.cgi
MS Office	Microsoft
NCBI databases	https://www.ncbi.nlm.nih.gov/
Optimizer	http://genomes.urv.es/OPTIMIZER/
Perseus	Max Planck Institute
PyMOL Version 1.7.2.1.	Schrödinger, LLC
SigmaPlot 13.0	Systat Software, Inc.
TriTrypDB (2019) release 46	https://tritrypdb.org/tritrypdb/
WatCut Silent mutagenesis tool	http://watcut.uwaterloo.ca/
Weblogo	https://weblogo.berkeley.edu/logo.cgi

2.1.8 Antibodies

Table 2.8 List of antibodies

Antibody	Host species	Source	Dilution
Anti 6x-His Tag Monoclonal A	Mouse	Thermo Scientific	1:1000
Anti-FLAG Monoclonal	Mouse	Thermo Scientific	1:1000
α-LtErv	Rabbit	Eckers *et al.*. 2012	1:500
α-Mouse IgG (H+L)-HRP conjugate	Goat	Biorad	1:10000
α-Rabbit IgG (H+L)-HRP conjugate	Goat	Biorad	1:10000

2.1.9 Bacterial strains

Table 2.9 List of bacterial strains

Strain	Genotype	Application	Source
XL-1 blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10 (Tetr)].	Cloning	Qiagen
SHuffle T7 Express	fhuA2 lacZ::T7 gene1 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbC (SpecR, lacIq) ΔtrxB sulA11 R(mcr-73::miniTn10--TetS)2 [dcm] R(zgb-210::Tn10 --TetS) endA1 Δgor Δ(mcrC-mrr)114::IS10	Protein production	New England Biolabs

2.1.10 *Leishmania tarentolae* strains

Unless otherwise stated, all shuttle plasmids were transfected transiently in wild-type *Leishmania tarentolae* cells and are listed in Table 2.11. To generate stable transgenic parasites by the CRIPRS/Cas9 system, the pTB007 strain was used and is further considered in this work as the parental strain.

Table 2.10 List of *Leishmania tarentolae* strains

Wild-type strain	Features	Source
Parrot TarII/UC	Isolated by Parrot from an Algerian gecko in 1939	André Schneider*
Parental strain		
pTB007	Transfectant expressing Cas9 and T7 RNA polymerase	This work
Transgenic		
ΔPF16	Knockout strain	This work
PF16::mCherry	C-terminal tagging mCherry	This work
PF16::*glmS*^{wt}	C-terminal tagging glmSwt	This work
PF16::*glmS*^{M9}	C-terminal tagging glmSM9	This work
ΔAPRT	Knockout strain	This work
APRT::*glmS*^{wt}	APRT gene fused with glmSwt (C-terminal)	This work
APRT::*glmS*^{M9}	APRT gene fused with glmSM9 (C-terminal)	This work
ERV$^{\Delta KISS}$-His_8	Truncated mutant lacking (KISS) domain. His_8-tagged (C-terminal)	This work
His_8-*ERV*^{C17S}	Cysteine-to-serine mutation in position 17. His_8-tagged (N-terminal)	This work
ERV-His_8	ERV gene tagged with His8-tag (C-terminal)	This work
ERV-Streptag	ERV gene tagged with Strep-tag (C-terminal)	This work

*Department of Chemistry and Biochemistry, University of Bern, Bern, Switzerland

2.1.11 Vectors

Table 2.11 List of shuttle vectors used to transfect *L. tarentolae*

Plasmids	Features	Reference
pTB007	AmpR, Hygro. FLAG::NLS::Cas9::NLS, NLS::T7RNAP. Generation of *L. tarentolae* cell line expressing Cas9 nuclease and T7RNAP polymerase	Beneke *et al.**
pTPuro_v1	AmpR, PuroR. For generation of null mutants.	Beneke *et al.**
pTBlast_v1	AmpR, BlastR. For generation of null mutants.	Beneke *et al.**
pPLOTv1 P-mCherry-puro	AmpR, PuroR. For generation of N- or C-terminal mCherry fusions.	Beneke *et al.**
pPLOTv1 blast-mNeonGreen-blast	AmpR, BlastR. For generation of N- or C-terminal mNeonGreen fusions.	Beneke *et al.**
pMOTag-4-*glmS^{wt}-PURO*-5	AmpR, PuroR. For generation of endogenous 3'-terminal glmS^{wt} fusions.	This work
pMOTag-4-*glms^{M9}-PURO*-5	AmpR, PuroR. For generation of endogenous 3'-terminal *glmS^{M9}* fusions.	This work
pMOTag-4-*glmS^{wt}-BLAST*-5	AmpR, BlastR. For generation of endogenous 3'-terminal *glmS^{wt}* fusions.	This work
pMOTag-4-*glmS^{M9}-BLAST*-5	AmpR, BlastR. For generation of endogenous 3'-terminal *glmS^{M9}* fusions.	This work
pX-*NEO-T2A-yFCU*	AmpR, NeoR. Rescue plasmid with yFCU negative marker.	This work
pPLOT-*BLAST-T2A-APRT-mCherry*	AmpR, BlastR. Rescue plasmid to test APRT as negative marker; mCherry reporter.	This work
pPLOT-His$_8$-*ERVrecodon*-His$_8$-*PURO*	AmpR, PuroR. Plasmid for generation of ERV point mutations and *ERV$^{\Delta KISS}$*	This work
pPLOT-His$_8$-*ERVrecodon*-His$_8$-*BLAST*	AmpR, BlastR. Plasmid for generation of ERV point mutations and *ERV$^{\Delta KISS}$*	This work
pX-mCherry-*glmS^{wt}*	AmpR, NeoR. Plasmid for generation of mCherry-*glmS^{wt}* reporter line.	This work
pX-mCherry-*glmS^{M9}*	AmpR, NeoR. Plasmid for generation of mCherry-*glmS^{M9}*reporter line.	This work

* Beneke *et al.*[109]

2.1.12 Oligonucleotides

All primers listed were purchased from Metabion (Planegg, Germany), desalted without further purification with exception of long primers G00, 166, 167, and 172, that were purified by HPLC. Restriction sites are underlined. Mutations are highlighted in red. Target sites for sgRNAs are underlined. All sequences are written in the 5' to 3' orientation.

Table 2.12 List of primers used for plasmid construction

pLPC	**Number**	**Sequence**
p/SPneogRNAH/Rv	14	GATCACATGTGTGAGTTATGAGGTCTGCG
p/ pSPneogRNAH/Fw	15	GATCACATGTGATCCCTAGGTCTACAATTGCGAATTCGAGCTCTGGAG
p/KOAPRT/01sgRNAtop	84	TTGTATCGATGACGTTCTGGCAAC
p/KOAPRT/01sgRNAbott	85	AAACGTTGCCAGAACGTCATCGAT
Long arm donor *APRT*	**Number**	**Sequence**
P/LtgDNA/5'aHR1/Fw	47	GATCGGATCCAGTTGGCCTTCTCGCTCGC
P/LtgDNA/5'aHR1/Rv	48	GATCACTAGTATTAGCACCACACGCGAATTC
P/LtgDNA/3'aHR2/Fw	49	GATCGCGGCCGCAACGGGTGGAACGGCGCTG
P/LtgDNA/3'aHR2/Rv	50	GATCTCTAGACCCTCAGCGACCGCTTTGC
***yFCU* rescue plasmid**	**Number**	**Sequence**
p/p57/T2oYFCU1/Fw	120	GGCGACGTGGAGGAGAAtCCtGGtCCGGTGACGGGCGGTATGGCC
p/p57/T2oYFCU2/Fw	121	AGGGCCGCGGCAGCCTGCTGACGTGCGGCGACGTGGAGGAGAAtC
p/p57/T2oYFCU3/XbaI/Fw	122	GATCTCTAGAcgagttcttcGGCAGCGGCGAGGGCCGCGGCAGCC
p/p57/oYFCU/XbaI/Rv	123	GATCTCTAGACTACACGCAGTAGTAGCGATC
***APRT* rescue plasmid**	**Number**	**Sequence**
p/gDNA/APRT/XmaI/Fw	177	GATCCCCGGGCCGTCCCTCAAGGAAATCGGACC
p/gDNA/APRT/NcoI/Rv	178	GATCCCATGGCTACTGGTTCACCAGTAGATC
mCherry *glmS* reporters	**Number**	**Sequence**
p/TOPO/mCherry/Fw	127	GATCGGATCCATGGTGAGCAAGGGCGAGG
p/TOPO/mCherry/Rv	128	CTACTTGTACAGCTCGTCCATG
p/TOPO/*glmS*M9/Fw	129	GATCTGTACAAGTAGATGTCCAGACCTGCAGTAATTATccCGCCCG
p/TOPO/*glmS*/Rv	130	GATCGGGCCCAAGCTTAGATCATGTGATTTCTCTTTGTTC
p/TOPO/*glmS*wt/Fw	131	GATCTGTACAAGTAGATGTCCAGACCTGCAGTAATTATAG
***glmS* tagging constructs**	**Number**	**Sequence**
p/pMOTag/ApaI/Fw	173	GATCGGGCCCggttctggtagtggttccggttccggttctTACCCTTACGATGTGCCTG

p/pMOTag/NcoI/Rv	174	GATCCCATGGAGATCATGTGATTTCTCTTTGT
p/pPLOT/NcoI/Fw	175	GATCCCATGGCCACCACCACCACTGAGAATTC
p/pPLOT/BamHI/Rv	176	GATCGGATCCCCAATTTGAGAGACCTGTGCGG
***ERV* recodonised plasmid**	**Number**	**Sequence**
p/ERVrcodnhis/NdeI/Fw	215	GATCCATATGtcggatgacggcgtgcaagag
p/ERVrcodnhis/BamHI/Rv	216	GATCGGATCCtcagtgtgtcagctcctcatcttcagg
***ERV* heterologous expression**	**Number**	**Sequence**
p/deltaKISS/pET/Fw	244	GATCCCATGGCCTCAGACGATGGTGTACAGG
	245	GATCCTCGAGTGTCAGCTCCTCATCTTCAG

Table 2.13 List of primers used for genetic modifications

PF16	**Number**	**Sequence**
p1/PF16KO/Fw	133	ACGTACCATCGCTACACAGTGGCTTCGCCTgtataatgcagacctgctgc
p5/PF16KO/Rv	134	TATTCGCTTTCCTCAGCCTTCAGTTGCCAccaatttgagagacctgtgc
p4/PF16Ctag/Fw	138	AAGATCGAGAACTACCACGTGCAGCAGCACggttctggtagtggttccgg
p/PF16sgRNA5'/01	135	gaaattaatacgactcactataggCGGTGGGAAGAGAAGAGTACgttttagagctagaaatagc
p/PF16sgRNA3'/01	136	gaaattaatacgactcactataggAGAGGAGGCAGCACGAGCTCgttttagagctagaatagc
p/PF16/5'UTR/Fw	142	ATCGTCAGTGTCAAAGTGATTCT
p/PF16/3'UTR/Rv	143	AGCGTGCGCAGGCGTGATTC
APRT	**Number**	**Sequence**
p1/APRTKO/Fw	144	CTCGCTCGCTCGCACGCTCTTCACGCTCCTgtataatgcagacctgctgc
p5/APRTKO/Rv	145	TGCCGCACAGCAAAGACAAAGATTCAACTCccaatttgagagacctgtgc
p4/APRTCtag/Fw	180	CTGAGCTGCAGCGATCTACTGGTGAACCAGggttctggtagtggttccgg
p/APRTsgRNA5'	146	gaaattaatacgactcactataggCGGTGGGAAGAGAAGAGTACgttttagagctagaaatagc
p/APRTsgRNA3'	147	gaaattaatacgactcactataggAGAGGAGGCAGCACGAGCTCgttttagagctagaatagc
p/APRT/5'UTR /Fw	045	AGCAGTGCACGGTTCATGC
P/APRT/3'UTR /Rv	046	TTGTGGCATACCACCTGGC
p/APRTORF/Fw	156	ATGTCCCTCAAGGAAATCGG
p/APRTORF/Rv	157	CTACTGGTTCACCAGTAGATC

ERV	Number	Sequence
p1/LTERV1KO/Fw	148	GGTGCCTATCCTCATAGCCAGGCTCCACTTgtataatgcagacctgctgc
p5/LTERV1KO/Rv	149	TCAGTGCCTCTTGCGATATCTACACTGCCGccaatttgagagacctgtgc
p4/LTERV1Ctag/Fw	153	TACTGCCCAGAGGACGAAGAACTCACGCACggttctggtagtggttccgg
p/LTERV1sgRNA5'/01	150	gaaattaatacgactcactataggTCCCTCGCTCGCTCGCTCGCgttttagagctagaaatgc
p/LTERV1sgRNA3'/02	151	gaaattaatacgactcactataggCAGCCATGATGGCTCGCAGAgttttagagctagaaaagc
p/LTERV/5'UTR/Fw	158	TCACCAATCACCGACCCTTTC
p/LTERV/3'UTR/Rv	159	GCGTATCTCTTACACGTATACC
p2/Plot/ LtErvNHis8/Rv	166	TGTGAGACGTTCCTGTACACCATCGTCTGAGTGGTGGTGGTGGTGGTGTCGTGcatgcttgacaagtggaaga
p4/Plot/ LtErvCHis8/Fw	167	TACTGCCCAGAGGACGAAGAACTCACGCACCACCACCACCACCACCACCACCACtaaccaccaccaccactgag
p4/Plot/LtErvCStrep/Fw	172	TACTGCCCAGAGGACGAAGAACTCACGCACTGGAGCCACCCGCAGTTCGAAAAGtaaccaccaccaccactgag
p4/LTERVY303P/Fw	187	GTTCTCAAACGTCTCAAGCGTTGTCAAGTGccaTGCCCAGAGGACGAAG
p4/LTERVY303D/Fw	188	GTTCTCAAACGTCTCAAGCGTTGTCAAGTGgatTGCCCAGAAGACGAAG
P4/LTERVCPC/Fw	189	GATGCCGTTCTCAAACGTCTCAAGCGTTGTccaTGCCCAGAAGACGAAG
p/ERVsgRNAORF/02	190	gaaattaatacgactcactataggTTGTCAAGTGTACTGCCCAGgttttagagctagaaaagc
p/LTERVNterSeq/Fw	191	CCACCGCGGAGGTGTACCAG
p4/LTERVY303P/r/Fw	203	GTTCTCAAACGTCTtAAGCGTTGTCAAGTGccatgtcctgaagatgagg
p4/LTERVCPC/r/Fw	204	GATGCCGTTCTCAAACGTCTtAAGCGTTGTccatgtcctgaagatgagg
p4/LTERVY303D/r/Fw	205	TTCTCAAACGTCTtAAGCGTTGTCAAGTGgattgtcctgaagatgagg
P4/LTERVR299D/r/Fw	206	GACATCGATGCCGTTCTCAAACGTCTtAAGgattgccaggtatactgtc
p/ERVsgRNANORF/03	207	gaaattaatacgactcactataggTGGTGAGTGCCCCACCCCGCgttttagagctagaaaagc
p2/LTERVC17S/rec/Rv	208	CCCTGACACGCCCAGCTCCAGCGGGGTGGGtgattcgcccgggattgtg
p/LtErvRecodonseq/Fw	209	catatgcaccaccaccaccacc
p/LTERVsgRNAKISS/01	210	gaaattaatacgactcactataggGGCGTATCCTCCATCTTGTTgttttagagctagaaatgc

ERV	Number	Sequence
P4/LTERVKISS/Fw	211	GTGGTGCTGCGTCGGTGGCACCCTGGGTACccaaccctgccgcgcg
P5/LTERVKISS/Rv	212	TTCGATGCCGCTAGTTGTTCCTCGATCGTTccaatttgagagacctgtgc
p5/LTERV1KO/Rv/02	214	TGCCTGCCCTAGTGCGTGAGTTCTTCGTCCTccaatttgagagacctgtgc

2.2 Molecular biology methods

2.2.1 Transformation of *E. coli*

Plasmid DNA was introduced into chemocompetent XL1-Blue cells (self-made) by chemical transformation. Aliquots of 50 µL competent cells were thawed on ice and mixed with plasmid DNA. Depending on the DNA source, 10 µL or 1 µL were used for a ligation mixture or purified plasmid DNA respectively. After mixing, cell suspensions were incubated 30 min on ice, heat-shocked at 42 °C for 90 s and re incubated on ice for 5 min. Subsequently, cells were diluted with 500 µL SOC medium without antibiotics and incubated for 1 h at 37 °C and 500 rpm in a thermomixer comfort. Cell suspensions were spread on LB plates containing the corresponding antibiotic using 20 µL and 200 µL of the cell mixture. Plates were incubated overnight at 37°C and stored at 4°C for up to 2 weeks.

SOC medium 5% (w/v) yeast extract, 2% (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ autoclaved, 20 mM glucose sterile filtered (0.2 µm).

2.2.2 Extraction and purification of plasmid DNA from *E. coli*

Minipreparation of plasmid DNA was done by the alkaline lysis with SDS method described previously [132]. For high yield and purity results required in *L. tarentolae* transfections, plasmid isolation was performed with the QIAprep spin Miniprep Kit according to the manufacturer's high-yield protocol. When the alkaline method was used, 2 mL of LB medium containing the appropriate antibiotic was inoculated with a single colony of transformed bacteria and grown overnight at 37 °C and 230 rpm. Cultures were pelleted (13.000 x g, 1 min, 4°C) and resuspended in 100 µL ice-cold buffer P1. Cells were lysed in 200 µL buffer P2 up to 5 min at room temperature. The reaction was neutralized with 150 µL buffer P3. Cell debris and genomic DNA were pelleted (20000 x g, 10 min, 4 °C) and the DNA present in the

supernatant was precipitated with 600 µL isopropanol (20000 x g, 5 min, 4 °C). Isolated plasmid DNA was washed with 600µL ice- cold 70% ethanol (20000 × g, 5 min, 4 °C), dried under red light for 10 min, and dissolved in 40 µL ddH_2O.

Buffer P1	50 mM Tris, 10 mM EDTA, pH 8.0, 0.1 mg/ml RNase A
Buffer P2	0.2 M NaOH, 1% (w/v) SDS
Buffer P3	1.8 M KAc/HAc, pH 5.2

2.2.3 Standard culture of *E. coli*

For propagation of plasmid DNA 3 mL of *E. coli* strains (transformed in XL-1 blue) were grown in 3 mL LB or 2xYT medium at 37°C or 30 °C for T7 Shuffle express overnight at 200 rpm. Plasmid DNA for a single *L. tarentolae* transfection was obtained by using 6 mL of 2xYT medium following a high- yield minipreparation protocol (QIAGEN). Isolated colonies were obtained after plating a transformation experiment on LB agar plates containing the appropriate antibiotic. All strains were conserved at -80 °C in 25% (w/v) glycerol stocks.

LB-medium	25 g LB- medium, ad 1 L ddH_2O, autoclaved
LB-agar	40 g LB-agar, ad 1 L dd_H2O, autoclaved
2× YT-medium	2.5% (w/v) tryptone, 2% (w/v) yeast extract, 0.5% (w/v) NaCl, pH 7, autoclaved

2.2.4 Agarose gel electrophoresis of DNA

Nucleic acids were identified, separated, and purified through electrophoresis in agarose gels. Gels were cast in 1-2% agarose (w/v) 1x TAE buffer, 6% (v/v) Gel Red installed in a horizontal electrophoresis tank containing 1x TAE buffer, and resolved at 5-10 V/cm. Samples were mixed in 6x loading dye and loaded together with 5 µL of appropriate 1kb or 100 bp DNA ladder. DNAs were visualized and documented in a CCD digital gel imaging system.

1× TAE buffer	40 mM Tris/HAc, 1 mM EDTA, pH 7.6

2.2.5 Purification of DNA fragments

In order to clone recombinant DNA constructs, fragments were first amplified by PCR or obtained by plasmid restriction enzyme digestion before agarose electrophoresis. Then, they were cleaned and purified by silica spin column-based nucleic acid purification Wizard according to the manufacturer's instructions. This way, nucleic acids were cleaned from upstream buffer components and eluted in 50 µL ddH_2O.

2.2.6 Quantitation of DNA by UV spectrophotometry

Nucleic acid concentration and purity of samples were determined by measuring absorption at 260 and 280 nm using a NanoDrop 1000 Spectrophotometer.

2.2.7 Ligation and cloning of constructs

DNA fragments obtained by PCR or plasmid minipreparation were digested by restriction endonucleases to generate cohesive fragments used to generate recombinant vectors or to perform analytical screenings. All reactions were carried out following the manufacture's recommendations with cleaned DNAs. Experimental conditions to prepare DNA fragments for cloning and to perform analytical DNA digest are described in Table 2.14 and Table 2.15 respectively.

Recombinant plasmids were generated by ligation of fragments of the target DNA into the destiny plasmid by T4 DNA ligase. A regular ratio scheme is summarized in Table 2.16. In order to melt any reannealed cohesive termini, nucleic acids were heated at 45 °C for 5 min and briefly incubated on ice, before adding ligase and buffer. To evaluate the ligation efficiency, a negative control was usually included were the target insert was excluded in the reaction. Positive clones were identified by analytical digests or colony PCR.

2.2.8 Sanger sequencing

All DNA sequences were validated by Sanger sequencing of plasmid constructs or PCR amplicons obtained from *L. tarentolae* mutants. The service was offered by GATC Biotech

(Konstanz, Germany) or Seq-it (Kaiserslautern, Germany). Samples were prepared according to the core facilities' instructions.

Table 2.14 Assemble reaction to prepare DNA fragments with protruding ends

Restriction digest		Conditions
39 µL	ddH_2O	2 h or overnight at optimal enzyme's temperature
4 µL	Purified vector (*or insert**)	
5 µL	10x Buffer	
1 µL	Enzyme A	
1 µL	Enzyme B	
50 µL	Final volume	

*when DNA inserts were obtained by PCR, up to 45 µL of clean amplified product was digested

Table 2.15 Assemble reaction for analytical screens

Restriction digest		Conditions
14 µL	ddH_2O	2 h at optimal enzyme's temperature
3 µL	Vector	
2 µL	10x Buffer	
0,5 µL	Enzyme A	
0,5 µL	Enzyme B	
20 µL	Final volume	

Table 2.16 Assemble reaction to ligate fragment DNAs

Ligation		Conditions
8µL	ddH_2O	16 h overnight
5 µL	Vector (*digested*)	
4 µL	Insert	
1 µL	T4	
2 µL	10x Buffer	
20 µL	Final volume	

2.2.9 Amplification of DNA by the Polymerase Chain Reaction (PCR)

DNA samples were amplified by PCR to yield DNA fragments further used for cloning or for screening positive recombinant plasmids by colony PCR. The high fidelity polymerase phusion was used to the former (Table 2.17) and taq polymerase for the latter purposes (Table 2.18). Primers were designed to reach 20-25 bp length, with one G or C at the 3' end when possible and a melting temperature ranging within 50-60 °C.

To screen transformed *E. coli* cells by colony PCR, a single colony was tested by transferring it to a labelled plate and resuspending it in 20 µL of master mix.

Table 2.17 PCR assemble reaction and conditions for cloning

Components*		**Thermocycling conditions**		
36.5 µL	ddH_2O	95 °C	30 s	
10 µL	5x Buffer	95 °C	10 s	
1 µL	10 mM dNTPs	50-65 °C	20 s	x 35 cycles
0.25 µL	100 µM Fw Primer	72 °C	30 s/ kb	
0.25 µL	100 µM Rv Primer	72 °C	5 min	
1 µL	DNA template	4 °C	Hold	
1 µL	Phusion			
50 µL	Final volume			

*DMSO was included when necessary at 3% (v/v)

Table 2.18 Assemble reaction and conditions for PCR colony

Components		**Thermocycling conditions**		
17.3 µL	ddH_2O	95 °C	30 s	
2 µL	10x Buffer	95 °C	20 s	
0.4 µL	10 mM dNTPs	50-65 °C	30 s	x 30 cycles
0.1 µL	100 µM Fw Primer	68 °C	60 s/ kb	
0.1 µL	100 µM Rv Primer	68 °C	5 min	
-	DNA template	4 °C	Hold	
0.125 µL	taq			
20 µL	Final volume			

2.3 Plasmid construction

2.3.1 Construction of pLPC

In order to clone plasmid pLPC, a fragment encoding the sgRNA transcription cassette was PCR-amplified using primers 14 and 15 using pSPneogRNAH as a template. The amplicon obtained was subsequently subcloned into pLPhygCAS9 (Gift from Greg Matlashewski addgene plasmids #63556 and #63555 respectively) using the PciI restriction site to yield pLPhygCas9- sgRNAH. The Cas9 sequence was replaced with an engineered Cas9 with improved specifity fused with a skipped 2P and GFP [133] by first PCR- amplifying the construct with primers 17 and 18 using pX458 (Gift from Antonino Moltalbano) as a template and HindIII and XhoI restriction sites. In order to clone a sgRNA-encoding DNA to target the *APRT* locus, a double-strand DNA sandwich oligo with primers 84 and 85 was generated and cloned using BbsI restriction site as described previously [107]. Primer sequences are specified in Table 2.12.

2.3.2 Construction of targeting cassette with long homology arms

For the construction of pCR2.1 TOPO-5HR-*BLE*-3HR, in order to obtain a knockout targeting cassette with 500 nucleotide homology arms, the expression cassette for the bleomycin resistant marker was PCR-amplified with primers 29 and 30 and subcloned into pCR 2.1 TOPO-TA using pSPble as a template (Gift from Greg Matlashewski addgene plasmid #63561). To clone 5' and 3' homologous arms, the *APRT* coding sequences were obtained separately by PCR from genomic DNA and subcloned into pCR 2.1 TOPO-BLE. For the construction of the 5' arm, primers 47 and 48 were used together with BamHI and SpeI restriction sites. Likewise, for the 3' arm, primers 49 and 50 were employed together with NotI and XbaI restriction sites. Primer sequences are listed in Table 2.12.

2.3.3 Construction of *yFCU* and *APRT* rescue plasmids

In order to clone pX-Neo-T2A-yFCU, the yeast cytosine deaminase and uridyl phosphoribosyl transferase (*yFCU*) negative marker sequence [100] was codon optimised for *L. tarentolae* and cloned into a pUC57 (Genscript, Piscataway, USA). To fuse the T2A sequence to the yFCU

maker, three consecutive PCR amplifications were carried out with long primers 120,121,122 and 123) using plasmid pUC57-*yFCU* as a template (Jordi Vilurbina Pérez). The T2A sequence was also optimized for *L. tarentolae* using the web tool optimizer (Table 2.7). The amplicon generated was subsequently cloned into pX-*NEO* [134] using the XbaI restriction site to obtain pX-*NEO*-*T2A*-*yFCU*.

In order to clone pPLOT-*BLAS*T-*T2A*-*APRT*-mCherry, a sequence module containing the blasticidin-S deaminase (*BLAST*) and self-cleaving peptide sequence T2A, was synthesized and cloned into a pUC57 vector (General Biosystems, Durham, USA). The module was further excised and subcloned into pPLOT-mCherry [109] using SpeI and NcoI restriction sites. The *APRT* gene was PCR-amplified from *L. tarentolae* genomic DNA using primers 177 and 178 and subcloned into pPLOT-BLAST-T2A-mCherry using restriction sites XmaI and NcoI to yield pPLOT-*BLAS*T-*T2A*-*APRT*-mCherry. Primer sequences are listed in Table 2.12.

2.3.4 Construction of pPLOT-*ERV* recodonised plasmids

In order to clone pPLOT-His_8-*ERVrecodon*-His_8-*PURO* and pPLOT-His_8-*ERVrecodon*-His_8-*BLAST* used for the amplification of the targeting cassettes to obtain *ERV* mutants that cannot skip the mutation by homologous recombination, a recodonisied version of *ERV* was synthesized and cloned into a pUC57 vector (General Biosystems, Durham, USA). The new codons were selected to maintain the codon adaptation index (CAI) value from wild-type *ERV* [135]. The recodonised *ERV* sequence was PCR-amplified using primers 215 and 216 and pUC57-His_8-*ERVrecodon*-His_8 as a template, and further subcloned into pPLOTv1 puro-mCherry-puro or pPLOTv1 blast-mNeonGreen-blast [109] using NdeI and BamHI restriction sites. Primer sequences are listed in Table 2.12.

2.3.5 Construction of pX-mCherry-*glmS* reporter plasmids

In order to clone pX-mCherry-$glmS^{wt}$ and pX-mCherry-$glmS^{M9}$, an amplicon containing the mCherry gene was generated by PCR using primers 127 and 128 and the vector pLEXY_IE-blecherry4 (JenaBioscience) as a template. The amplified product was cloned into pCR 2.1 TOPO. The M9 mutation reported previously [123] was introduced by site-directed mutagenesis using primers 129 and 130, and vector pARL-GFP-*glmS* as a template. The wild-

type *glmS* and mutant M9 ribozyme encoding sequences were joined to the mCherry reporter using primers 131 and 130 and BsrGI and HindII restriction sites of pCR 2.1 TOPO-mCherry. The inserts mCherry-*glmS*^{wt} and mCherry-*glmS*^{M9} were subsequently excised and inserted into pX using BamHI and HindIII restriction sites. Primer sequences are listed in Table 2.12.

2.3.6 Construction of pMOTag plasmids for endogenous *glmS* tagging

To clone vectors used to amplify 3'-terminal *glmS* tagging cassettes compatible with LeishGEedit, four plasmids were constructed (Table 2.11): Plasmids containing the *glmS* wt and M9 sequences pMOTag-*glmS*^{wt}-4H and pMOTag-*glmS*^{M9}-4H (Gift from Roberto Docampo, addgene plasmid #106378 and #106379 respectively) were further modified to contain the puromycin *N*-acetyltransferase and blasticidin-S deaminase cassette with primer binding sites 4 and 5 from the LeishGEdit system. To generate these plasmids (i) the 3xHA-*glmS*^{wt} or 3xHA-*glmS*^{M9} sequence was PCR-amplified with primer 173 which contains the ApaI restriction site and the primer binding site 5, and primer 174 which contains the NcoI restriction site using pMOTag-*glmS*^{wt}-4H or pMOTag-*glmS*^{M9}-4H as a template. (ii) Antibiotic resistance cassettes for puromycin *N*-acetyltransferase or blasticidin-S deaminase suitable for 3'-terminal tagging were PCR-amplified with primer 175 and 176 containing the NcoI and BamHI restriction sites respectively, using pPLOTv1 puro-mCherry-puro or pPLOTv1 blast-mNeonGreen-blast as a template. (iii) Fragments generated from step (i) and (ii) were ligated using the NcoI restriction site and finally subcloned into a pMOTag-*glmS*^{wt}-4H backbone using the ApaI and BamHI restriction sites. Primer sequences are listed in Table 2.12.

2.3.7 Construction of pET-28a-*ERV*$^{\Delta KISS}$-His$_6$

The *ERV*$^{\Delta KISS}$ gene was PCR-amplified the coding sequence using primers 244 and 245 and genomic DNA extracted from the *L. tarentolae ERV*$^{\Delta KISS}$-*His*$_8$ transgenic line as a template (Robin Schumann). The resulting amplicon was cloned into pET-28a using NcoI and Xho restriction sites (C-terminal His$_6$-tag fusion). Primer sequences are listed in Table 2.12.

2.4 Protein biochemistry methods

2.4.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were separated and analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis using the discontinuous denaturing buffer system [136]. Handcast polyacrylamide gels were prepared and electrophoresed using the Mini-PROTEAN system according to manufactures' instructions. To cast gels, the resolving gel solution was poured into 1 mm glass cassettes and overlaid with 100% isopropanol. Once polymerized, the layer was covered with the stacking gel solution and the comb was placed into the glass sandwich. After polymerization, gels were directly used or stored at 4 °C in wet paper towels for up to one week. Running conditions and gel preparation are shown in Table 2.19

Table 2.19 Recipes for gels solutions and conditions for SDS-PAGE

	Stacking gel	**Resolving gel**		**Running conditions**	
	4%	**10%**	**12%**		
ddH_2O	1.4 mL	4.0 mL	3.3 mL	Voltage	200 V
30% Acrylamide/bis	0.33 mL	3.3 mL	4.0 mL	Run time	40-50 min
1.5 M Tris-HCl, pH 6.8	0.25 mL	-	-	Current /gel	
1 M Tris-HCl, pH 8.8	-	2.5 mL	2.5 mL	Initial	25-50 mA
10% SDS	20 µL	100 µL	100 µL	Final	20-30 mA
10% APS	20 µL	100 µL	100 µL		
TEMED	4 µL	4 µL	4 µL		
Total volume	~4.0 mL	~10 mL	~10 mL		

Separating gel buffer 1.5 M Tris/HCl, pH 8.8

Stacking gel buffer 1 M Tris/HCl, pH 6.8

SDS running buffer 25 mM Tris, 250 mM glycine, 0.1% (w/v) SDS, pH 8.3

2.4.2 Protein precipitation

Protein samples were concentrated when needed by trichloroacetic acid (TCA) precipitation according to standard protocols [137]. When highly diluted samples were expected, protein samples were precipitated by adding 10 volumes of ice-cold 100% acetone. The solution was incubated for 1 h at -20 °C and centrifuged (30000 x g, 1 h, 4 °C). Protein pellets were air-dried for 20 min at room temperature and resuspended in 5x Laemmli buffer.

5× Laemmli buffer 50 mM Tris/HCl pH 6.8, 25% (w/v) glycerol, 10% (w/v) SDS, 0.1% bromophenol blue

2.4.3 Production and purification of Erv$^{\Delta KISS}$

The Erv$^{\Delta KISS}$ protein was heterologously produced in *E. coli* with slight modifications of a previous protocol reported for LtErv [72]. The expression construct was transformed into *E. coli* T7 SHuffle express bearing a GroEL/S plasmid (Gift from Antonio Pierik). First, a pre-culture was prepared by inoculating 4 mL of LB medium supplemented with 50 mg/L kanamycin and 34 mg/L chloramphenicol with a single transformed colony and grown at 30 °C, 230 rpm for approximately 16 h. This pre-culture was then used to inoculate 400 mL of auto-induction medium (Table 2.20) [138] supplemented with 100 mg/L kanamycin, 34 mg/L chloramphenicol, 10 µM FAD, trace metals, and 0.1% (w/v) galactose [139]. Cultures were grown for 8 h at 30 °C and approximately 24 h at 16 °C at 230 rpm in 2000 mL baffled flasks. Cells were cooled down in an ice bath and collected by centrifugation at 4500 x g for 30 min at 4 °C. The pellet was resuspended in 10 mL lysis buffer and stored at -20 °C. The recombinant proteins were further purified using immobilized metal ion affinity chromatography with Ni-NTA agarose. Cells were disrupted by first stirring a thawed pellet supplemented with a spatula tip of DNaseI and 10 mg lysozyme on ice for 30 min. Subsequently, cells were sonicated on ice for 10 cycles of 10 s pulses with 30 s intervals. Soluble proteins were fractionated from cell debris by centrifugation for 30 min at 10000 x g at 4 °C. The supernatant was diluted with 30 mL binding buffer and binding was carried out using 250 µL of resin in batch mode for 1 h at 4 °C on gentle rotation. The resin was then loaded into a column and washed thrice with 10 mL washing buffer containing 70 mM imidazole. The protein LtErv$^{\Delta KISS}$ was finally eluted with elution buffer containing 250 mM imidazole and immediately diluted

with TCEP buffer. Samples were shipped on the day of the purification on ice and dry ice to perform crystallization trials.

Table 2.20 Auto-induction medium components

Stock	Components	Concentration
ZY	Peptone from casein	1% (w/v)
	yeast extract	0.5% (w/v)
$MgSO_4$	$MgSO_4$	1 mM
5052	glycerol	0.50% (w/v)
	glucose	0.05% (w/v)
	α-D-galactose	0.10% (w/v)
NPS	$(NH_4)_2SO_4$	25 mM
	KH_2PO_4	50 mM
	Na_2HPO_4	50 mM
Trace metals	$FeCl_3 \bullet 6H_2O$	50 µM Fe
	$CaCl_2$	20 µM Ca
	$MnCl_2 \bullet 4H_2O$	10 µM Mn
	$ZnSO_4 \bullet 7H_2O$	10 µM Zn
	$CoCl_2 \bullet 6H_2O$	2 µM Co
	$CuCl_2 \bullet 2H_2O$	2 µM Cu
	$NiCl_2$	2 µM Ni
	$Na_2MoO_4 \bullet 2H_2O$	2 µM Mo
	H_3BO_3	2 µM H_3BO_3

Lysis buffer	50 mM Na_xHPO_y pH 8.5, 300 mM NaCl, 0.25% 20 (v/v) Tween 20, 1 mM $MgCl_2$, 20 mM Inmidazole
Binding buffer	50 mM Na_xHPO_y pH 8.0, 300 mM NaCl, 20 mM imidazole
Washing buffer	50 mM Na_xHPO_y pH 8.0, 300 mM NaCl, 75 mM imidazole
Elution buffer	50 mM Na_xHPO_y pH 8.0, 300 mM NaCl, 250 mM imidazole
TCEP buffer	50 mM Na_xHPO_y pH 8.0, 300 mM NaCl, 1 mM EDTA, 1 mM TCEP

2.4.4 Determination of protein concentration by Bradford

Protein concentrations were estimated by the Bradford Assay [140]. The reaction was monitored at 595 nm in a semi-automatic microtiter plate-based assay using a microplate spectrophotometer (CLARIOstar, BMG). A standard curve was prepared using BSA in technical triplicates containing seven points covering the range of 0 to 1 µg/µL. The data was fitted to

a hyperbolic curve and parameters were determined using the formula y=a*x/(b+x). Sample concentrations were determined by using two different sample dilutions in technical triplicates.

2.4.5 In-gel protein staining by Coomassie Brilliant Blue

SDS-PAGE resolved proteins were detected by incubating gels with colloidal or non-colloidal Coomassie Brilliant Blue, either when high specify was required or when general purification steps were analysed.

For general purposes, after electrophoresis gels were incubated with Coomassie Brilliant Blue solution for 1 h or overnight at RT on gentle orbital rotation. Afterwards gels were destained in destaining solution for 1 h at RT on a rocking platform and washed twice for 10 min each in ddH_2O.

To achieve high sensitivity detection, a colloidal Coomassie Brilliant Blue solution was used as reported previously [141]. After electrophoretic separation, gels were washed three times for 10 min each in ddH_2O on a rocking platform. Detection was achieved by incubating gels in colloidal Coomassie solution overnight at RT on orbital rotation. The next day, gels were washed twice for 10 min each in ddH_2O and destained for 10-60 min in distaining solution C in agitation. Finally, gels were washed twice for 10 min each in ddH_2O.

Non-colloidal Staining solution	25% (v/v) isopropanol, 10% (v/v) acetic acid, 0.05% (w/v) Coomassie Blue G250
Destaining solution	25% (v/v) isopropanol, 10 % (v/v) acetic acid
Colloidal Staining solution	2% (v/v) phosphoric acid (85%) 10% (v/v) Ethanol (96%)**,** 5% (w/v) aluminium sulphate-(14-18)-hydrate, 0.05% (w/v) Coomassie Blue G250
Destaining solution C	10% (v/v) ethanol (96%), 2% (v/v) phosphoric acid (85%)

2.4.6 Silver staining of proteins in polyacrylamide gels

Silver staining after SDS-PAGE was employed when high sensitivity was required and samples were not further analysed by mass spectrometry analysis. The procedure was carried out as described previously [142] and all reagents were provided by the Roti Black silver staining kit.

Following electrophoresis, contaminant compounds were eliminated by fixing gels in fixative solution for 1 h at RT on a rocking platform. To increase sensitivity and contrasts, gels were first washed twice with agitation for 10 min each in 20% ethanol and twice for 10 min each in ddH_2O. Gels were sensitized for 1 minute in sensitizing solution, rinsed twice for 1 min each in ddH_2O and impregnated in silver nitrate solution for 1-2 h. To develop bands, gels were transferred promptly with gloved hands in a series of steps, thorough four different containers. First, gels were pulled out from the silver nitrate solution and dipped for 10 seconds in ddH_2O, immediately afterwards, gels were transferred into the develing solution. Depending on the result desired, gels were developed between 5 and 45 min. The reaction was stopped by incubating gels for 10 min or 2 h in stopping solution and rinsed twice for 30 min each in ddH_2O.

2.4.7 Immunodetection of proteins by western blotting

Proteins were transferred to a solid membrane after polyacrylamide electrophoresis and subsequently probed with specific antibodies. Gel-resolved proteins were blotted electrophoretically by using either the Wet/Tank Mini-Trans blot cell system or the Semi-Dry PerfectBlue system. With few exceptions, all blots shown in this work were prepared using the former system.

When performing a wet transfer, gels, Whatman papers, sponges and methanol-activated PDVF membranes were pre-incubated in ice-cold wet transfer buffer. Gel sandwiches were prepared using two Whatman papers. Blotting cassettes and cell chambers were assembled according to the manufacturer's instructions. The electrophoretic transfer was completed in 75 min at 100 V at 4 °C, where the initial and final current was usually 280 and 400 mA, respectively.

In order to transfer proteins using the semi-dry system, gels, Whatman papers, and methanol activated PDVF membranes were pre-incubated in semi-dry transfer buffer. Gel sandwiches were prepared using five Whatman papers. Blotting sandwiches, base and lid electrodes were assembled according to the manufacturer's instructions. The electrophoretic transfer was completed in 60 min at 100 mA.

Following Western transfer, the blotting efficiency was analysed by incubating the membrane in Ponceau S staining solution for 1 min. The blots were washed briefly in ddH_2O, documented

and completely destained for 20 min in TBS on a rocking platform. Afterwards, non-specific binding sites on the membranes were masked by incubating the blots with blocking buffer for 1 h at RT on a rocking platform.

To enable detection, target antigens were incubated with specific primary antibodies (Table 2.8) overnight at 4 °C on a roller shaker. The next day, the membrane was washed five times for 10 min each in TBST. Subsequently, the blots were incubated with HRP-conjugated secondary antibodies for 1 h at RT on a rocking platform. Thereafter, membranes were washed in agitation five times for 10 min each in TBST. As a final step, blots were developed by enzymatic chemiluminescence by incubating membranes with a luminol-based buffer solution and immediately analysed. Signals were detected using a CCD camera documenting system.

Transfer buffer	20 mM Tris, 150 mM glycine, 20% (v/v) methanol, 0.02% (w/v) SDS
Ponceau S solution	3% (w/v) Ponceau S, 3% (w/v) trichloroacetic acid
TBST	10 mM Tris/HCl, 0.9 % (w/v) NaCl, pH 7.4, 0.1% (v/v) Tween 20

2.4.8 Stripping and reprobing of western blots

When the analysis of more than one antigen was required in the same blot to achieve a precise comparison of specific bands, blots were recycled by stripping and reprobing. This was only carried out when antigens were identified by two unalike primary antibodies (e.g. mouse vs. rabbit).

After blot development, membranes were reused in a series of washing steps at RT on a rocking platform. First, the blots were washed for 10 min in TBS and subsequently stripped by washing five times for 10 min each in an acidic stripping buffer. To neutralize them, membranes were washed twice for 10 min each in TBS and twice for 5 min each in TBST. After this blots were reprobed by immunostaining as described in section 2.4.7

Stripping buffer	1.5% (w/v) glycine, 0.1% (w/v) SDS, 1% (v/v) Tween 20, pH 2.2

2.5 Biology and molecular parasitology of *L. tarentolae*

2.5.1 Standard cultivation of *L. tarentolae*

L. tarentolae (UC strain) obtained from Andre Schneider (University of Berne, Switzerland) was grown in axenic promastigote state at 27 °C in brain heart infusion (BHI) medium containing 10 µg/mL hemin. Cells were maintained in suspension in ventilated TC flasks upright with continuous movement on an orbital platform shaker. Cells were kept in mid-growth phase at approximately 5 x 10^7 cell/mL in 10 mL medium by diluting cultures 1:10 to 1:20 depending on parasitemia, appearance, and motility assessment. Cells were usually cultured for up to 1 month. Unless otherwise stated all standard cultivations were done in supplemented BHI medium.

BHI medium 37 g BHI powder, ad 1 L ddH2O, 10 µg/mL hemin solution

Hemin solution stock 2 mg/mL hemin chloride in 0.05 M NaOH

2.5.2 Cell density determination by cell counting

Parasitemia was determined by counting cells using a hemocytometer. Cell suspensions were diluted 1:2 or 1:3 in 1 × SSC fixation solution, assembled into a chamber, and directly counted thorough 16 squares (0.04 mm^2) using a brightfield inverted microscope.

20× SSC 3 M NaCl, 0.3 M Na_3citrate, pH 7.4

Fixation solution 10% (w/v) paraformaldehyde in 1× SSC

2.5.3 Cryopreservation and thawing of *L. tarentolae* cultures

Cells were preserved and maintained in liquid N_2. For long term storage, healthy looking mid-growth phase parasites were diluted 1:1 in sterile BHI freezing solution and aliquoted in cryotubes. Cells were first incubated at -80°C overnight wrapped in paper towels before being finally stored in a N_2 tank.

In order to thaw strains, cryotubes were thawed at 27 °C in a water bath and immediately washed with freshly pre-warmed BHI medium for 10 min at 1500 x g. Afterwards, cells were pelleted and resuspended in 2 mL of fresh BHI and incubated overnight. When high parasitemia was reached, cells were split and cultivated under standard conditions as

described in section 2.5.1. When antibiotic-resistant strains were used, cells were exposed to medium supplemented with antibiotic/s immediately after thawing at concentrations specified in Table 2.4.

BHI freezing solution BHI medium supplemented with 30% (v/v) glycerol and 10 µg/mL hemin

2.5.4 Cell lysates

Cell lysates used for general westernblot analysis were prepared by harvesting cells for 10 min at 1500 x g. Pellets were resuspended in denaturing buffer supplemented with protease inhibitors to reach a final concentration of 1 x 10^6 cells/µL of buffer. Afterwards cells were lysed by the addition of 0.25 volumes of 8% SDS and 1x Laemmli buffer solution supplemented with β-mercaptoethanol at 10% when required. Samples were boiled at 96 °C for 20 min.

2.5.5 Plasmid transfection into *L. tarentolae*

Plasmids or fragment DNAs used in stable episomal lines or classic genetic experiments respectively, were introduced in *L. tarentolae* cells by a protocol based on the Basic Parasite Nucleofector kit (Lonza). Although later adopted in this work, all DNA transfections were replaced by a better cost-effective protocol described in section 2.5.12, avoiding the use of expensive reagents and substituted with self-made buffer solutions.

Healthy-looking wild-type cultures that were passed several times for at least one week, were transfected at mid-growth phase. A total of 5 × 10^7 cells were harvested by centrifugation for 5 min at 1500 x g at RT. Cells were washed by resuspending them in 1 mL transfection buffer and then centrifuged for 5 min at 1500 x g at RT. These pellets were then resuspended in 100 µL of Nucleofactor Solution and mixed with 10-20 µg sterile DNA pre-dissolved in 10 µL ddH2O. The mixed cell DNA suspension was transferred into an electroporation cuvette and the cells were pulsed once using the program U-033 in the electroporator Amaxa Nucleofector 2b. Cuvettes were immediately washed with 500 µL pre-warmed BHI medium and cell suspensions were transferred into a 2 mL Eppendorf tube each. Cells were settled by diluting them up to 2 mL of fresh BHI without antibiotics and incubated overnight without shaking. The next day cells were pelleted by centrifugation for 5 min at 1500 x g at RT. Cells

were resuspended in approximately 300 µL BHI medium and distributed on a fresh BHI agar plate (section 2.5.7) with the corresponding selection marker. Plates were wrapped in parafilm and incubated at 27 °C (agar side facing upwards) until colonies spawned within one to three weeks after transfection.

Transfection buffer 21 mM HEPES, 137 mM NaCl, 5 mM KCl, 0.7 mM NaH_2PO_4, 6 mM glucose, pH 7.4

2.5.6 DNA sterilization

Plasmids and PCR cassettes used for transfection were heat-sterilised in a thermocycler at 94°C for 5 min without any further purification.

2.5.7 Clonal selection on BHI agar plates

Isogenic parasite lines were obtained by picking single colonies on BHI agar plates. Agar plates were freshly prepared on the day of plating, by melting sterilized BHI containing 0.08% (w/v) folic acid and 0.8% (w/v) agar in a microwave. The medium was then cooled down on top of a heat block at 60 °C and supplemented with 20 µg/mL hemin, 10% heat-inactivated FBS, and the desired antibiotic at concentrations described in Table 2.4. BHI agar was let to solidify and to dry under a laminar flow hood for 10 min.

Colonies were picked with a disposable inoculating loop and transferred into a flask containing 1-2 mL BHI with antibiotics. Cells were then cultured as described in section 2.5.1

BHI agar 3.7% BHI, 0.08% (w/v) folic acid, 0.8% (w/v) agar, 10% (v/v) heat- inactivated FBS, 20 µg/mL hemin

2.5.8 Genetic manipulations of *L. tarentolae* by CRISPR using the LeishGEdit system

Parasite mutants were obtained by reproducing the CRISPR/Cas9 system reported by Beneke *et al.* [109] in *L. tarentolae*. Donor DNA cassettes and sgRNA were obtained by PCR amplification. Transfections were performed in a cell line that expresses *Streptococcus pyogenes* Cas9 nuclease and T7 RNA polymerase (RNAP) transiently from an episome pTB007.

Plasmids pT, pPLOT and PTB007 described in further sections were the kind gift from Eva Gluenz (University of Oxford, United Kingdom).

2.5.9 Amplification of sgRNA templates

sgRNAs were generated by extending the 3' ends of two complementary oligos, a target-specific custom made oligo containing the T7 promoter and a protospacer sequence, and an overlapping oligo containing the *S. pyogenes* Cas9 sgRNA scaffold (G00 primer). The design of the target-specific sequences can be found in section 2.5.11. All polymerase reactions were performed using the HiFi Polymerase and transfected after sterilization (section 2.5.6) directly without further purification. Assemble PCR reaction and conditions are detailed in Table 2.21.

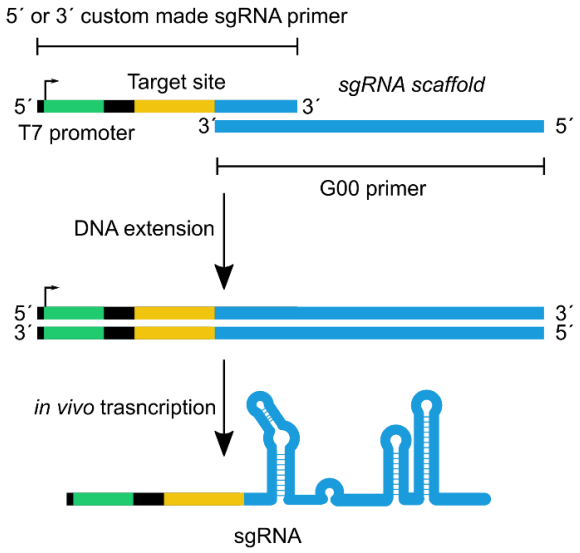


Figure 2.1 sgRNA cassette generation.
Schematic representation of the generation of the sgRNA cassette. The custom made oligo contains 24 nucleotides, the T7 promoter and 20 nucleotides that match exactly the target sequence without the PAM motif. This target-specific oligo also contains a 20 nucleotide overlap sequence complementary to a universal G00 primer. The G00 primer has the *S.pyogenes* sgRNA scaffold oligo. A DNA extension using a thermocycler is needed to extend both 3' ends generating a double-stranded DNA sgRNA cassette. This cassette is electroporated into TB007 transfectants expressing T7 RNA polymerase. An *in vivo* transcription is then initiated by the RNA polymerase generating a target-specific sgRNA. Figure modified from [109].

Table 2.21 Assemble reaction for sgRNA synthesis

Components		Thermocycling conditions		
16.5 µL	ddH_2O	98 °C	30 s	
2.0 µL	10x Buffer with Mg^{2+}	98 °C	10 s	
0.4 µL	10 mM dNTPs	60 °C	30 s	x 35 cycles
0.4 µL	100 µM Primer G00	72 °C	15 s	
0.4 µL	100 µM Primer FW target	4 °C	Hold	
0.3 µL	HiFi polymerase			
20.0 µL	Final volume			

2.5.10 Amplification of donor DNA cassettes by PCR

Donor DNA carrying selectable marker genes was PCR-amplified with primers containing 30 nucleotides target-specific sequences that served as homology arms for the repair of Cas9- induced double-strand breaks. All primer sequences can be found in Table 2.13. Donor DNA cassettes used for knockout or knockin experiments were generated by PCR amplification using pT (Figure 2.2) or pPLOT plasmids (Figure 2.3) as a template respectively. *glmS* or *ERV* chromosomal mutants (LtErv^{C17S} LtErv$^{\Delta KISS}$ and CQVYC-motif mutants) were generated with pPLOT modified plasmids pMOTag and pPLOT-*ERV* recodonised respectively (section 2.3). Primer design is described in section 2.5.11 and PCR conditions are summarized in Table 2.22

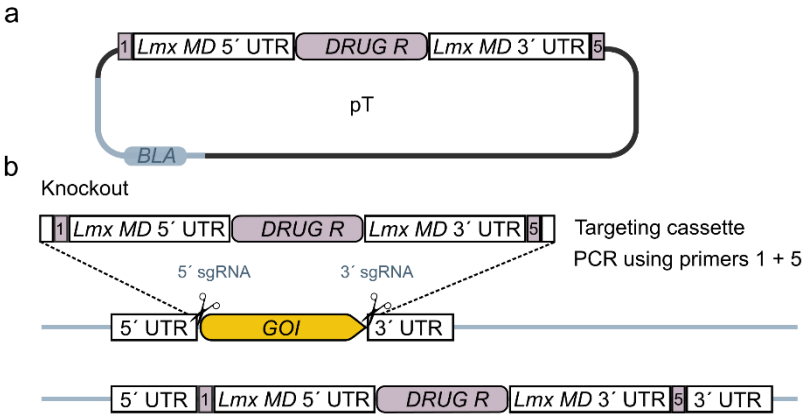


Figure 2.2 Knockout experiment.
(a) General map of pT plasmid used to generate knockout resistance cassettes or donor DNA templates. *DRUG R*, indicates either the puromycin *N*-acetyltransferase or the blasticidin-S deaminase. The antibiotic resistance cassette contains 5′ and 3′ UTR regions of the *Leishmania mexicana* malate dehydrogenase gene (MD). **(b)** The knockout cassette is generated by PCR using primers 1 and 5 containing 30 nucleotides target-specific homology arms (white boxes). The cassette is transfected along with two sgRNA targeting up and downstream the gene of interest (*GOI*) ORF sequence. Figure modified from [109].

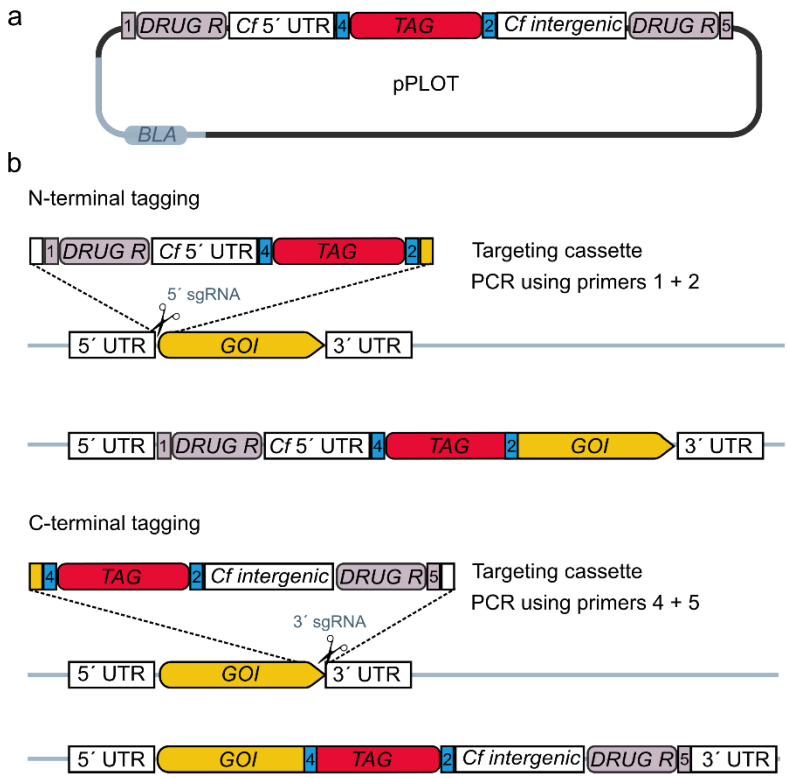


Figure 2.3 Knockin experiment.
(a) General map of pPLOT plasmid used to generate knockin resistance cassettes or donor DNA templates. *TAG* indicates the tagging construct of interest fused with 3xGS linker (blue box, primer binding sites 4 and 2). *DRUG R*, indicates either the puromycin *N*-acetyltransferase or the blasticidin-S deaminase resistance cassette. Expression cassettes contain UTR regions of *Crithidia fasciculata* (*Cf*) from the *PGKB* gene. **(b)** For N-terminal tagging the targeting cassette is generated by PCR using primers 1 and 2 containing 30 nucleotides target-specific homology arms (white and yellow boxes). The cassette is transfected along with one sgRNA targeting upstream the gene of interest (*GOI*) ORF sequence. For C-terminal tagging the targeting cassette is generated by PCR using primers 4 and 5 containing 30 nucleotides target-specific homology arms (yellow and white boxes). The cassette is transfected along with one sgRNA targeting downstream or upstream of the *GOI* ORF sequence. Figure adapted from [109].

Table 2.22 PCR assemble reaction for donor DNA cassettes

Components		**Thermocycling conditions**		
28.1 µL	ddH_2O	94 °C	5 min	
1.2 µL	100% DMSO	94 °C	30 s	
4.0 µL	10x Buffer with Mg^{2+}	65 °C	30 s	x 40 cycles
3.0 µL	25 mM Mg^{2+}	72 °C	2 min, 15 s	
0.8 µL	10 mM dNTPs	72 °C	7 min, 30 s	
0.8 µL	Primer Fw 1 or 4	4 °C	Hold	
0.8 µL	Primer Fw 2 or 5			
0.3 µL	HiFi polymerase			
1.0 µL	30 ng/µL plasmid			
40.0 µL	Final volume			

2.5.11 Primer design for LeishGEdit

Donor DNAs and sgRNA nucleic acids used to repair and induce targeted double-strand breaks respectively, were generated by PCR. To generate proof-of-concept deletions and tagging in this work, the *in silico* generated LeishGEdit primer list including all the theoretical open readings frame in *L. tarentolae* was used. For constructs that could not be obtained with primers included in the LeishGEdit primer list (His and Strep tag editions, single mutations, and deletions in the *ERV* gene) sgRNA and donor DNAs were designed manually. The targeting site selection and corresponding sgRNA sequence were evaluated with the Cas-OFFinder online tool (see table Table 2.7). The sgRNA was selected if the sequence was unique in the *L. tarentoale* genome and without computed off-target sites. The homology arms flanking the donor DNA were selected to have a length of 30 nucleotides, to be closest to the targeted double-strand break, exclude the 3 nucleotides NGG sequence, and if possible none of the 20 nucleotide sgRNA's protospacer sequence. All primer sequences used to generate targeting cassettes and sgRNA templates can be found in Table 2.13 .

2.5.12 Transfection of DNA cassettes using the LeishGEdit system

Targeting and sgRNA cassettes were transfected into *L. tarentolae* expressing Cas9 and T7 RNA polymerase transiently from the episome pTB007. Parental cultures that were passed several times for at least one week, were transfected at mid-growth phase using the transfection protocol described by Schumann Burkard *et al.*, [143]. First, 1 x 10^7 or 5 x 10^6 cells either for knockout or tagging experiment respectively, were harvested by centrifugation for 5 min at 1500 x g. Cells were subsequently washed by discarding the supernatant and suspending pellets in 1 mL of transfection buffer and centrifuging for 5 min at 1500 x g. Pellets were once again resuspended in the desired volume of transfection buffer. Afterwards, the cell suspensions were mixed with targeting and sgRNA cassettes in the proportions described in Table 2.23. The mixed cell DNA suspension was then transferred into an electroporation cuvette and pulsed once using the program X-001 in the electroporator Amaxa Nucleofector 2b. Cuvettes were immediately washed with 500 µL of pre-warmed BHI medium and cell suspensions were transferred to a 2 mL Eppendorf tube. Cells were allowed to recover by diluting them up to 2 mL in fresh BHI without antibiotics for 16-20 h. The next day, cells were

resuspended in approximately 300 µL of BHI medium and distributed on a fresh BHI agar plate with corresponding selection markers. Plates were wrapped in parafilm and incubated at 27 °C (agar side facing upwards) until colonies spawned 1 to 3 weeks later. Clonal lines were obtained as described in section 2.5.7.

Table 2.23 Transfections proportions per reaction tube for LeishGEdit

DNA cassette proportion	**Knockout experiment**	**Tagging experiment**
sgRNA templates	2	1
Knockout cassettes	2	-
Tagging cassette	-	1
Transfection		
DNA cassettes	100 µL	50 µL
Cells	150 µL	200 µL
Total Volume	250 µL	250 µL

3x Tb-BSF buffer	200 mM Na_2HPO_4**,** 70 mM NaH_2PO_4**,** 15 mM KCl 150mM HEPES pH 7.3
$CaCl_2$ stock	1.5 mM $CaCl_2$
Knockout buffer proportions	1.25 mL $CaCl_2$ stock: 4.15 mL 3x Tb-BSF: 2.1 mL ddH_2O (7.5 mL total)
Knockin buffer proportions	1.25 mL $CaCl_2$ stock: 4.15 mL 3x Tb-BSF: 4.6 mL ddH_2O (10 mL total)

2.5.13 Extraction and purification of genomic DNA from *L. tarentolae*

Chromosomal DNA from *L. tarentolae* cell lines was purified using the DNA Blood Mini Kit with the following modifications. Cultures that were passaged at least twice with the corresponding antibiotic were harvested to collect 2 x 10^7 cells by centrifugation for 5 min at 1500 x g. Cells were resuspended in 200 µL of PBS and supplemented with 20 µL proteinase K and 200 µL of buffer AL. Cells were thoroughly vortexed and the purifications steps were followed according to the manufacturer's instructions without protease incubation. Nucleic acid quantity and quality was determined as described in section 2.2.6. PCR genotyping was carried out as detailed in sections 2.5.14.

PBS 1.84 mM KH_2PO_4, 10 mM Na_2HPO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4

2.5.14 Genotyping of *L. tarentolae* by PCR

Targeted genome integration was analysed by analytical PCR using either primers annealing at the wild-type locus or at the targeting cassette and UTR region to test the amplification and size of the mutant locus or the presence of a recombinant amplicon respectively. Genomic DNA from a parental *L. tarentolae* Cas9 T7 RNAP strain was used as control. PCR products were later separated and analysed as described in section 2.2.4. Reaction conditions for PCR genotyping are described in Table 2.24.

Table 2.24 PCR assemble reaction and conditions for genotyping

Components		**Thermocycling conditions**		
40.75µL	ddH_2O	95 °C	30 s	
5.0 µL	10x Buffer	95 °C	10 s	
1.0 µL	10 mM dNTPs	60 °C	20 s	x 35 cycles
1.5 µL	DMSO	68 °C	60 s/ kb	
0.25 µL	100 µM Fw Primer	68 °C	7 min	
0.25 µL	100 µM Rw Primer	4 °C	Hold	
1.0 µL	100 ng gDNA			
0.25 µL	taq*			
50.0 µL	Final volume			

*When sequencing was required phusion polymerase was used instead. Fragments were purified and sequenced as described in 2.2.5 and 2.2.8 respectively.

2.5.15 *In situ* cysteine-cysteine trapping by alkylation treatments

To stabilize and protect potential LtErv transient protein-protein interactions via covalent disulphide bonds, three different alkylation treatments were performed on whole *L. tarentolae* cell lysates. Alkylation agents were freshly prepared, immediately prior to use and protected from light.

Treatment A was adapted from a protocol described previously [144]. *L. tarentolae* cultures were harvested (4 x 10^7 cells) at 1500 x g for 3 min, washed once in 500 µL ice-cold PBS

containing 100 mM NEM and incubated for 30 min at RT. Parasites were subsequently resuspended in 40 µL IP-Buffer supplemented with 100 mM NEM and lysed by adding 10 µL 8% SDS.

In treatment D, protein alkylation was carried out as described in [145] with the following modifications. Cells were harvested (4 x 10^7) at 1500 x g for 3 min and immediately resuspended in 500 µL of ice-cold 10% TCA. Parasites were then disrupted by freeze thawing in N_2(l) and incubated 20 min on ice. Precipitated proteins and cell debris were pelleted at 30.000 x g for 5 min, washed twice with 2 mL ice-cold acetone and dried 5 min at RT. Subsequently, the alkylation was carried out by resuspending the pellet in 50 µL NEM alkylation buffer D and incubating the samples 30 min at 37 °C.

Treatment C included a pre-treatment with diamide and is based on a protocol reported previously [146]. First, cells were harvested (4 x 10^7) at 1500 x g for 3 min and resuspended in 500 µL of warm BHI medium supplemented with 20 mM diamide. Parasites were incubated 10 min at 27 °C and further processed as above descried for treatment D.

PBS	1.84 mM KH_2PO_4, 10 mM Na_2HPO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4
Alkylation buffer D	100 mM Tris-HCl pH 7.0, 2% (w/v) SDS, 100 mM NEM

2.5.16 Stable isotope labelling by amino acids in cell culture (SILAC) in *L. tarentoale*

To label the whole *L. tarentolae* proteome, cells were grown in a defined medium reported by Silverman *et al.* [147] with modifications. Promastigotes were grown in a custom made M199 without L-lysine and L-arginine containing 10% dialysed heat-inactivated FBS, 1 M HEPES, 10 µg/mL hemin, 2 mM L-glutamine, 10 µM folic acid, 100 µM adenosine. For light medium, M199 was supplemented with 42 mg/L arginine and 73 mg/L lysine. For heavy medium, M199 was supplemented with 43.32 mg/L $[^{13}C]_6$-L-arginine and 75.44 mg/L $[^{13}C]_6$-L-lysine. The control medium consisted of light (or heavy) M199 supplemented with dialysed instead of nondialysed FBS. All solutions were sterile filtered using a vacuum filter unit (0.22 µm). To adapt cells, parasites were first cultured in the control medium for three days. Afterwards, promastigotes were passaged with fresh medium at a ratio of 1:20 every third

day. Parasites were harvested after passage 2 and 3 and subsequently used to perform affinity purifications and mass spectrometry analysis.

2.5.17 Affinity purifications of endogenously tagged LtErv-His8 and LtErv-strep

For SILAC-based affinity purifications, endogenously His-tagged or Strep-tagged Erv were purified essentially as described previously [144]. *L. tarentolae* "heavy" and "light" cells (10 mL), were cultured as described in section 2.5.16 and collected by centrifugation at 1500 x g for 5 min at 4°C. Parasites were subsequently washed once with 3 mL ice-cold PBS-NEM by centrifugation at 1500 x g for 5 min at 4 °C. Pellets were resuspended with 2.5 mL PBS-NEM and incubated for 10 min at 4 °C. Suspensions "heavy" and "light" were mixed 1:1, centrifuged at 1500 x g for 5 min at 4 °C and resuspended in 4 mL denaturing IP buffer supplemented with 1 µM or 10 µM EDTA for His-tagged or Strep-tagged Erv respectively. To prepare cell lysates, 200 µL fractions from the suspension were divided into 1.5 reaction tubes, lysed with 50 µL 8% SDS, and boiled 20 min at 96 °C. The aliquots were then supplemented with 750 µL of 2.5% Triton X-100 and incubated for 1 h at 4 °C. The lysates were subsequently cleared by centrifugation at 30000 x g for 1 h at 4°C. Baits were precipitated by adding 100 µl Ni-NTA Agarose beads or Strep-Tactin Sepharose beads for His-tagged or Strep-tagged Erv respectively. After incubation overnight at 4 °C on gentle rotation, beads were washed thrice with 1 mL denaturing IP buffer supplemented with 1.6% SDS and 1.875% Triton X-100, and thrice with 1 mL PBS. Finally, proteins were eluted with 100 µL PBS supplemented with 400 mM Imidazole or 30 mM biotin for His-tagged or Strep-tagged Erv respectively and heated at 95 °C for 10 min. Samples were stored at -80 °C until analysis.

2.5.18 Glucosamine treatment and phenotype analysis in knockdown lines

The knockdown mediated phenotype of episomal and chromosomal lines was measured after exposing mutant parasites to glucosamine in BHI or MEM medium. Stock solutions of 500 mM D-(+)-glucosamine were prepared in BHI medium or MEM. The pH was adjusted to 7.5 with 5 M NaOH before the stock solutions were sterile filtered (0.22 µm) and stored at -20 °C.

Parasites were cultured in BHI (section 2.5.1) or in MEM supplemented with 10 µg/mL hemin, 5% (v/v) FBS and 25 mM HEPES, pH 7.4.

To analyse knockdown mediated effects by an episomal mCherry reporter, clonal *L. tarentolae* cell lines transfected with vectors pX-*mCherry-glmS*^{wt} or pX-*mCherry-glmS*^{M9} were grown at 27 °C in either MEM o BHI medium supplemented with 0.1 mg/mL G418 disulfate. Cultures (10 mL) in 25 cm^2 T-flasks were diluted to an initial density of 5 x 10^6 cells/mL or an OD_{600nm} of 0.2, supplemented with 5, or 10 mM glucosamine and incubated for 24 h. Every 24 h, parasites were diluted to an OD_{600nm} of 0.2 in fresh supplemented MEM medium with 5, or 10 mM glucosamine. Three days after inoculation, cultures were harvested at 1500 x g for 5 min. The fluorescence from reporter cell lines was determined in a microplate fluorescence spectrophotometer. Parasites were resuspended in 400 µL 100 mM MES/Tris buffer pH 6.0 to a final OD_{600nm} of 1.5. Technical duplicates of the suspensions (190 µL each) were transferred to a flat-bottom 96-well microplate. The microplate was centrifuged for 5 min at 30 x g and the fluorescence was subsequently measured (λ_{Exc} = 583-15 nm, λ_{Emis} = 623-20 nm). The data was averaged from three or four biological replicates as indicated. Statistical analyses were performed in SigmaPlot 13 using the One-way ANOVA method.

To evaluate the phenotype of chromosomal *APRT-glmS*^{wt} and control lines, cells were diluted in six well plates at 4 mL to a parasitemia of 1 x 10^6 cells/mL in MEM medium supplemented with 10 µg/mL hemin, 25 mM HEPES and 5% FBS and exposed to 5, 10 and 15 mM glucosamine to induce knockdown regulation. The later effect was further tested by the addition of 10 or 100 µM of APP (section 2.5.19). A pretreatment was also evaluated in which parasites were cultured for two passages with 5 mM glucosamine, before simultaneously exposure of glucosamine and APP, as previously described. Three days after inoculation, parasitemia was measured as detailed in section 2.5.2.

2.5.19 Phenotype analysis of wt and APRT knockout lines in the presence of APP

A 500 mM APP stock was prepared as described previously [148] by first dissolving the drug in 0.05 M HCL and back-titrate the solution by adding an equal volume of 3.75 mM Tris-HCl. The stock solution was then sterile filtered (0.22 µm) and stored at 4 °C.

To assess the effect of APP on L. *tarentolae* wt and *APR*T knockout lines, exponential growing parasites were diluted to 1 x 10^6 cells/mL in BHI and distributed into a 6-well plate at 1 mL per well in three independent biological replicates. Each well was further supplemented with 3 mL drug, previously diluted in series to the desired concentration to cover 14 points (0.05- 400 µM APP). Plates were incubated for 72 h and growth was determined as described in section 2.5.2. Data was analysed using SigmaPlot 13 and EC_{50} values were calculated using a linear regression.

2.5.20 yFCU and *APRT* negative marker experiments

To probe the *yFCU* gene as a negative marker in *L. tarentolae,* the pX-*NEO-T2A-yFCU* vector (section 2.3.3) was transfected in *L. tarentolae* promastigotes. A clonal line was diluted to 1 x 10^6 cells/mL in BHI and fractionated into a 6-well plate at 4 mL per well. Cells were exposed to 3 mM 5-FC with and without 0.1 mg/mL G418 disulfate. Parasites transfected with pX-NEO alone were used as control. Parasitemia was measured 72 h post-treatment by cell counting as described in section 2.5.2

In order to test the suitability of the *APRT* as counter-selectable marker, the pPLOT-*BLAS*T-*T2A-APRT*-mCherry vector (section 2.3.3) was transfected into a *ΔAPRT L. tarentolae* line. Untransfected *ΔAPRT* and wt cells were further used as control. To assess the ability of the *APRT* marker to kill cells, a mid-log phase pPLOT-*BLAST-T2A-APRT*-mCherry line was diluted to 1 x 10^6 cells/mL in BHI and distributed into a 6-well plate at 4 mL per well. Parasites were incubated for 72 h with APP (100 and 400 µM) with and without blasticidin. Viability was determined by measuring parasitemia. To evaluate plasmid loss, cells were treated essentially as described above, except that parasites were treated with 10 µM APP and medium lacked blasticidin. The presence of plasmid was evaluated by measuring mCherry fluorescence. After incubation cells were harvested at 1500 x g for 5 min, washed once in 4 mL of PBS, and resuspended in 100 µL of PBS. Cells suspensions were further seeded into a flat-bottom 96- well microplate and fluorescence was measured on a microplate spectrophotometer (λ_{Exc} = 583-15 nm, λ_{Emis} = 623-20 nm).

CHAPTER

RESULTS 3

3.1 Establishment of the CRISPR/Cas9 system in *L. tarentolae*

3.1.1 Cas9 and sgRNA episomal encoded strategy:

In order to test the CRISPR/Cas9 system in *L. tarentolae,* we initially chose a strategy based on a previously study reported by Zhang [107]. A new plasmid pLPC (Fig 3.1a) was constructed to encode both CRISPR elements, the endonuclease Cas9 and the sgRNA expression cassette. To generate the donor DNA, a plasmid was designed to contain an antibiotic resistance cassette for bleomycin flanked by 500 nucleotide homology arms for the targeted gene and was digested to generate a linearized fragment (Figure 3.1b). Furthermore, the original Cas9 sequence from Zhang's plasmid was replaced with an engineered version of Cas9 with improved specificity and fused to a skip peptide and to the fluorescent protein GFP [133].

To test the plasmid pLPC function in *L. tarentolae*, cells were electroporated with pLPC generating a transfectant cell line and the expression of Cas9 and GFP was subsequently validated. The presence of the endonuclease Cas9 was visualized by SDS-PAGE followed by western blot probed against FLAG-tag. The expected molecular weight of 165 kDa revealed that the Cas9 encoding gene was expressed and that the skip peptide was also functional. Positive *GFP* expression was also verified by fluorescence microscopy (Figure 3.1c).

Previous studies in *Leishmania donovani* have shown that knockout mutants of the adenine phosphoribosyltransferase gene (*APRT*) exhibit resistance to the toxic prodrug 4-aminopyrazolopyrimidine (APP) when compared to wild-type cells [149-151]. To validate the CRISPR/Cas9 system, the *APRT* gene was selected as a proof-of-concept knockout experiment. A sgRNA was designed and cloned into the pLPC to target at the central coding region position of the *APRT* gene. The donor DNA was designed so that after repair, a dysfunctional disruption of the *APRT* locus would be generated by the insertion of the bleomycin resistance cassette. In order to generate Δ*APRT*, a cell line transfected with pLPC harboring the sgRNA for *APRT* was generated. Likewise, in a second round of transfection, cells were electroporated with the corresponding targeting cassette fragment and selected against bleomycin. Five unsuccessful attempts were made to generate homozygous or heterozygous knockout *APRT* mutants (Supplementary figure 7.1). We further included a donor DNA generated by PCR with 25 nucleotide homology arms and used a classical genetic approach with the targeting cassette only, in a cell line that did not express endonuclease Cas9 as a control. Genotyping with specific primers visualizing the size of the *APRT* locus revealed that in all replicates, the wild-type locus was still present (Supplementary figure 7.1b).

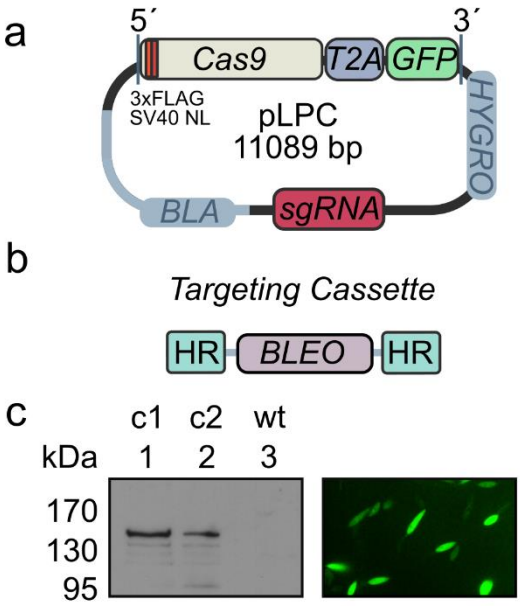


Figure 3.1 Plasmid encoded CRISPR/Cas9 strategy.
(a) Plasmid pLPC encodes both CRISPR/Cas9 system elements: a cassette expressing Cas9 with a skip peptide for GFP and a sgRNA cassette for cloning the target sgRNA. **(b)** The donor DNA included a bleomycin resistance cassette flanked by two long arms with APRT DNA allowing gene replacement by homologous double-crossover recombination. **(c)** Left: cell lysates of pLC-containing *L. tarentolae* promastigotes were resolved by reducing SDS-PAGE and analysed by western blotting against a FLAG-tagged Cas9 with an expected size of 165 kDa. Right: GFP expression revealed by fluorescence microscopy in pLC transfectants.

In summary, the system here tested was not successful presumably because of the lack of activity of one of its components. However, the skip peptide T2A was functional in *L. tarentolae.*

3.1.2 Successful genetic editions by the LeishGEdit CRISPR/Cas9 system

Employing the CRISPR/Cas9 system previously reported by Beneke *et al.* [109] we were able to successfully generate genetic editions in *L. tarentolae*. Both sgRNA and donor DNA are generated by PCR reactions, making this system simple and versatile. To test the CRISPR/Cas9 system function, a transient line expressing Cas9 and T7 RNA polymerase was generated by transfecting the plasmid TB007 (Figure 3.2a). The expression of the endonuclease Cas9 was probed by western blot using an anti-FLAG antibody visualizing the expected molecular mass of 165 kDa (Figure 3.2b). T7 RNA polymerase activity was tested by transfecting a plasmid harbouring the mCherry reporter under the control of the T7 promoter where fluorescent parasites were visible by microscopy. (Figure 3.2c).

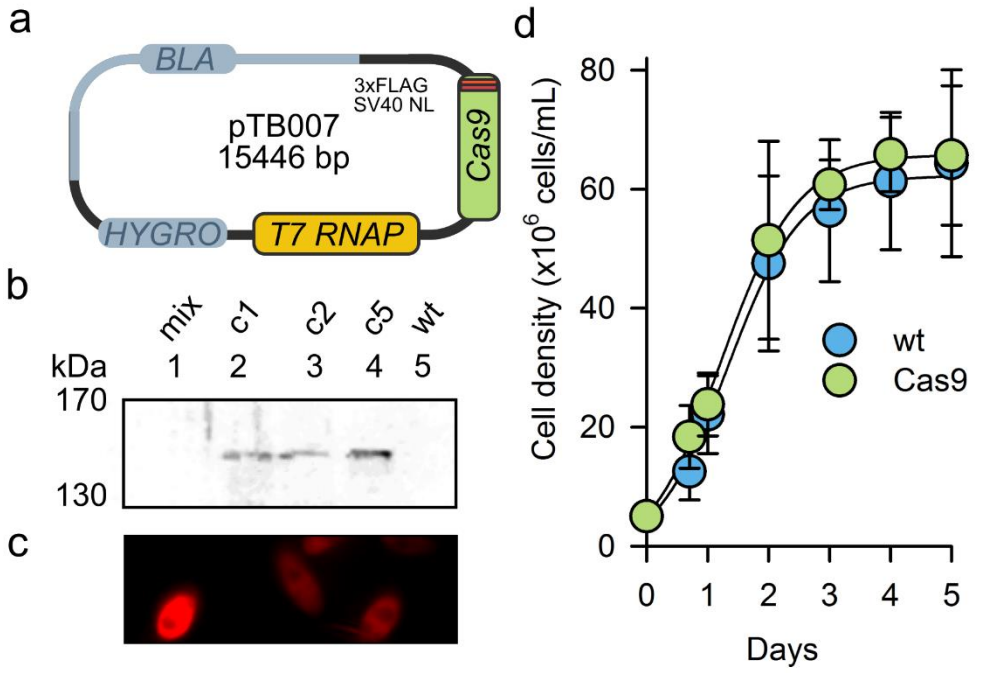


Figure 3.2 Validation of Cas9-T7 RNAP line.
(a) Plasmid with cassettes to express the endonuclease Cas9 and the RNA polymerase T7, allowing site Cas9-directed double-strand breaks and transcription of guide RNAs *in vivo*, respectively. **(b)** Protein samples obtained from TB007 transfectants were resolved by SDS-PAGE and analysed by western blot using FLAG-tag antibodies. **(c)** Fluorescence microscopy showing mCherry fluorescent parasites transfected with TB007 and a plasmid encoding mCherry under control of a T7 promoter. **(d)** TB007 transfectants and wild-type cells growth curves. Each data point represents the mean ± standard deviation from three independent biological replicates. Adapted from Turra *et al.*[152].

To exclude a potential growth impairment caused by vector TB007, growth curves were examined in TB007 transfectants. No alteration between the plasmid-harbouring line and the wild-type strain was found (Figure 3.2d).

Thereafter, the system was used to ablate and tag the *PF16* gene. PF16 is a conserved central protein in the structure of the motile flagellum in *Trypanosomes* [153], as well as in members found across eukaryotic kingdoms [154]. The *PF16* gene was previously reported in *Leishmania mexicana* to be non-essential [109], single copy and its ablation to render cells with an easily recognized immotile phenotype [109, 153]. We therefore selected the *PF16* homologue in *L. tarentolae* to test the CRISPR/Cas9 system and generate a proof-of-concept knockout and knockin mutant in the *PF16* gene (Figure 3.3). To generate Δ*PF16*, a TB007 transfectant line was co-electroporated with two sgRNA-encoding DNA templates targeting upstream and downstream of *PF16,* together with two knockout donor DNAs containing the puromycin *N*-acetyl-transferase resistance cassette. Two clonal lines were obtained and bright-field microscopy revealed that the typical flagellar beats present in wild-type cells were absent and no directional movement was observed (Figure 3.3c). The expected edition was also further characterized by PCR, showing the absence of the wild-type locus amplicon, indicating that the targeting occurred in both copies (homozygous mutant) and that the *PF16* gene was not found elsewhere in the genome (Figure 3.3b). A *PF16* knockin was also tested by genetically fusing the reporter mCherry at the C-terminus of PF16. Fluorescent microscopy showed the expected localization at the flagella (Figure 3.3f) and genotyping revealed the expected integration into both chromosomes (Figure 3.3e). Thus, knockout and knockin mutants were successfully generated in *L. tarentolae* using the CRISPR/Cas9 reported by Beneke *et al.* [109].

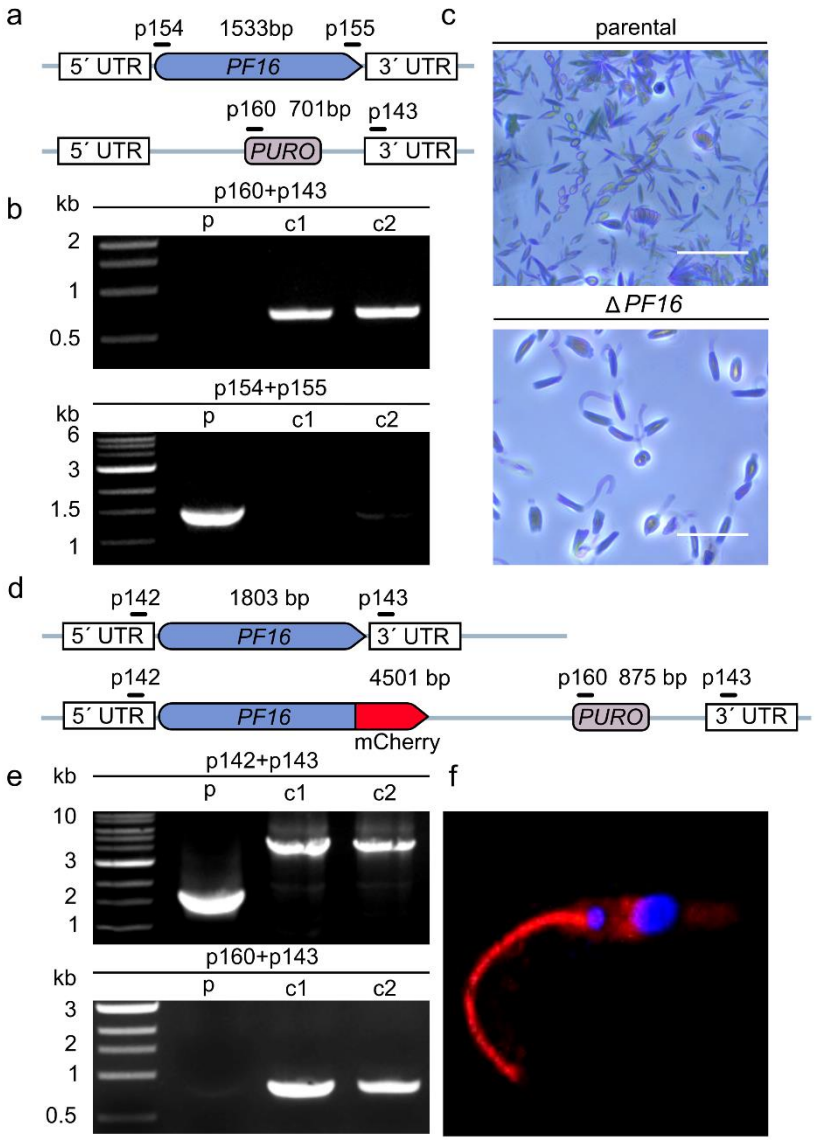


Figure 3.3 Knockout and knockin studies on *PF16*.
(a) Schematic of the *PF16* locus and the expected deletion. Black lines indicate primer binding sites aligned with the expected PCR product sizes. **(b)** Disruption of the locus analysed by PCR showing targeting of both *PF16* copies (homozygous mutant) and amplification failure of *PF16* coding sequence in clonal line mutants (c1 and c2). Parental strain (p) was used as control. **(c)** Cells observed by bright field microscopy. To denote movement, 25 consecutive frames recorded during 23.5 seconds were stacked in one single image. Top: wt cells with typical movement. Bottom: Δ*PF16* cells without a functional flagellum. Scale bars: 10 µm. **(d)** Diagram illustrating wild-type *PF16* locus and expected mCherry-tagging. **(e)** Analytical PCR visualizing the absence of wild-type amplicon (homozygous mutant) and the expected recombinant PCR product. **(f)** Cells observed by fluorescent microscopy. mCherry (red) localizes in the flagellum. The nuclear and kinetoplast DNAs were stained with DAPI (blue). Modified from Turra *et al.* [152].

3.2 Deletion of the adenine phosphoribosyltransferase gene (*APRT*) causes resistance to 4-aminopyrazolopyrimidine (APP)

As explained earlier (3.1.1), it was previously shown that the deletion of the *APRT* gene using classic genetics techniques rendered *Leishmania donovani* resistant to the otherwise toxic prodrug 4-aminopyrazolopyrimidine (APP) [151]. Drug resistance linked to a single gene is a rare striking phenotype and is attractive to use as a proof-of-concept gene mutant because of its distinctive phenotype. To generate Δ*APRT*, parasites expressing Cas9 and T7 were transfected with two sgRNA-encoding DNAs targeting the 5' and 3' UTR of the *APRT* ORF. The Cas9-induced double-strand breaks were repaired with a knockout targeting cassette that was obtained by PCR from the pT-Puro plasmid. After transfection, two clonal lines were studied and the predicted genotype was analysed by PCR. The integration of the cassettes was visualized testing the PCR amplification of a recombinant amplicon, using primers binding in the chromosome region and within the antibiotic resistance cassette. Amplification of the *APRT* gene was unsuccessful, indicating that the gene was not present elsewhere in the genome (Figure 3.4b).

After validation of the expected genotype, the susceptibility towards APP was analysed by calculating EC_{50} values for Δ*APRT* and wild-type strain (Figure 3.4c). *L. tarentolae* wild-type was susceptible to APP, as previously reported for *L. donovani* [149, 151], with an EC_{50} of 4 µM, an inhibition that can be considered potent for a drug. In contrast, the loss of the *APRT* gene did not lead to growth inhibition over the whole concentration range tested ($[APP]_{max}$ = 400 µM). The characterization of this gene-phenotype relationship was of value in the present work and was further used to test the knockdown effect of a *glmS* riboswitch (section 3.3.2) and to develop a new negative marker system (section 3.5.2).

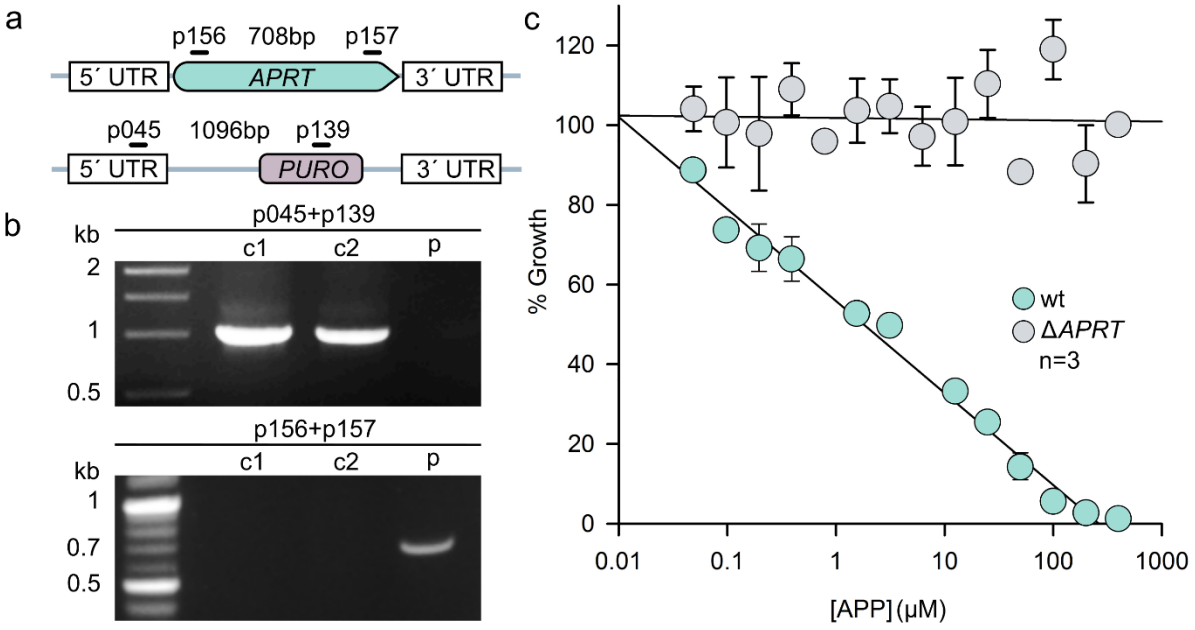


Figure 3.4 Generation of *APRT* knockout.
(a) Schematic representation of the *APRT* locus before and after deletion. Primer binding sites are indicated in lines aligned with the expected PCR amplicon size. **(b)** Top: Amplification visualizing the integration of the knockout cassette. Bottom: PCR reactions showing the absence of the *APRT* gene. **(c)** EC_{50} curves comparing the susceptibility of APP drug at different concentrations in Δ*APRT* and wild-type parasites. Each data point represents the mean ± standard deviation from three independent biological replicates. Adapted from Turra *et al.* [152].

3.3 Negative selection studies in *L. tarentolae*

3.3.1 Yeast cytosine deaminase and uridyl phosphoribosyl transferase (*yFCU*) as negative selection marker

In order to perform plasmid shuffling assays or to rescue an essential gene deletion with a rescue plasmid, a functional negative marker is necessary. This system usually consists of a plasmid harbouring a gene that is responsible for growth inhibition that is induced upon selection, therefore ensuring counter-selection of the episome. To achieve this, the negative marker *yFCU* was tested in *L. tarentolae*. Parasites electroporated with pX-*NEO-T2A-yFCU* (Figure 3.5a) were first obtained by selecting transfectants using the neomycin resistance as a positive marker. To assess if the expression of the *yFCU* marker prevents the growth of pX-*NEO-T2A-yFCU* transfectants when exposed to the prodrug 5-fluorocytosine (5-FC), two different sets of experiments were tested: Parasites were either exposed to 3 mM 5-FC in the absence of neomycin to allow plasmid curing using wild-type cells as control, or exposed simultaneously with 5-FC and neomycin to maintain selective pressure using parasites

transfected with pX-*NEO* as control (Figure 3.5.b). Parasite growth was determined 72 h post treatment. The expected non viability of pX-*NEO-T2A-yFCU* transfectants in the presence of 5-FC was not evident. The treatment here described is a representative experiment carried out for this line. Several other treatments that included longer incubation times, higher concentrations of 5-FC, or repeated on/off cycles with 5-FC were also tested without prevention of growth (Data not shown). Thus the yFCU/5-FC system was not successful as a negative selection marker in the experiments tested in this work. Presumably, because of an incompatibility of the system with *L. tarentolae* or an inefficient conversion of the 5-FC in to the toxic compound 5' fluorocytosine triphosphate.

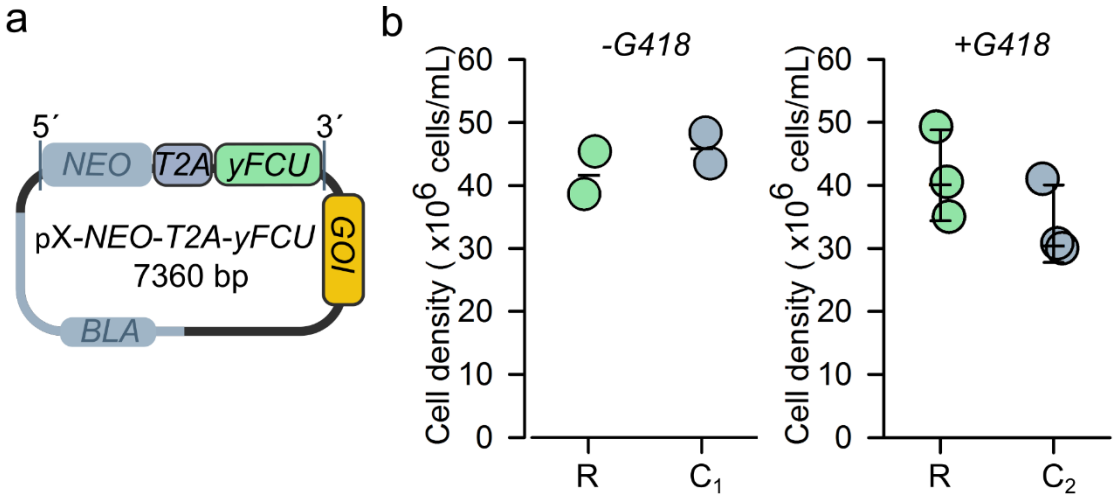


Figure 3.5 Growth effect of *yFCU*-expressing parasites with presence of 5-FC.
(a) Map of the pX-*NEO-T2A-yFCU* rescue plasmid to test yFCU function. The expression cassette bearing the neomycin positive and yFCU negative marker was fused by the skip peptide T2A. A multiple cloning site was also present to clone a future gene of interest (GOI) for rescue experiments. **(b)** Cell density of yFCU-expressing parasites treated with 3 mM 5-FC with (right) or without (left) plasmid selection pressure by G418. Data from two (left) and three (right) independent biological replicates are shown. Wild-type cells (C_1) and pX-NEO-transfectants (C_2) were used as controls.

3.3.2 *APRT* as novel negative selection marker

Given the strong APP-mediated toxicity found in *L. tarentolae* wild-type cells and the high values of resistance to the same drug in Δ*APRT* mutants (Figure 3.4), the suitability of *APRT* as a novel negative marker was further tested. We first constructed a rescue vector pPLOT-*BLAS*T-*T2A-APRT*-mCherry (Figure 3.6a), bearing the blasticidin-S deaminase and *APRT* gene fused with skip peptide T2A. We also included an mCherry reporter and an extra expression cassette to clone a gene of interest (*GOI*) for further rescue or shuffling experiments. The rescue vector was transfected into a Δ*APRT* line to yield Δ*APRT*/pPLOT-*BLAS*T-*T2A-APRT*-

mCherry. Fluorescent parasites were distinctly different in mCherry intensity from the *ΔAPRT* line allowing to assess the presence or absence of the plasmid by measuring fluorescence. The toxic effect of APP on Δ*APRT*/pPLOT-*BLAS*T-*T2A*-*APRT*-mCherry transfectants was first tested by exposing parasites to 10 µM APP excluding the blasticidin positive selection pressure and assessing growth 72 h post-treatment (Figure 3.6b). Parental Cas9/T7RNA polymerase and Δ*APRT* strains served as controls. Under these conditions, *APRT* expressing transfectants had a growth inhibition of ~ 60% when exposed to APP. When treated with APP, parental and Δ*APRT* controls had 95% or no growth inhibition, respectively. The presence of the episome was also evaluated finding a reduction of ~98% of fluorescence when APP was added, indicating an inducible APP-mediated plasmid loss or counter-selection. In order to determine if it is possible to completely inhibit growth at all when the rescue plasmid expressing *APRT* is present, parasites were forced to keep the episome by adding blasticidin in the medium while APP was added simultaneously (Figure 3.6c). Under these conditions, when exposed to 100 and 400 µM APP, parasites had ~90% and ~98% growth inhibition, respectively.

To further understand the behaviour and segregation of the plasmid-encoded *APRT* gene, the growth inhibition and plasmid presence in transfectant parasites were followed up by assessing parasitemia and fluorescence over the course of four passages in the presence and absence of APP. A representative experiment is depicted in Figure 3.6d. The experiment showed that the plasmid was partially lost even after one passage with medium containing 400 µM APP, highlighting that the plasmid curing was efficient. In summary, the APRT/APP system was tested for the first time to segregate a plasmid-encoded *APRT* gene. The results showed the potential of the technique to develop a novel negative selection marker.

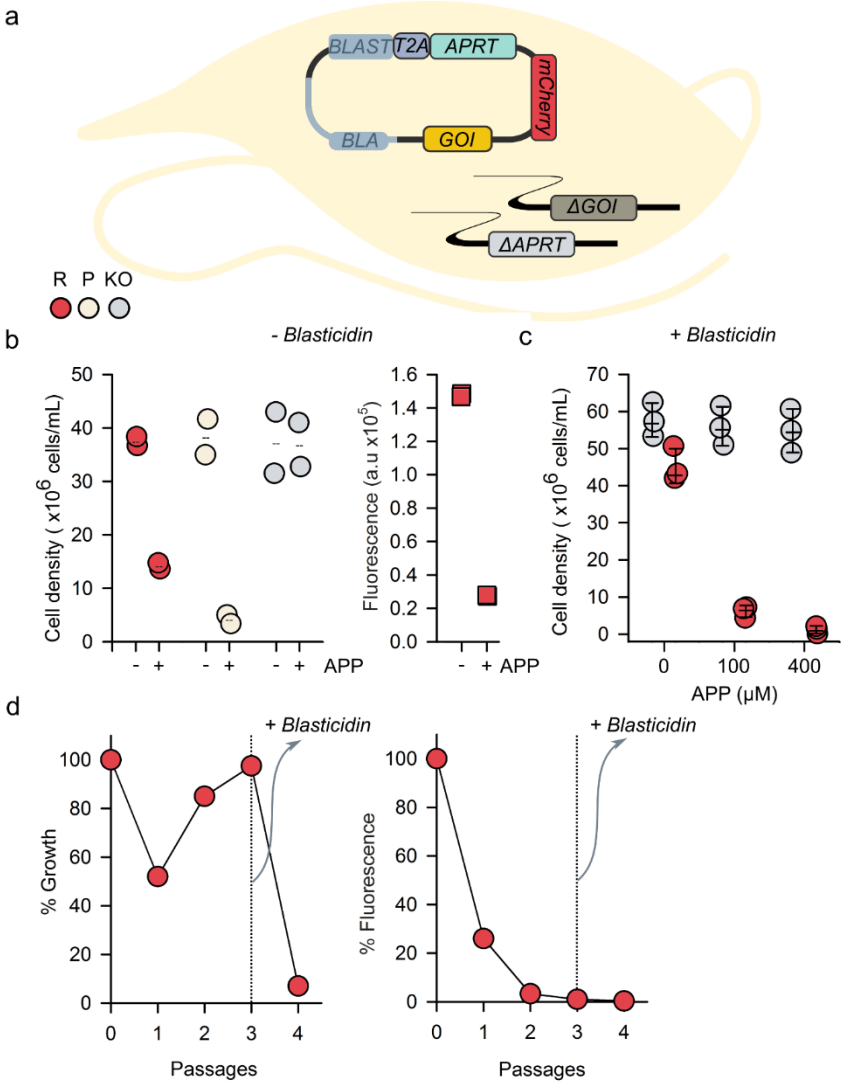


Figure 3.6 Testing *APRT* as a novel negative selection marker.
(a) Map of the pPLOT-*BLAST-T2A-APRT* rescue plasmid to test *APRT* as a negative marker. The *blasticidin* and *APRT* gene were expressed in a single cassette by fusing the genes with skip peptide 2TA. Two additional expression cassettes from pPLOT plasmid were used to express mCherry and to accommodate a gene of interest (GOI) for future rescue or shuffling experiments. **(b)** Cell density measurements to assess plasmid curing. Rescue plasmid-transfectants expressing APRT (R) in a Δ*APRT* line were exposed to 10 µM APP without positive selection pressure. TB007-transfectants (P) and Δ*APRT* (KO) were used as controls. Parasitemia and fluorescence in rescue plasmid-transfectants were measured after 72 h exposure (left and right respectively). **(c)** Growth curve to evaluate the killing potential of APP in *APRT*-expressing plasmid parasites under positive selection pressure. Rescue plasmid-transfectants (R) expressing APRT in a *APRT* knockout line were exposed to 100 and 400 µM APP in the presence of blasticidin. Δ*APRT* (KO) line served as a control. **(d)** Representative experiment following up growth rate profiles and plasmid content by fluorescence of rescue plasmid-transfectants under exposure of APP. Rescue-plasmid transfectants without adding APP were used as a control. A passage consists of a 1:20 dilution every 72 h with fresh medium containing 400 µM APP. In passage 3, the presence of the episome was tested by adding blasticidin. Growth (left) and fluorescence (right) was assayed every 72 before splitting cells with fresh medium. Each data point is expressed as a percentage of control lines (parasitemia or fluorescence) without the addition of APP. Data from two (b) and three (c) independent biological replicates are shown.

3.4 Genetic and proteomic studies on sulfhydryl oxidoreductase Erv

3.4.1 Generation of endogenous His and Strep-tagged Erv baits for affinity purification experiments

One of the possible approaches to gain a better understanding of the so-far unknown *Leishmania* Erv interactome, is to identify Erv-client/s interactions by affinity purification and mass spectrometry. For this purpose, CRISPR/Cas9 was used to direct editions in endogenous *ERV* and fuse it with DNA encoding small His_8-tag and Strep-tag II. To generate strains encoding C-terminus tagged Erv-His_8 and Erv-Strep-tag II, a sgRNA-encoding DNA was designed to target immediately downstream of the *ERV* ORF (Figure 3.7a and e respectively). A DNA cassette was generated containing 54 nucleotide homology arm at the 5' end, bearing the His-tag or strep-tag II coding sequence using a pPLOT plasmid as a template. Both elements were simultaneously electroporated in TB007 transfectants expressing Cas9 and T7 RNA polymerase. After transfection and selection with appropriate antibiotics, His_8-tag and Strep-tag II lines were genotyped by PCR (Figure 3.7b and f, respectively). The targeting of both chromosomes and the absence of a wild-type *ERV* copy was corroborated by a PCR visualizing the *ERV* locus. While the His-tagging experiment was successful with one single selection marker to target both gene copies, the Strep-tagging experiment required simultaneous selection with two resistance markers to obtain homozygous mutants. The correct integration and expected coding sequences for the His-tag and Strep-tag were further validated by sequencing of a PCR amplicon that spans the junction between the targeting cassette and the chromosomal DNA (Figure 3.7d and g, respectively). Furthermore, the presence of the His-tag in Erv-His_8 lines was shown by western blot using anti-His antibodies (Figure 3.7c). Upon validation of the correct genotype of His- and Strep-tagged Erv in each of the endogenous lines generated, growth curves were analysed. A comparison of edited and unedited cell lines revealed no growth defect that was caused by the His- or Strep-tag (Figure 3.8a and b), in the conditions tested. Finally, we conducted a preliminary affinity purification to evaluate the isolation of Erv-His_8 and Erv-Strep baits from total cell lysates. The efficiency of the purification was visualized by western blot and silver staining of eluates with the tagged baits using lysates from the parental strains as a control (Figure 3.8c and d).

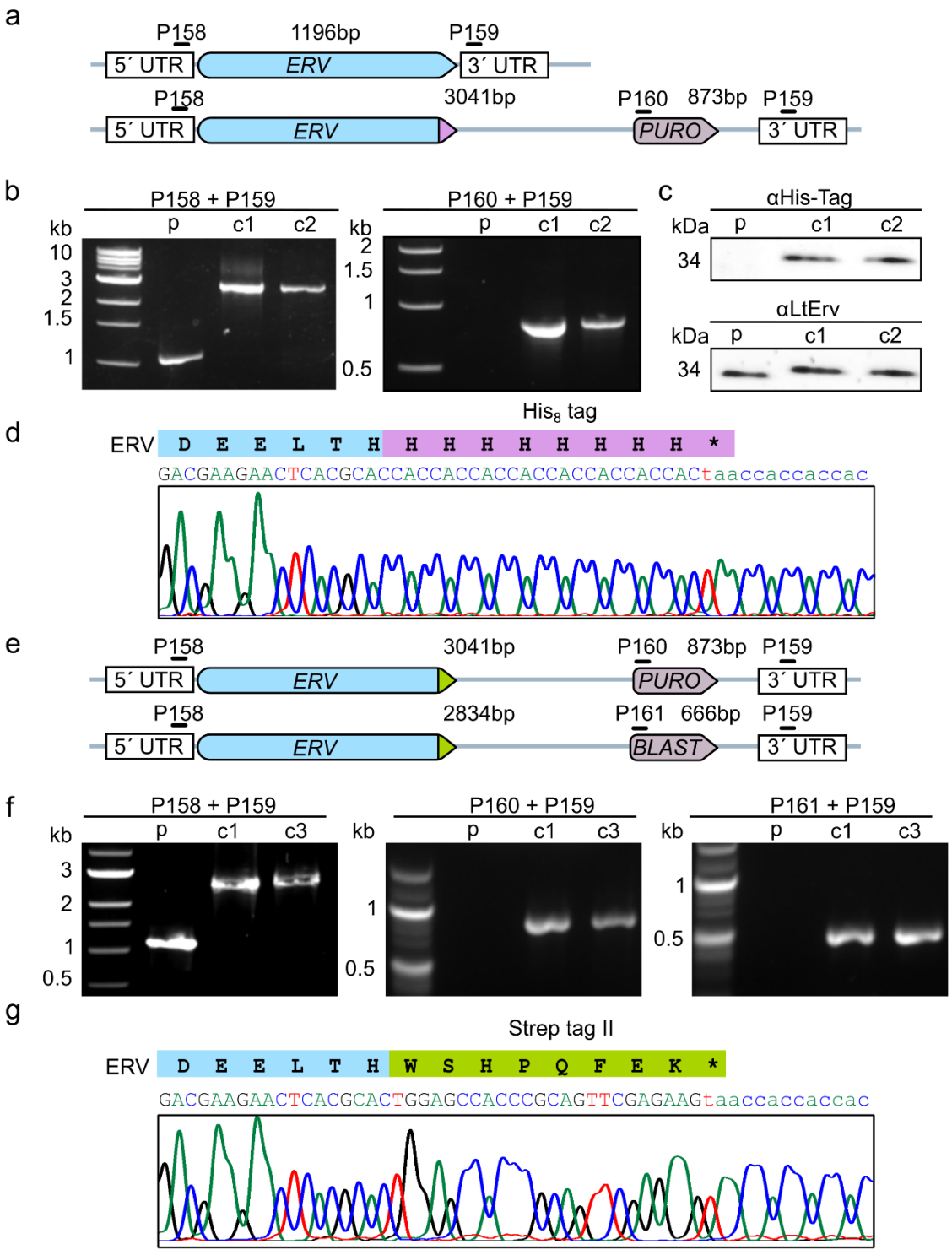


Figure 3.7 Generation of Erv-His$_8$ and Erv-Strep endogenous tagged lines.
(a) Schematic representation of *ERV* wild-type locus and expected His-tag and **(e)** Strep-tag editions. Primer binding sites are indicated in lines with corresponding product sizes for analytical PCRs. Erv-His and Erv-Strep mutants were obtained by targeting donor DNAs containing one or two drug-resistance genes, respectively. **(b)** and **(f)** Genotyping of Erv-His and Erv-strep tagged lines by PCR reactions visualizing the targeting of both chromosomes by the absence of wild-type amplicon and the expected integration of targeting cassettes. **(c)** The presence of His-tag in Erv-His$_8$ was further shown by western blot probing mutants with antibodies against His-tag. **(d)** and **(g)** Sanger sequencing chromatogram to corroborate the correct integration and expected coding sequence for the tags. Analytical PCRs were performed under my supervision by Jordi Vilurbina Pérez.

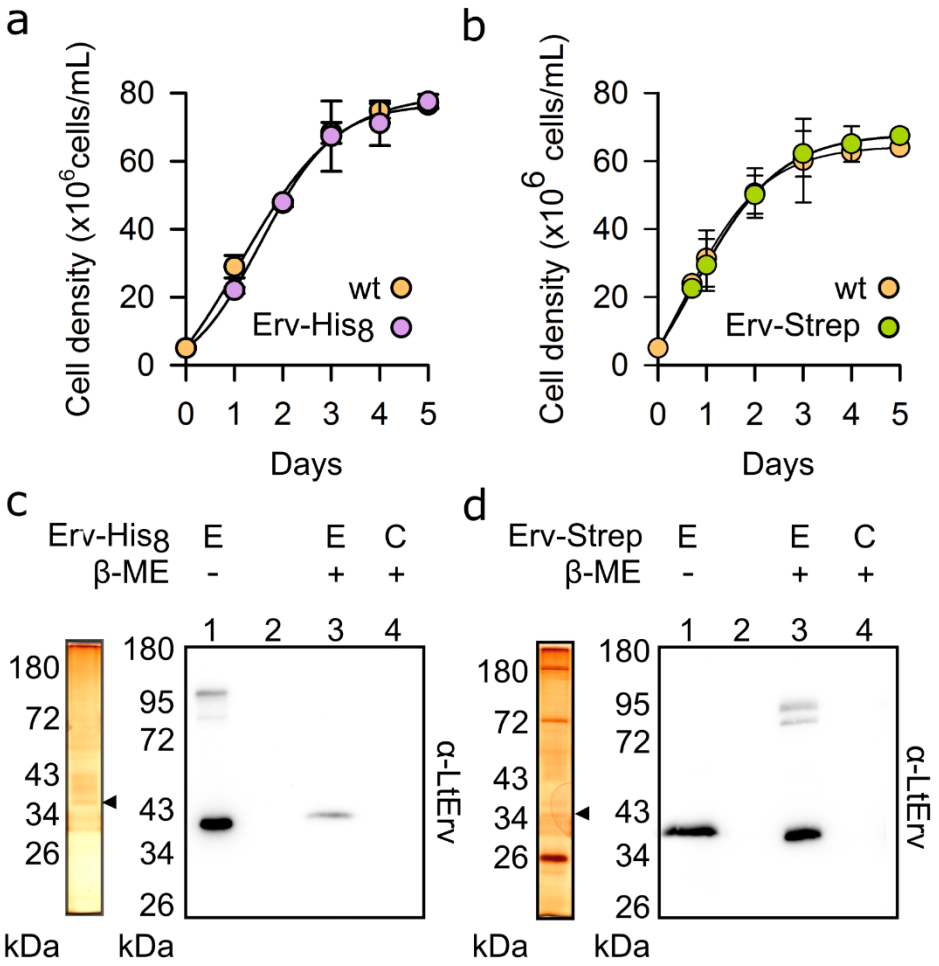


Figure 3.8 Characterization of LtErv-His and LtErv-strep tagged lines.
(a) and **(b)** Growth curves showing no growth impairment for Erv-His and Erv-strep tagged lines. Each data point represents the mean ± standard deviation from two independent biological replicates. **(c)** and **(d)** Affinity purification using Ni-NTA or Strep-Tacin to bind Erv from *L. tarentolae* lysates endogenously expressing Erv-His or Erv-Strep respectively. Eluates were resolved by SDS-PAGE and visualized by silver staining (left) and western blot (right). To evaluate the relative enrichment of baits, a control sample was included in western blot analysis for a parallel affinity purification that was performed using wild-type cell lysates. E, eluates, C, control, β-ME, β-mercaptoethanol. Cell counting was performed by Jordi Vilurbina Pérez.

In summary, the experiments showed the successful tagging and purification of His- and Strep-tagged lines.

3.4.2 Attempts to generate single point mutations in the Erv redox pair motif failed to produce viable cells

In the quest to generate a substrate-trapping Erv mutant that allows to slow down the fast resolving oxidation of Erv's unknown client/s, we aimed to introduce single point nonsynonymous substitutions in the intervening sequence of the CXXXC redox pair to generate Erv^{CPC}, Erv^{Y303P}, and Erv^{Y303D} mutants. The strategy consisted of generating donor DNA repair cassettes bearing the desired mutations together with a His_8 tag encoding

sequence, allowing for future affinity purification experiments (Figure 3.9a). After the selection of mutants with respective selection antibiotics and generation of clonal transgenic lines, mutant parasites were genotyped to corroborate that the donor DNA was targeted in both chromosomes and that homozygous mutants were obtained (Figure 3.9d). To further verify the integration of the mutation, an amplicon spanning the junction between the chromosome and the integrated cassette was sequenced to corroborate the desired genotype. Surprisingly, although the cassette was integrated and the His_8-tag encoding sequence was present, the desired point mutation was presumably skipped during homologous recombination. A representative sequencing result is shown for the attempted generation of mutant *ERV*^{Y303D}-His_8 (Figure 3.9). Similar results were obtained for *ERV*^{Y303P}-His_8 and *ERV*^{CPC}-His_8. The former results showed that presumably, after the Cas9-induced double-strand break, the cells that survived after the homologous recombination were only the ones that did not include the desired mutations. To avoid that the repair event uses a downstream sequence as a template and therefore skips the desired mutation, we fully recodonised the *ERV* sequence and constructed two new plasmids, pPLOT-His_8-*ERVrecodon*-His_8-*PURO* and pPLOT-His_8-*ERVrecodon*-His_8-*PURO,* suitable to generate targeting cassettes. This time, after repeating transfections in independent biological duplicates using two antibiotics to target both chromosomes to generate *ERV*^{Y303D}-His_8, *ERV*^{Y303P}-His_8 and for *ERV*^{CPC}-*His*$_8$, no viable cells were obtained. Thus, these experiments revealed that modifications at the tyrosine in position 303 are lethal, suggesting an essential role of this amino acid in the Erv redox catalysis.

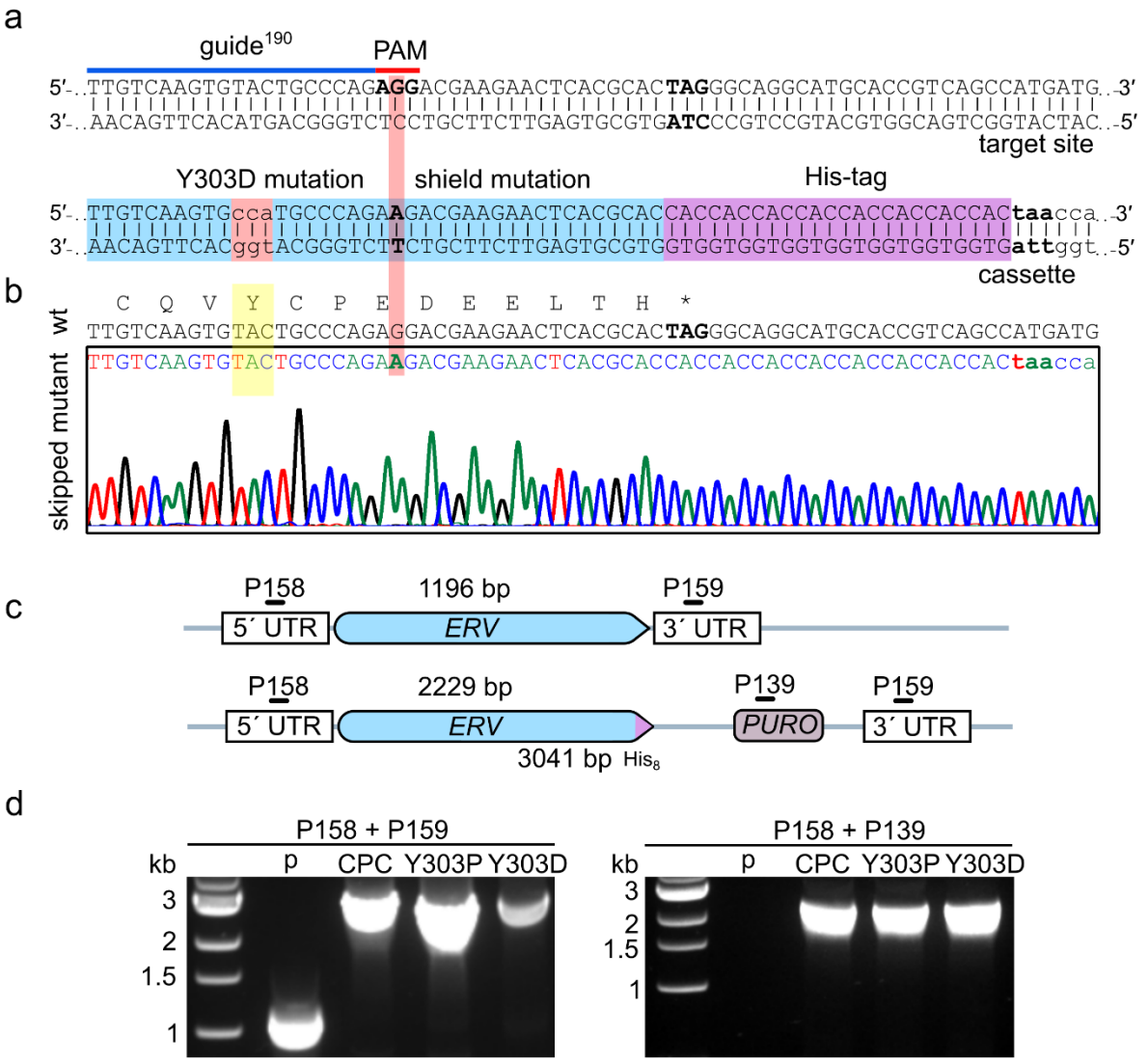


Figure 3.9 Unsuccessful attempts to generate intervening sequence mutants in the redox pair of Erv.
(a) Schematic for single point mutation editions of the Erv redox pair intervening sequence. The strategy for ErvY303D is shown as an example. A guide was selected to introduce a double-strand break close to the codon for amino acid 303. Homologous repair was induced by co-transfecting a donor DNA carrying the desired mutation in the left homology arm. To prevent the re-cleavage of the edited mutants, the PAM sequence was silently mutated (shield mutation). **(b)** Sanger sequencing chromatogram showing puromycin-resistant mutants for which the desired mutation was skipped while the shield mutation and His-tag were integrated. **(c)** Schematic representation of the ERV locus and expected editions. Primer binding sites are indicated in lines with corresponding product sizes for analytical PCRs **(d)** Genotyping results for clones obtained after transfections to generate Erv^{CPC}, Erv^{Y303P}, Erv^{Y303D}. Left: PCR reaction visualizing of the ERV locus showing the targeting of two chromosomal copies. Right: PCR reaction amplifying a recombinant amplicon is only expected in mutant parasites.

3.4.3 The conserved Erv clamp cysteine 17 is dispensable for *L. tarentolae* viability

In order to test if the clamp-forming cysteine in position 17 is structurally related to the essential function of Erv in *L. tarentolae,* we attempted to generate a Erv^{C17S} mutant using a single point mutation strategy (Figure 3.10a). Two donor DNAs using pPLOT-His_8-*ERVrecodon*-His_8-*PURO* and pPLOT-His_8-*ERVrecodon*-His_8-*Blast* as templates were obtained by PCR amplification in which the substitution mutation was present in the 3' homology arm (Figure 3.10b). After transfection of the targeting cassettes with the corresponding sgRNAs, two cell lines were obtained and further characterized. Diagnostic PCRs to genotype both lines were carried out to test the loss of the wild-type locus, showing that both clones were homozygous mutants (Figure 3.10c). To visualize the integration of the cassettes, a PCR using a forward primer located within the blasticidin-S deaminase or puromycin *N*-acetyltransferase genes and a reverse primer located in the 3' UTR region of *ERV* was used. To further exclude that the viable cells obtained had an ectopic wild-type copy elsewhere in the genome or even in an episome, primers that only amplified the wild-type ORF were used, showing the absence of wild-type *ERV*. The desired genotype was further corroborated by the sequencing of a PCR amplicon spanning the junction between the chromosome and the integrated cassette, showing the desired mutation (Figure 3.10b). The presence of the N-terminal His_8 tag included in this modification was also assessed by a western blot showing the expected signal when decorated with an anti-His tag antibody (Figure 3.10d). Finally, to evaluate if the His_8-Erv^{C17S} mutant conferred a fitness cost compared to the parental cell line, a growth curve was carried out in BHI medium showing no significant difference when compared to the control line (Figure 3.11). In summary, genotyping and growth curve analysis provided evidence that the Erv cysteine 17 is dispensable for *L. tarentolae* viability and that N-terminal His_8-tagging did not cause a growth defect.

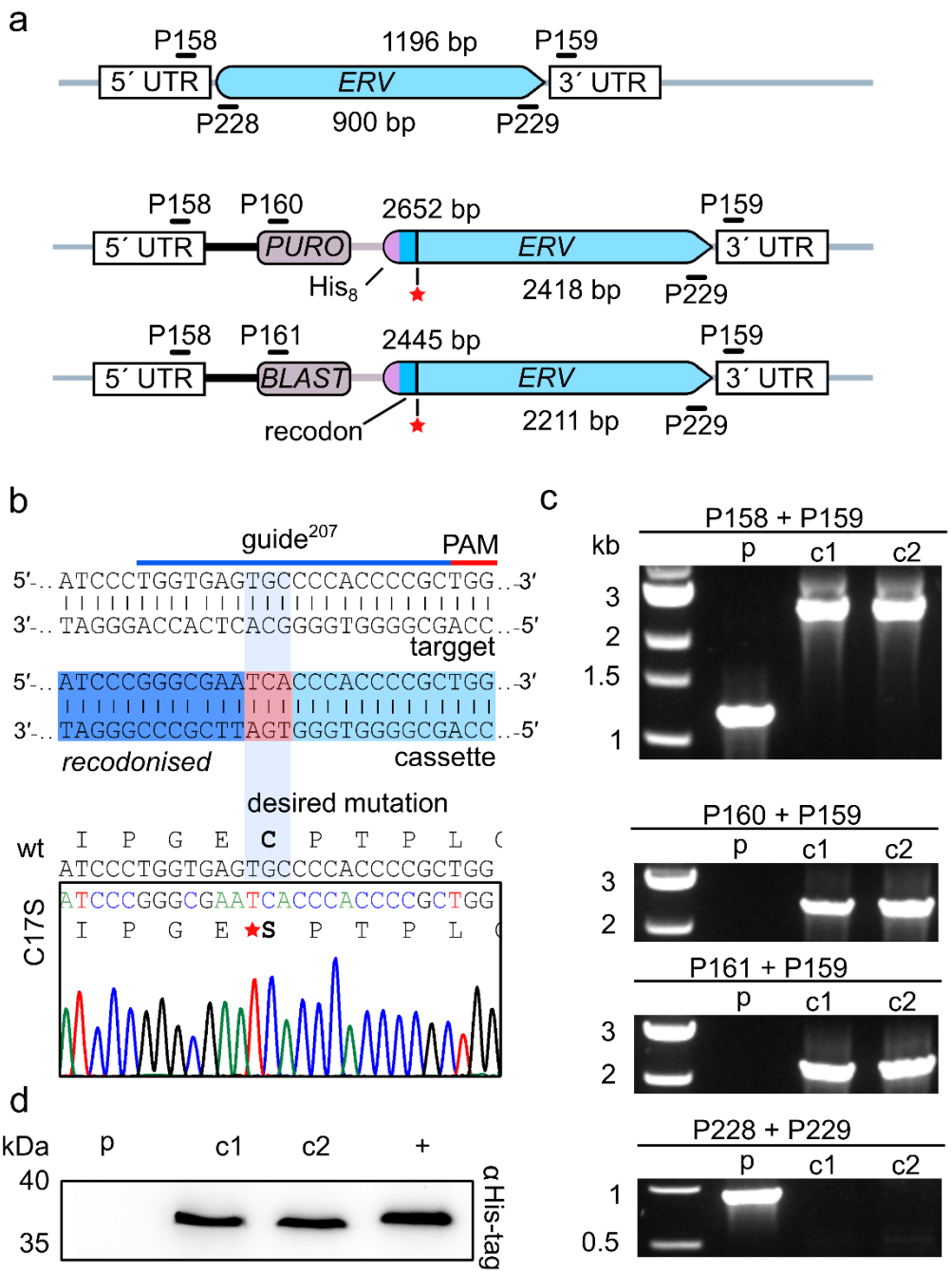


Figure 3.10 Generation of the clamp-forming cysteine mutant His_8-Erv^{C17S}.
(a) Schematic of the wild-type *ERV* locus and expected single nonsynonymous point mutation by targeting the gene using two resistant cassettes. Primer binding sites are depicted in black lines with expected PCR product sizes. **(b)** Schematic of the strategy used to generate a single cysteine replacement in position 17. Top: $sgRNA^{207}$ designed to recognize the target sequence and induce a Cas9-mediated DNA double-strand break 6-7 nucleotides downstream of the desired mutation. A donor DNA cassette was designed to contain in the right homology arm the desired mutation. To avoid skipping of the mutation during homologous recombination, the antibiotic resistance cassette was directly flanked in the right arm by a recodonised Erv (N-terminal) until amino acid 17 (dark blue) followed by the typical 30 nucleotide homology arm (light blue). Bottom: Successful C17S point mutation was corroborated by Sanger sequencing of a PCR amplicon. **(c)** Analytical PCR visualizing the targeting of both chromosomes and expected integration of puromycin and blasticidin cassettes in two mutant clonal lines (c1 and c2). The existence of an *ERV* wild-type copy in the mutant parasites was probed by using primers that only amplify the *ERV* wild-type ORF. **(d)**. The presence of the N-terminal His_8-tag also introduced during the edition was visualized by western blot using mutant parasite lysates (c1 and c2) and decorating with anti His-antibody. The Erv-His_8 line served as a positive control (+).

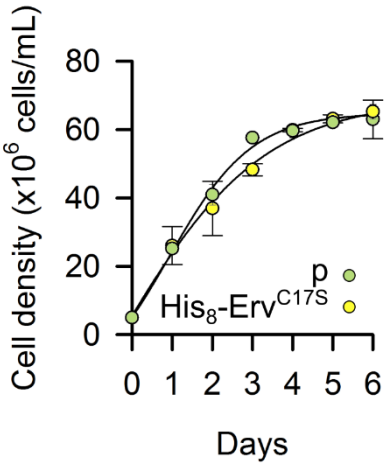


Figure 3.11 Growth curve of His_8-Erv^{C17S}.
Growth curve of His_8-Erv^{C17S} mutant analysing a putative fitness defect relative to parental line (p). No significant difference was observed. Each data point represents the mean ± standard deviation from two independent biological replicates.

3.4.4 Attempts to ablate the *ERV* gene failed to produce viable cells

The essential role of *ERV* for cell viability was already shown in *Leishmania infantum* using classic genetics [131]. Since most of the interpretations for the Erv mutants in this work assume that Erv is also essential in *Leishmania tarentolae,* we further reproduced this result by attempting to ablate the *ERV* coding sequence using CRISPR/Cas9. Two knockout targeting cassettes containing the puromycin *N*-acetyltransferase and blasticidin-S deaminase genes, together with 2 sgRNA-encoding DNA templates targeting upstream and downstream the *ERV* coding sequence were co-transfected in a TB007 line expressing Cas9 and T7 RNA polymerase. The transfections were carried out in two biological replicates and no colonies were obtained. Thus, the data suggest that Erv is essential for survival in *L. tarentolae.*

3.4.5 The Kinetoplastida-specific second (KISS) domain of Erv is dispensable for *L. tarentolae* viability

One of the most conspicuous features of all Erv homologues from kinetoplastid parasites is a long C-terminal domain comprising almost 50% of the total protein sequence, termed the Kinetoplastida-specific second (KISS) domain (Figure 3.12a) [72]. To test whether this additional region is essential for cell viability and to gain knowledge about its structure-function, 139 amino acids belonging to the KISS domain were deleted, fusing the C-terminal redox shuttle pair directly to the core FAD domain to obtain $ERV^{\Delta KISS}$-His_8 (Figure 3.12b). The

strategy used to generate this mutant consisted of the design of a sgRNA that induced a double-strand break near the coding sequence of amino acid 122. The donor DNAs were designed so that after subsequent repair the desired wild-type KISS sequence would be excluded (Figure 3.12d). These targeting cassettes were generated using pPLOT-*His$_8$-ERVrecodon-His$_8$-PURO* and pPLOT-*His$_8$-ERVrecodon-His$_8$-BLAST* as templates. The transfection was performed as usual using *L. tarentolae* cells expressing Cas9 and T7RNA polymerase, together with the sgRNA-encoding DNA template and two targeting cassettes. After electroporation and clonal selection, two lines were isolated and further characterized. PCR analysis of these clonal transfectants verified the loss of the wild-type locus, indicating the generation of homozygotes lines (Figure 3.12e). To further corroborate the insertions, the presence of recombinant PCR products was tested for both chromosomes by using primers binding to the UTR *ERV* locus region along with marker-specific primers that bind within the targeting fragment. In order to exclude ectopic recombination events where the wild-type copy is integrated somewhere else in the genome, or present in the episomal form, specific primers amplifying the wild-type copy were also tested revealing no amplification in the *ERV$^{\Delta KISS}$* lines. Because we were also interested in generating mutants to perform affinity purification experiments, we tested the presence of the His tag by western blot analysis using anti-His antibodies (Figure 3.12f). To our surprise, the Erv$^{\Delta KISS}$-His$_8$ mutant showed a significantly altered band pattern when compared to wild-type Erv-His$_8$. Under non-reducing conditions, full-length wild-type *Lt*Erv typically remains almost entirely as a monomer of 34 kDa. In contrast, the Erv$^{\Delta KISS}$-His$_8$ mutant showed a marked decreased signal for the expected monomer of 19 kDa, and a signal was enriched in the 34 kDa range. The disulphide nature of this difference was further corroborated by analysing the mutant under reducing conditions. In this case, both mutant and wild-type Erv showed the expected monomer size of 19 and 34 kDa, respectively. To evaluate if the viable Erv$^{\Delta KISS}$-His$_8$ exhibited a growth impairment, mutant growth rates were compared to parental lines. No significant differences were observed (Figure 3.13). Thus, the KISS domain appears to constitute a dispensable element of *L. tarentolae* Erv and Erv$^{\Delta KISS}$-His$_8$ has a shifted equilibrium towards mixed or/and dimer disulphide formation.

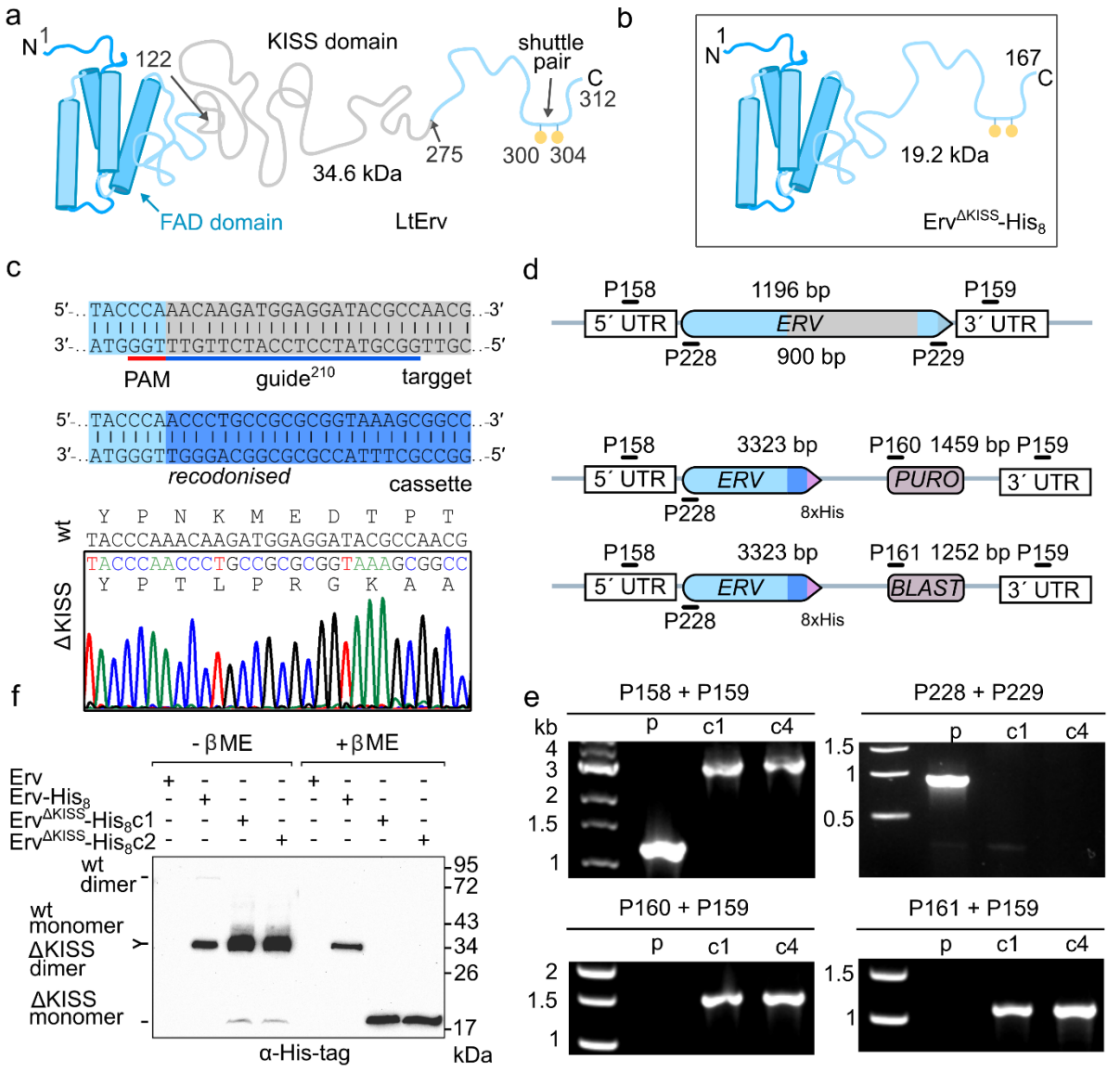


Figure 3.12 Generation of Erv$^{\Delta KISS}$-His$_8$.
(a) Schematic representation of the *L. tarentolae* Erv structure depicting the FAD, KISS domain and active redox pair. Numbers indicate amino acid positions. **(b)** Structure representation of a C-terminally His-tagged mutant that lacks the KISS domain but contain the redox-active CX3C-motif. The expected molecular mass of the mutant is given. **(c)** Schematic of the edition planned to delete a genomic fragment encoding the KISS domain. Top: sgRNA^{210} targeting a position close to a codon coding for amino acid 122 within the KISS domain encoding sequence (grey). The donor DNA had in the left homology arm the typical 30 nucleotide homology region for *ERV* (light blue) upstream a recodonised fragment (dark blue) encoding a C-terminal fragment that contains the redox pair and His$_8$-tag. Bottom: The successful deletion of the KISS domain was validated by Sanger sequencing of a PCR amplicon. **(d)** Scheme of the *ERV* locus depicting the expected KISS deletion using two selectable markers. Black lines indicate primer binding sites along with expected PCR product sizes. **(e)** Genotype analysis by PCR showing the generation of a homozygous mutant and the correct integration of both targeting cassettes in two mutant clonal lines (c1 and c4). The absence of a wild-type *ERV* sequence was probed by PCR including a reverse primer (p229) that only binds to the 3′ end of the wild-type *ERV* gene. **(f)** Western blot analysis of Erv$^{\Delta KISS}$-His$_8$ mutants cell lysates under non-reducing and reducing conditions. A band matching the expected mutant dimer and monomer was observed under non-reducing and reducing conditions respectively. Strains Erv-His$_8$ and Erv served as controls.

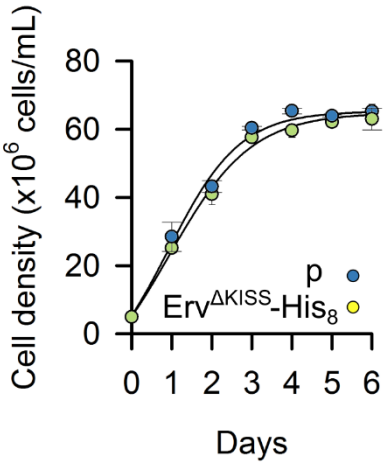


Figure 3.13 Growth curve of Erv$^{\Delta KISS}$-His$_8$.
Growth curve of mutant Erv$^{\Delta KISS}$-His$_8$ in BHI medium assessing a potential growth defect relative to the parental line (p). No significant differences were observed. Each data point represents the mean ± standard deviation from two independent biological replicates.

3.4.6 Alkylation treatments of Erv$^{\Delta KISS}$-His$_8$ revealed putative mixed disulphide intermediates

Although there is yet no direct evidence showing which is/are the Erv client/s in Kinetoplastida, comparative structural and evolutionary analysis together with experiments *in vitro* strongly support that the essential function of Erv in *Leishmania* is the disulphide bond formation of substrate/s via a disulphide exchange reaction [20, 57, 70, 72, 73]. Therefore, one of the major challenges in identifying an Erv interactome is that interactions that form and transfer disulphide bonds can be very short-lived [155]. This is because of an intramolecular reaction involving the attack of the mixed-disulphide bond by a free cysteine subsequently resolving the complex. To prevent this thiol-disulphide rearrangement, we used three chemical treatments (A, C, and D) in an attempt to stabilize the potential interactions in both Erv-His$_8$ and Erv$^{\Delta KISS}$-His$_8$ in whole-cell lysates. In treatment A, *N*-ethylmaleimide (NEM) was used in intact cells to irreversibly block free cysteine thiols and prevent them from attacking disulphide bonds. Treatments C and D consisted of a combined strategy that involved strong oxidation of whole-cell lysates with diamide and/or acid quenching of thiolates with trichloroacetic acetic acid (TCA) to decrease the reactivity of thiolate anions, followed by alkylation of free thiols. After treating cells, the visualization of high molecular weight complexes containing Erv was analysed by western blots under non-reducing conditions using anti-His and anti-Erv antibodies (Figure 3.14a). Full-length Erv showed the

expected monomer size around 35 kDa under all conditions, whereas treatments C and D made apparent a band in the 60 kDa range that is compatible with the predicted dimer form. In contrast, Erv$^{\Delta KISS}$-His$_8$ showed a strong upper pattern of higher molecular species above the expected size of the dimer form (38 kDa). To further test the nature of these higher molecular weight bands found in Erv$^{\Delta KISS}$-His$_8$, treatment C was repeated and visualized with an anti-Erv antibody (Figure 3.14b). When reduced with β-mercaptoethanol, higher bands disappeared and were replaced by the expected 19 kDa monomeric size for Erv$^{\Delta KISS}$-His$_8$, validating the detection of a homo or/and hetero-complex via disulphide bonds.

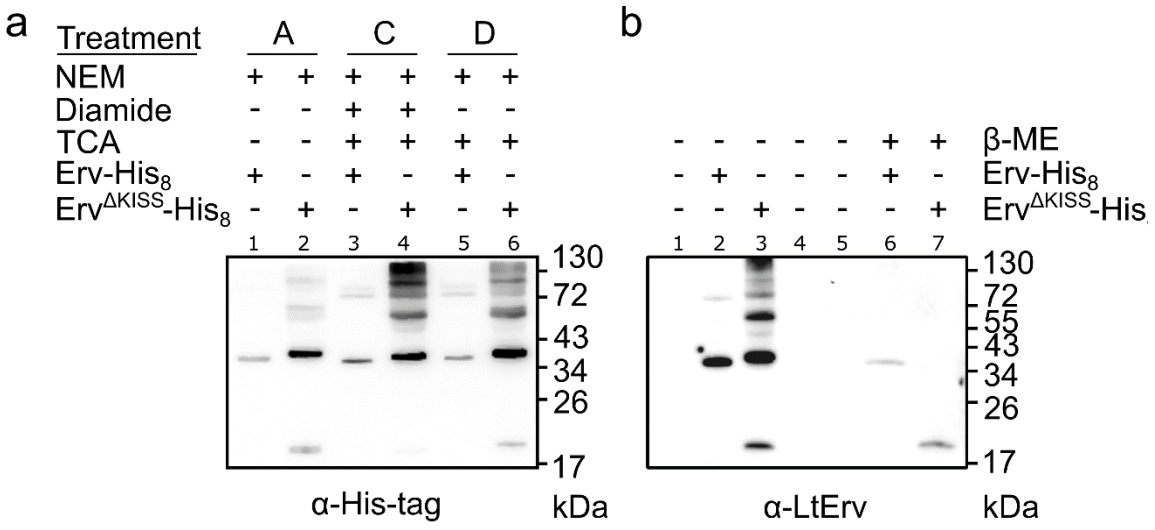


Figure 3.14 Alkylation treatments using Erv-His$_8$ and Erv$^{\Delta KISS}$-His$_8$.
free thiols from whole-cell lysates were protected to conserve potential mixed disulphides. **(a)** Western blot analysis testing three distinct alkylation treatments of whole-cell lysates separated by SDS-PAGE under non reducing conditions. Mutant Erv$^{\Delta KISS}$-His$_8$ showed His-tagged specific high molecular mass bands relative to the expected dimer (38.4 kDa). In contrast, Erv-His$_8$ showed mainly bands matching the size of monomer and dimer forms (35.7 and 71 kDa respectively). **(b)** Western blot analysis reproducing treatment C. Whole-cell lysates were separated by SDS-PAGE under non-reducing and reducing conditions to validate that Erv$^{\Delta KISS}$-His$_8$ was forming high molecular mass complexes via disulphide bonds.

3.4.7 A chemically defined *L. tarentolae* medium is suitable for stable isotopic labelling of amino acids in culture (SILAC) experiments

A chemically defined growth medium is essential when performing SILAC experiments in cell culture. SILAC consists of cells growing with either normal isotopic abundance amino acids or with its stable isotope counterparts until the whole cell proteome is labelled. To ensure incorporation of isotopes, usually two naturally occurring amino acids present in the medium need to be replaced. For this, one of the limitations of SILAC experiments is that a completely defined medium is required. Before this method was further used to perform affinity purifications coupled with mass spectrometry (AP-MS) as is detailed in the following section, a defined medium M199, previously reported in *L. donovani* [147] together with five different conditions, were tested in *L. tarentolae* (detailed in Figure 3.15a). Limitations in finding a suitable defined medium formulation include the replacement of serum with dialyzed serum and not including an essential component or cofactor that the cell is auxotroph for [156]. Considering that *Leishmania* spp. are auxotrophic for purines, pterins and folate [38, 157, 158] we tested three conditions (conditions 1-3) where the base medium M199 was further supplemented with adenosine, 6-hypoxanthine, 6-biopterin, and folic acid as reported for other defined medium formulations in *Trypanosomes* [32, 156, 159]. In all conditions tested, cells were viable, obtaining optimal growth rates (Figure 3.15b and c). Although serum-free conditions three and four allowed cells to divide, the morphology was typical for unhealthy late-phase parasites. Condition five was further selected and analysed for the incorporation efficiency of L-lysine-6-^{13}C and L-arginine-$^{13}C_6$ in the proteome by mass spectrometry at different passages (Figure 3.15d). High degrees of label incorporation of isotopes (accounting for a peptide ratio H/L of around 13) were obtained after 2 and 3 passages and these duplication times were further used in SILAC-based AP-MS experiments described in section 3.4.8 and 3.4.9. In summary, all conditions tested successfully grew *L. tarentolae* in a complete defined medium. Condition 5 was selected and used in SILAC experiments because the protocol, slightly modified from Silverman *et. al* [147] (section 2.5.16), presented optimal growth kinetics in *L. tarentolae*.

a

Components	Conditions				
	1	2	3	4	5
M199	+	+	+	+	+
FBS	-	+	-	-	-
dFBS	+	-	-	-	+
Hypoxanathine	+	+	+	-	-
6-Biopterin	+	+	+	-	-
Biotin	+	+	+	-	-
Folic Acid	+	+	+	-	+
Adenosine	+	+	+	-	+
L-Glutamine	+	+	+	-	+
L-Proline	-	-	-	-	+
HEPES	+	+	+	+	+
Hemin	+	+	+	+	+

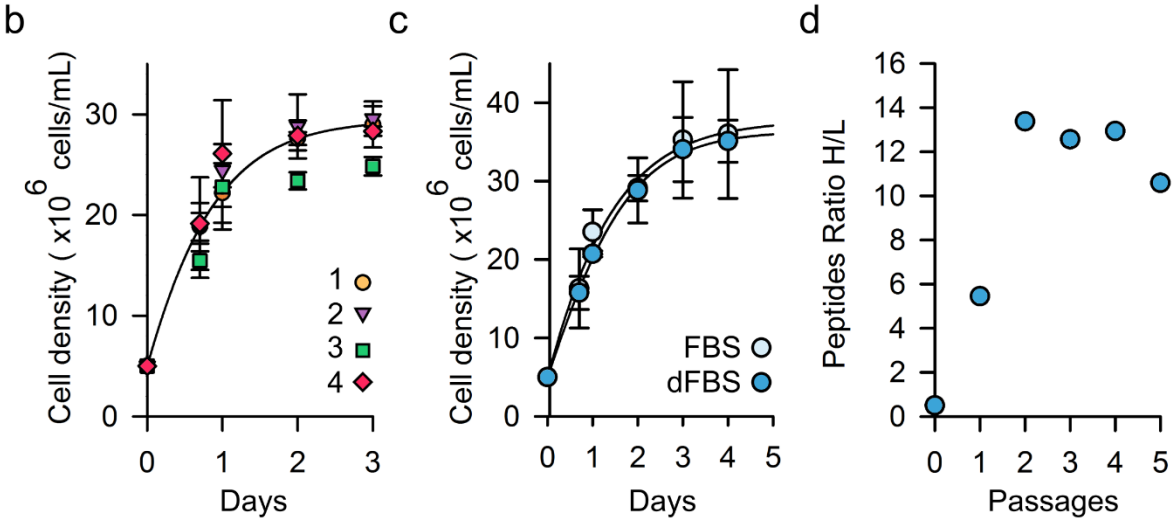
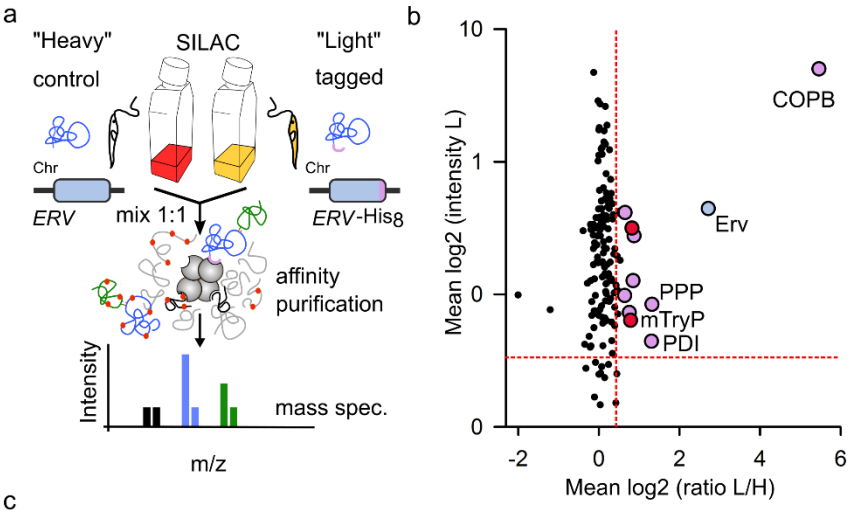


Figure 3.15 Defined medium formulation for SILAC experiments in *L. tarentolae*.
(a) Summary of conditions tested in defined medium formulation using M199 as base medium supplemented with non-dialyzed or dialyzed fetal bovine serum (FBS and dFBS respectively) and several medium components. **(b)** Growth curve depicting conditions 1-4. **(c)** Growth curve depicting condition 5 supplemented with FBS and dFBS. **(d)** Incorporation efficiency was calculated as ratios of all identified peptides from heavy versus light (H/L). Whole-cell lysates were prepared after splitting cells 1:20 every 72 h (one passage) and analysing the remaining sample by mass spectrometry. Each data point in growth curves (b and c) represents the mean ± standard deviation from two independent biological replicates. Mass spectrometry was performed by Frederik Sommer (AG Prof. M. Schroda, Kaiserslautern).

3.4.8 SILAC-based affinity purification coupled with mass spectrometry (AP-MS) studies on the Erv-His$_8$ interactome

In order to obtain an interaction profile of Erv we used SILAC combined with AP-MS. The strategy consisted of using an endogenous tagged Erv-His$_8$ line grown in light medium and a parental line with unedited Erv grown in heavy medium as control. Since we expected interactions to be covalent via disulphide bonds, a 1:1 mixing of bait and control cell populations was done before cell lysis, resulting in a denaturing affinity purification workflow (Figure 3.16a). This strategy sought to ensure standardization and to reduce technical variability. Baits were isolated by metal affinity chromatography using Ni-NTA and disulphide bonds were stabilized by alkylation with NEM in intact cells and during the purification process. Eluates were precipitated, resolved by SDS-PAGE followed by in-gel tryptic digestion and further analysed by mass spectrometry. Potential interactions were identified by analysing hits with high light-to-heavy ratios, which represented a theoretical enrichment compared to the control. In contrast, proteins that showed an abundant ratio near to the mixing ratio of 1 were considered co-purified contaminants in both conditions. The experiment was done in five biological replicates and more than 500 proteins were detected. The bait had one of the highest SILAC ratios (L/H 10) indicating the presence of an affinity enrichment of Erv-His$_8$ compared to the heavy control population. To select candidate hits, the data was plotted in a distribution of log2 fold changes intensity vs SILAC ratio on the premise that peptides should be present in at least 50% of the 5 replicates (Figure 3.16b). A summary of the potential true interactors and the predicted biological function are specified in Figure 3.16c. The most enriched protein was a beta prime COP protein. Furthermore, I detected a PDI protein and a tryparedoxin peroxidase. The enriched peptides did not present typical CX3C or CX9C-motifs. A systematic comparison with proteomic data was also conducted to find if the candidates are present in the Erv depletome and ATOM40 importome reported in *T. brucei* [160]. Additionally, the candidates were analysed *in silico* to predict an N-terminal mitochondrial targeting sequence by the algorithm MitoFates [161]. The comparison results are marked with symbols in the summary of the candidates (Figure 3.16c).



c

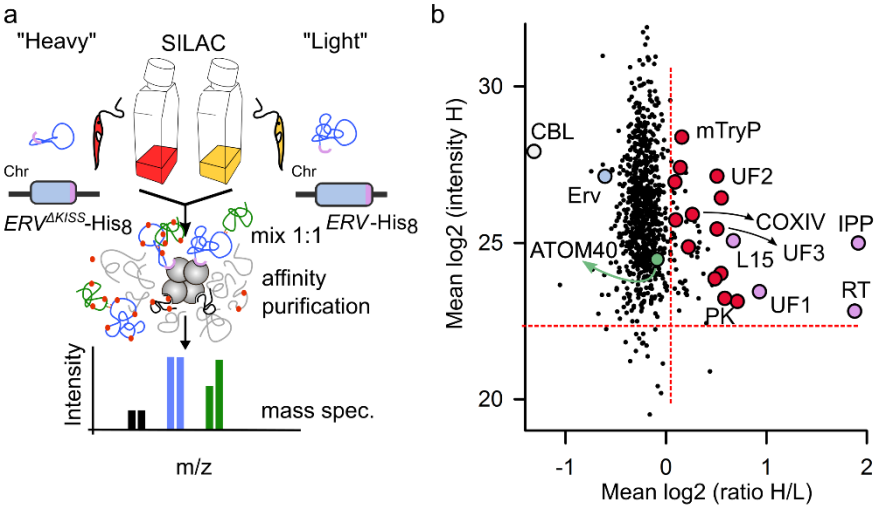
	Name	*L. tarentolae* accesion	Mean ratio /H
	Beta prime COP protein COPB	LtaP33.3460	50.47
△ ⬠	Erv	LtaP15.0980	9.56
	Putative phospho-protein phosphatase PPP	LtaP30.3320	2.56
	Protein disulfide isomerase PDI	LtaP36.7210	1.15
⬠ ●	Putative pteridine transporter	LtaP10.0370	1.96
	PUF RNA binding protein	LtaP33.1250	1.79
	Unknown function	LtaP29.2640	1.30
	Unknown function	LtaP31.1000	1.96
	Putative axonemal dynein	LtaP04.0680	1.75
△ □ ⬠ ●	Tryparedoxin peroxidase mTrypx	LtaP15.1080	1.37
	Putative exonuclease	LtaP36.1780	1.92

Figure 3.16 SILAC AP-MS for Erv.
(a) Schematic of the SILAC based affinity purification coupled with mass spectrometry (AP-MS) workflow used to detect potential Erv-His8 interactors. **(b)** Volcano plot showing putative interaction partners of Erv-His$_8$ under denaturing conditions. The peptides highlighted in purple within the right quadrant are candidates that overpassed a threshold for mean fold enrichment, intensity and were present in 50% of the replicates (n=5). **(c)** Summary of the candidates specifying the predicted gene function, accession gene number and fold mean fold enrichment. (●), Protein found in the *Trypnosoma brucei* Erv depletome [160], (⬠), protein found in the *T. brucei* ATOM40-dependent importome [160],(□) protein with a predicted N-terminal mitochondrial presequence using MitoFates [161],(△), protein found in the Erv$^{\Delta KISS}$-His$_8$ interactome determined in this study. Mass spectrometry data was performed by Frederik Sommer (AG Prof. M. Schroda, Kaiserslautern).

3.4.9 SILAC-based affinity purification coupled with mass spectrometry (AP-MS) studies on the Erv$^{\Delta KISS}$-His$_8$ interactome

To test putative interactions detected in Erv$^{\Delta KISS}$-His$_8$ we again performed SILAC combined with AP-MS. To reveal differences between mutant and wt Erv, Erv$^{\Delta KISS}$-His$_8$ was grown in heavy medium and Erv-His$_8$ in light medium as control. The workflow and affinity purification conditions were as described in the earlier section (Figure 3.17a). Eluates were concentrated, reduced, alkylated, and digested with trypsin using an ultrafiltration cartridge before further detection by mass spectrometry. Potential interactions were identified by analysing hits with high heavy-to-light ratios, which represented substrates that were enriched in the Erv$^{\Delta KISS}$-His$_8$ bait. In contrast, proteins that were present in both baits Erv$^{\Delta KISS}$-His$_8$ and Erv-His$_8$ showed a common abundant ratio (close to 0). The experiment was done in biological duplicates and

880 proteins were detected. To visualize and detect potential hits, the data was plotted in a distribution of log2 fold changes intensity vs SILAC ratio (Figure 3.17b). A summary of the potential true interactors and the predicted biological function are specified (Figure 3.17c).



c

Name	*L. tarentolae* accesion	Mean ratio H/L
Inositol polyphosphate phosphatase IPP	LtaP24.0900	3.73
Riboflavin transporter RT	LtaP05.0490	3.65
Unknown function 1 UF1	LtaP34.4370	1.88
Pantothenate kinase PK	LtaP28.0140	1.62
Ribosomal protein L15	LtaP35.0070	1.59
Ligand-effect modulator 3	LtaP09.1060	1.49
Fatty acyl CoA synthetase	LtaP03.0220	1.48
Folate transporter	LtaP10.0340	1.47
Unknown function 3 UF3	LtaP07.0980	1.43
Unknown function 2 UF2	LtaP32.0380	1.42
Mitotubule-associated protein	LtaP26.2080	1.40
Cytochrome c oxidase subunit IV	LtaP12.0690	1.21
Mitochondrial ribosomal protein L15	LtaP16.0020	1.20
Tryparedoxin peroxidase mTrypx	LtaP15.1080	1.11
Unknown function	LtaP33.2660	1.09
Unknown function	LtaP24.0150	1.08
Unknown function	LtaP14.1510	1.04
Import receptor subunit ATOM40	LtaP35.4850	0.93
Erv	LtaP15.0980	0.66
Cystathionine beta-lyase CBL	LtaP14.0440	0.39

Figure 3.17 SILAC AP-MS for Erv$^{\Delta KISS}$.
(a) Schematic of the SILAC based affinity purification coupled with mass spectrometry (AP-MS) workflow used to detect differences between Erv$^{\Delta KISS}$ and Erv. **(b)** Volcano plot showing proteins enriched in Erv$^{\Delta KISS}$. The peptides highlighted in the right quadrant are candidates that overpassed a threshold for mean fold enrichment and intensity. **(c)** Summary of the candidates specifying the predicted gene function, accession gene number and fold mean fold enrichment. (●), Protein found in the *Trypnosoma brucei* Erv depletome [160], (⬟), protein found in the *T. brucei* ATOM40-dependent importome [160],(□) protein with a predicted N-terminal mitochondrial presequence using MitoFates [161],(△), protein found in the Erv-His$_8$ interactome determined in this study. Mass spectrometry was performed by Proteome Sciences R&D (Frankfurt am Main, Germany).

The most enriched proteins were an inositol polyphosphate phosphatase (IPP) and a riboflavin transporter (RT). Furthermore, I detected a protein with unknown function (UF2) with a potential redox pair CHAYCRCSYC-motif located at the N-terminal that is conserved in Kinetoplastida. It is worth noting the detection of a tryparedoxin peroxidase also found in the

Erv-His_8 interactome in this work and that unknown function 3 (UF3) was the top of the list (the most depleted protein) in the Erv depletome of *T. brucei* [160]. However, the latter candidate has a predicted molecular mass of 240 kDa. Similar to the results using Erv-His_8 as a bait, I could not detect peptides with typical CX3C or CX9C-motifs. A systematic search comparing the candidates with existing proteomic data was also conducted as previously mention for the Erv-His_8 interactome. An overlap of the candidates was analysed to find if they are present in the Erv depletome and ATOM40 importome reported of *T. brucei* [160]. Furthermore, the candidates were analysed *in silico* to predict an N-terminal mitochondrial targeting sequence by the algorithm MitoFates [161]. The comparison results are marked with symbols in the summary of the candidates (Figure 3.17c).

3.4.10 High yield and purity of recombinant $Erv^{\Delta KISS}$-His_6 from *E. coli* using auto-induction medium

Erv in *L. tarentolae* and homologues in Kinetoplastida were extensively studied *in vitro* [72, 73]. However, the crystal structure was not yet described for members of this order. Having shown here that $Erv^{\Delta KISS}$-His_8 is functional *in vivo*, we speculated that the lack of the KISS domain might also be beneficial in protein purification and crystallization trials (Figure 3.18a). Therefore, we streamlined a previously established protocol [72], this time using $Erv^{\Delta KISS}$ and auto-induction medium with the aim to conduct crystallization trials. The combined strategy using a short-length mutant together with auto-induction medium resulted in high yield and purity (Figure 3.18b).

The oligomeric state of $Erv^{\Delta KISS}$-His_6 was also analysed *in vitro* by gel filtration chromatography under both reducing and non-reducing conditions (Figure 3.18c) where a single peak corresponding to a molecular mass of 38 kDa and two major peaks corresponding to a molecular mass of 38 and 89 kDa respectively are shown. Indicating the presence of a homodimer and possible a homotetramer. Selected peak fractions were further analysed by SDS-PAGE under non-reducing condition, revealing, bands corresponding to the 19 kDa and 38 kDa monomeric and homodimeric size, but also a band around the 57 kDa size, presumably matching a homotrimer (Figure 3.18d).

In summary, the purification of $Erv^{\Delta KISS}$-His_6 using auto-induction medium offered a great quality recombinant protein suitable for crystallization tests. Currently, protein crystallization

screening trials are conducted by Dr. Thomas Barends (Max Planck Institute, Heidelberg). Furthermore, the absence of the KISS domain apparently altered the oligomer state comparted to the full-length *Lt*Erv [72] and *Sc*Erv [86] in a mechanism that remains to be elucidated.

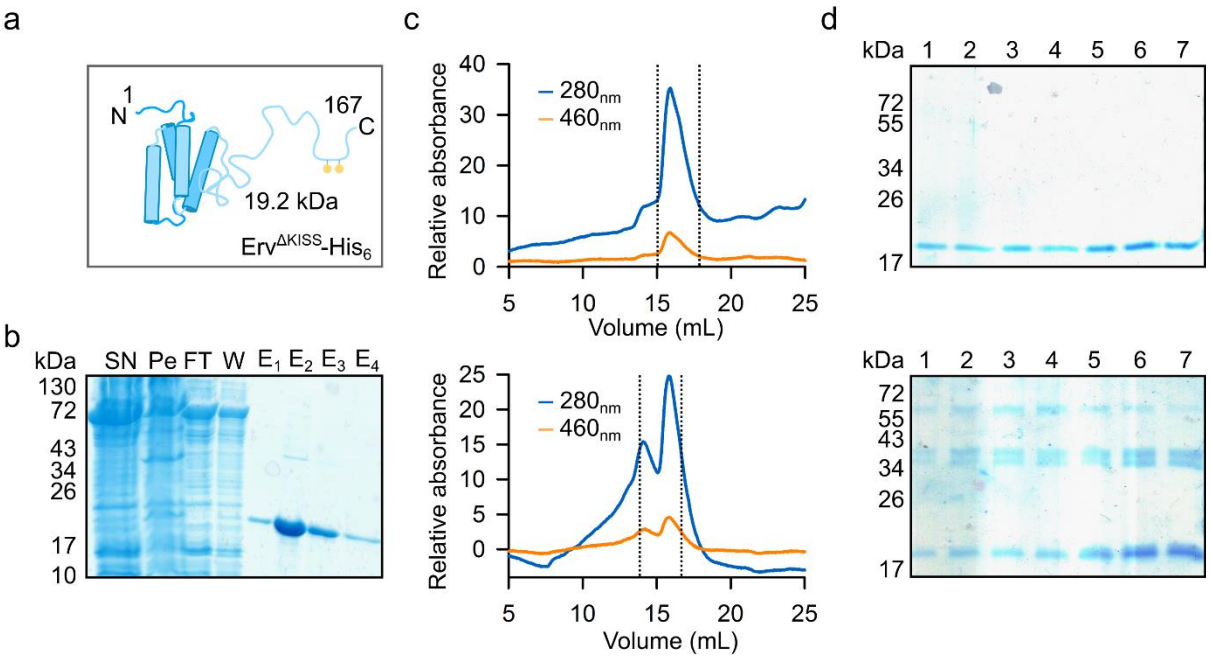


Figure 3.18 Heterologous expression of Erv$^{\Delta KISS}$-His$_8$.
(a) Schematic of mutant cloned in pET-28a with a C-terminal His$_6$-tag and an expected molecular mass of 19.2 kDa. **(b)** SDS-PAGE (12% w/v) of a representative purification of Erv$^{\Delta KISS}$-His$_6$ in *E. coli* using Ni-NTA chromatography. SN, supernatant fraction; Pe, Pellet or insoluble fraction; FT, flow through; W, wash, E, eluate. **(c)** Eluates were separated by gel filtration chromatography under reducing (top) and non-reducing (bottom) conditions. Dashed lines indicate fractions analysed. For reducing conditions 2 mM DTT in the running buffer was used. Protein content and the presence of flavin cofactor was analysed by measuring the absorbance at 280 nm and 460 nm respectively. **(d)** Fractions obtained from gel filtration chromatography were re-loaded and analysed by SDS-PAGE (12% w/v) under reducing (top) and non-reducing (bottom) conditions and stained with colloidal Coomassie blue. The representative expression, purification and SDS-PAGE was carried out under my supervision by Ogaga Daniel Ogbodu.

3.5 Investigations using the *glmS* riboswitch in *L. tarentolae*

3.5.1 A *glmS* riboswitch-induced partial knockdown regulation in episomal reporter assays

In order to test the *glmS* riboswitch function in *L. tarentolae*, a plasmid-based mCherry-*glmS* reporter was designed (Figure 3.19a). The expected knockdown effect was quantified by measuring fluorescence intensity in reporter transfectant lines grown in medium with and without glucosamine (GlcN). A mutant *glmS*^{M9} with catalytically inactive RNA was used as a control. The riboswitch reporter lines *glmS* and the control strain were obtained by transfecting parasites with pX-mCherry-*glmS*^{wt} and pX-mCherry-*glmS*^{M9} respectively. A first experiment where parasites were passaged with fresh medium every 72 h was conducted in MEM medium (Figure 3.19b). The effect of glucosamine on growth during the experiment was also evaluated (Figure 3.19c). After treatment, the fluorescence was reduced by ~35% compared to control lines in the absence of glucosamine, indicating a possible endogenous effect (Figure 3.19d). A reduction in fluorescence of ~75% was also detected when glucosamine was added exogenously, showing that the knockdown effect could also be induced. The behaviour of the *glmS* reporter lines was also further tested in BHI medium (Figure 3.20). There, an endogenous effect was observed whereas an exogenously GlcN-mediated knockdown regulation was not evident. The reporters consisted of cells transfected with a plasmid bearing the mCherry gene fused to either the *glmS*^{wt} or *glms*^{M9} sequence at 3' end. In summary, the experiments suggest that there is an inducible and endogenous knockdown effect from an episome. However, the extension of the knockdown effect indicate that the regulation might be partially effective.

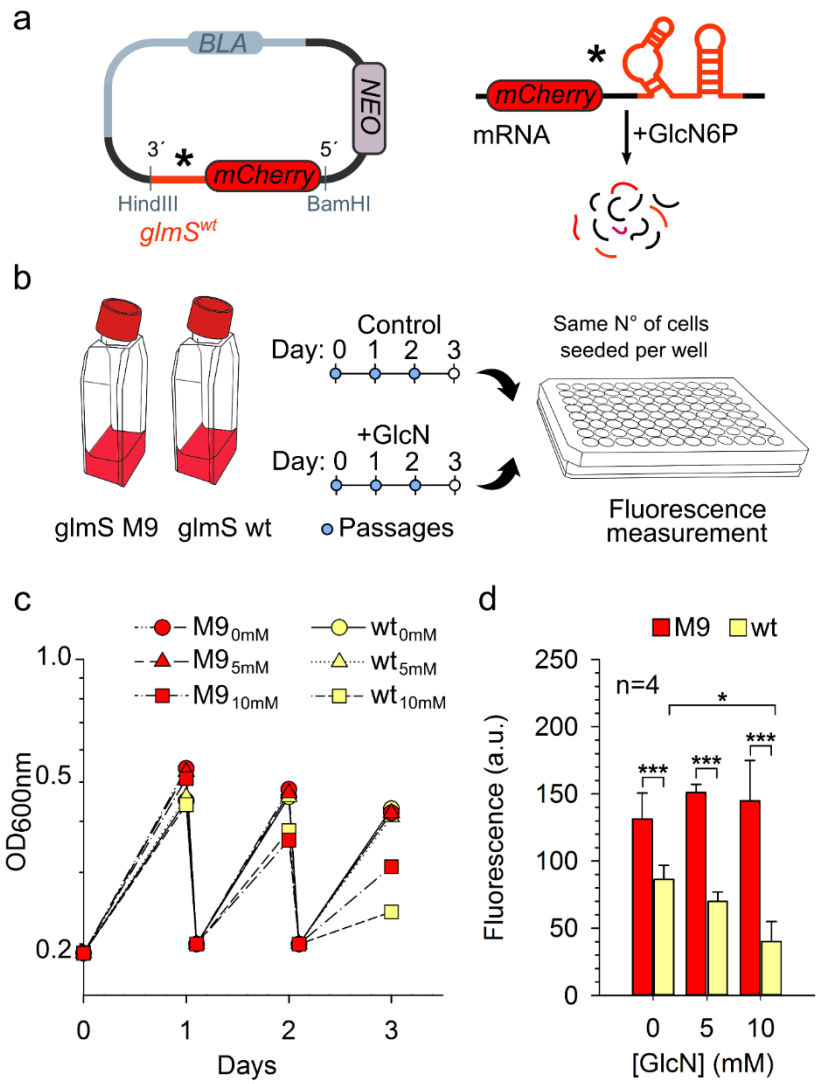


Figure 3.19 Episomal knockdown assay in mCherry reporter lines using MEM medium.
(a) Left: Schematic of reporter episome containing the mCherry fluorescent coding sequence fused to the wt or inactive M9 (*) *glmS* riboswitch. Right: expected self-cleavage and degradation of mCherry-transcript fused to wt riboswitch upon the addition of GlcN. **(b)** Experimental design to treat reporter-transfectants with GlcN. Parasites were exposed to GlcN for 2 passages and subsequently harvested. The knockdown effect was analysed by fluorescence normalizing the amount of cells by seeding the same number of parasites. Untreated cells and an inactive riboswitch M9 served as controls. **(c)** Representative growth analysis during the treatments. **(d)** Fluorescence obtained for treated and untreated parasites using 5 and 10 mM GlcN. (wt) active and (M9) inactive riboswitch. Each data point represents the mean of four biological replicates. Statistical analyses were performed using the one-way ANOVA method (* $p < 0.05$; *** $p < 0.001$). Modified from Turra *et al.* [152].

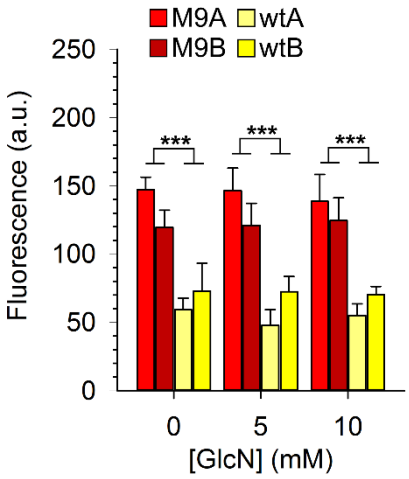


Figure 3.20 Episomal knockdown assay in mCherry reporter lines using BHI medium.
Two clonal mCherry reporter lines, including wild-type riboswitch and control (wtA, B and M9A, B respectively) were treated with 5 and 10 mM GlcN in BHI medium. Each data point represents the mean of four biological replicates. Statistical analyses were performed using the one-way ANOVA method (*** $p < 0.001$). Modified from Turra *et al.* [152].

3.5.2 The *glmS* riboswitch does not induce the expected knockdown phenotype in endogenous *APRT*, *PF16* and *ERV* lines.

In order to validate the *glmS* results obtained with reporter lines, the knockdown effect was analysed this time at the chromosomal level. CRISPR/Cas9 was used to target *APRT*, *PF16*, and *ERV* chromosomal loci with $glmS^{wt}$ and $glmS^{M9}$ sequence at the 3' terminus (Figure 3.21a, 3.22a, 3.23a). The selection of these genes to study their corresponding knockdown phenotypes was based on the knowledge acquired previously in this work in regards to the knockout phenotypes (see section 3.1.2, 3.2, and 3.4.4). An increase in APP drug resistance is expected when the *APRT* gene is knocked down. In the case of *PF16,* an impairment or reduction of movement was speculated. Regarding the essential *ERV* gene, we predicted that a depression in Erv synthesis will lead to a defect in parasite growth. To generate endogenous *glmS* mutants, specific sgRNAs were designed to target immediately downstream of the genes of interest. Donor DNAs containing $glmS^{wt}$ and $glmS^{M9}$ were generated with vectors pMOTag using two antibiotic resistance cassettes to ensure mutant homozygosis. After DNA repair genes were edited at the 3' terminus with riboswitch sequences $glmS^{wt}$ and $glmS^{M9}$. Parental lines expressing Cas9 and T7 were transfected and corresponding clonal lines for *APRT, PF16* and *ERV* were further characterized. PCR analysis showed the expected editions for *APRT* (Figure 3.21b), *PF16* (Figure 3.22b), and *ERV* (Figure 3.23b), visualizing the absence of an unedited locus. To further validate the introduction of the mutation, PCR amplicons spanning

the junction between the chromosome and the *glmS* sequence were sequenced, showing the correct modification for all genes (Figure 3.21c, Figure 3.22c,Figure 3.23c). To investigate the efficiency of the putative knockdown-mediated effect in *APRT*, a similar experiment as for Δ*APRT* was carried out for *APRT*::*glmS* lines using MEM medium and exposing the cells to glucosamine (Figure 3.21d). Lines treated with APP and supplemented with or without glucosamine did not show a significant difference in parasitemia compared to controls, indicating that under the conditions tested the expected knockdown regulation of *APRT* was not evident. Likewise, *PF16*::*glmS* lines were exposed to glucosamine in MEM medium and inspected under the microscope. Visual phenotypic evaluation in double-blind experiments did not show an apparent difference in motility when compared to the control lines.

A similar analysis assessing the impairment in growth was carried out by Luzia Schneider for *ERV*::*glmS* lines by exposing cells to glucosamine in MEM medium. Cell count and optical densities of cultures treated or not with glucosamine did not show a significant difference (Figure 3.23e,d).

Thus, using three independent genes tagged with *glmS* at the chromosomal level, I could not gather evidence for the expected knockdown phenotypes. The unsuccessful knockdown effect suggests that the *glmS* system is not efficient in *Leishmania* or that modifications are required to achieve the desired function of this technique.

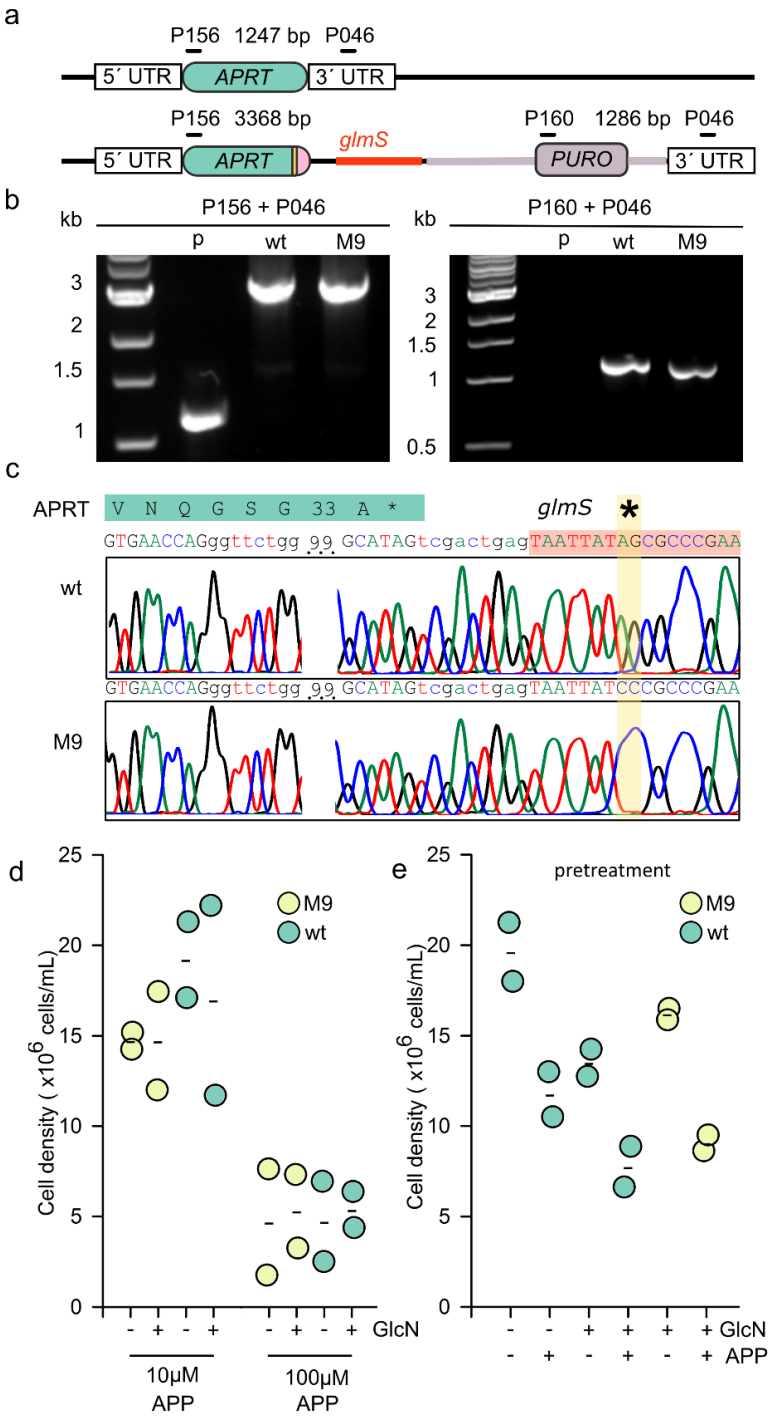


Figure 3.21 Endogenous *glmS* riboswitch-tagging of the *APRT* gene and phenotype analysis.
(a) Scheme of the *APRT* locus and expected knockin edition using one antibiotic resistance cassette. Black lines indicate primer binding sites aligned with expected PCR amplicon size. **(b)** Analytical PCR showing the genotypes for lines tagged with *glmS*^{wt} and control *glmS*^{M9}. Amplification of the *APRT* locus revealed the targeting of both copies (homozygote). The expected integration of the cassette was shown by the presence of recombinant PCR products. **(c)** The successful fusions of the *glmS* sequences were corroborated by Sanger sequencing of a PCR amplicon showing the junction between chromosomal *APRT* and the *glmS* riboswitch. **(d)** *APRT* phenotype of lines fused to *glmS*^{wt} and inactive *glmS*^{M9} by growth analysis in supplemented MEM medium. Left: Parasites were exposed to 10 or 100 µM APP with or without addition of 10 mM glucosamine (GlcN). Right: A pre-treatment was carried out where cells were exposed to 5 mM GlcN for 2 passages before exposure to 10 µM APP. A pre-treatment without GlcN in a line fused to *glmS*^{wt} was used as a control. Data points represent the mean of two biological replicates. Modified from Turra *et al.* [152].

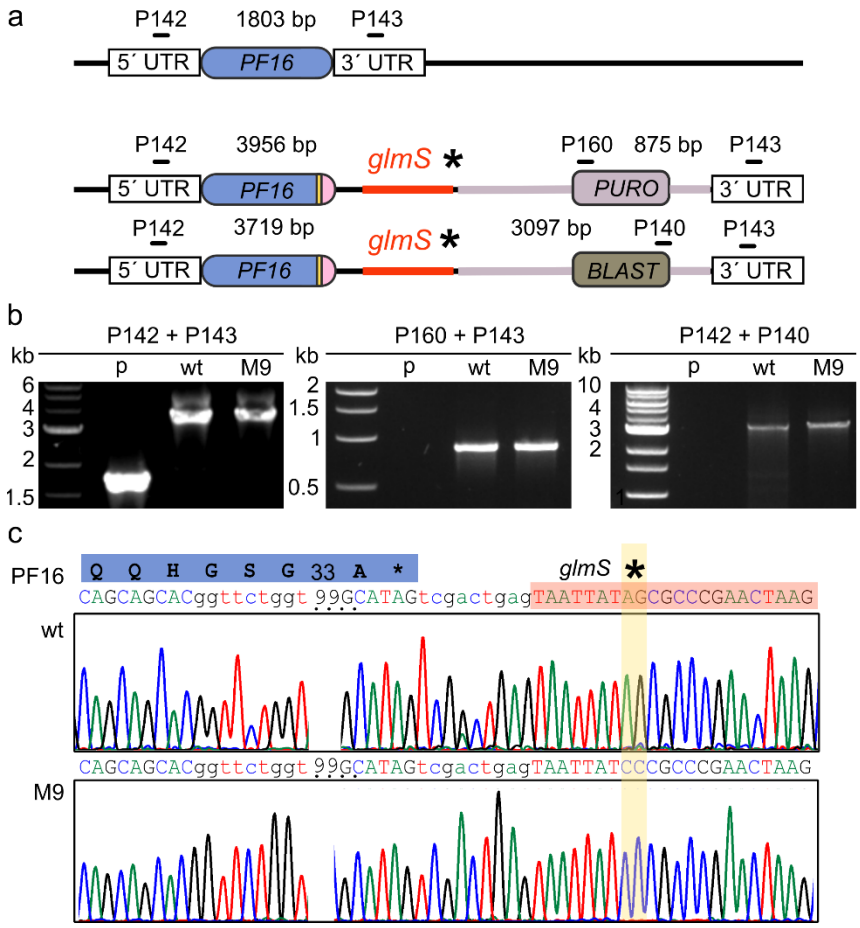


Figure 3.22 Endogenous *glmS* riboswitch-tagging of the *PF16* gene.
(a) Scheme of the *PF16* locus and expected knockin edition using two antibiotic resistance cassettes. Black lines indicate primer binding sites aligned with expected PCR amplicon size. **(b)** Analytical PCR showing the genotypes for lines tagged with *glmS*^{wt} and control *glmS*^{M9}. Amplification of the *PF16* locus revealed the absence of a wild-type amplicon (homozygote). The expected integration of both cassettes was shown by the presence of recombinant PCR products. **(c)** The successful fusion with *glmS*^{wt} and control *glmS*^{M9} was corroborated by Sanger sequencing of a PCR amplicon showing the junction between chromosomal *PF16* and *glmS* riboswitch. Modified from Turra *et al.* [152].

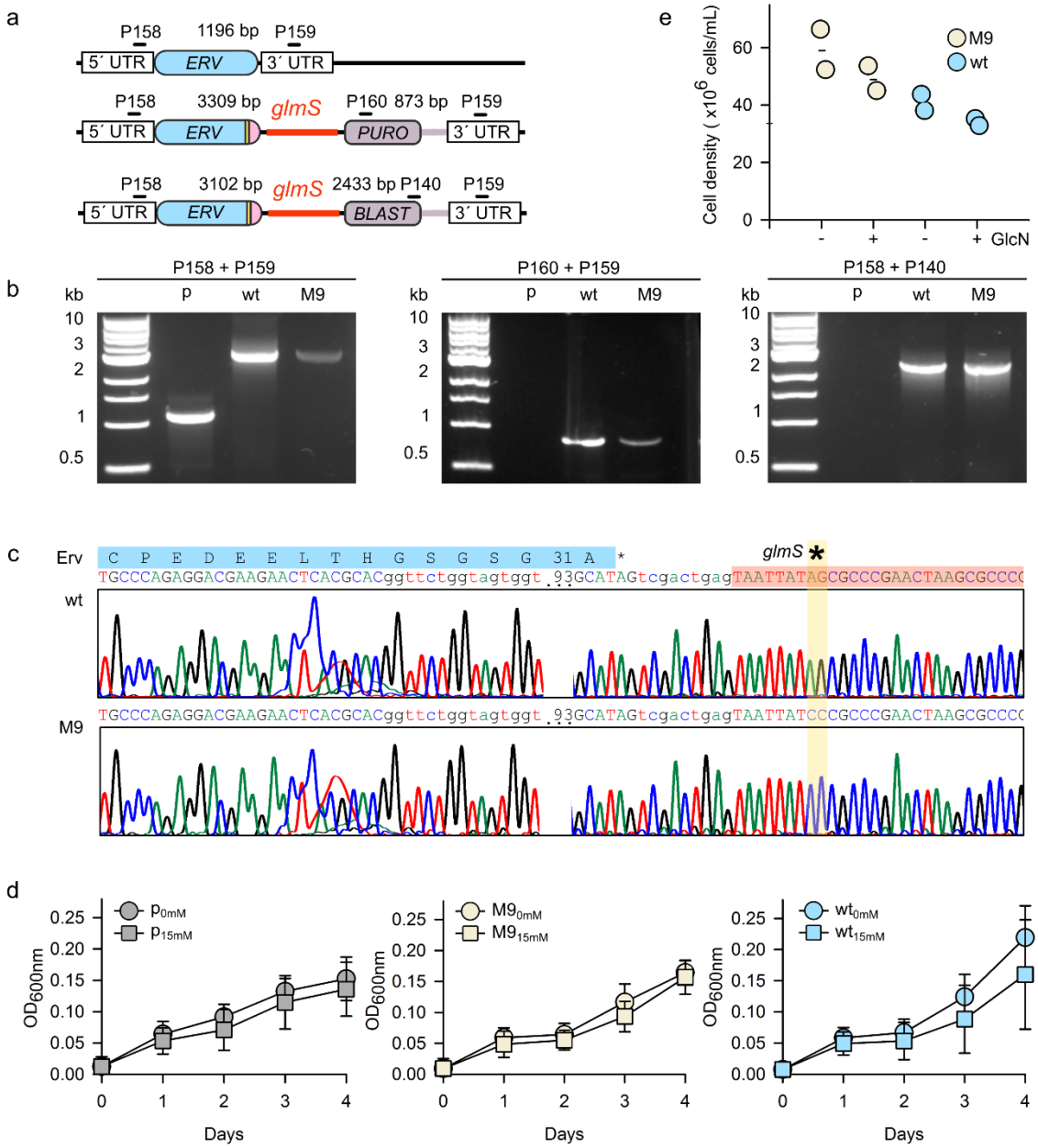


Figure 3.23 Endogenous *glmS* riboswitch-tagging of the *ERV* gene and phenotype analysis.
(a) Scheme of the *ERV* locus and expected knockin edition using two antibiotic resistance cassettes. Black lines indicate primer binding sites aligned with expected PCR amplicon size. **(b)** Analytical PCR showing the genotypes for lines tagged with *glmS*^{wt} and control *glmS*^{M9}. Amplification of the *ERV* locus revealed the targeting of both copies (homozygote). The expected integration of the cassette was shown by the presence of recombinant PCR products. **(c)** The successful fusion of the *glmS*^{wt} and control *glmS*^{M9} was corroborated by Sanger sequencing of a PCR amplicon showing the junction between endogenous *ERV* and the *glmS* riboswitch. **(d)** and **(e)** Growth analysis of lines fused to *glmS*^{wt} and control *glmS*^{M9} riboswitch in supplemented MEM medium. **(d)** Growth curve assessing parasitemia by optical density in cultures exposed to 15 mM glucosamine (GlcN). (p) parental TB007-transfectant line served as an additional control. Data points represent the mean of three independent biological replicates. **(e)** Erv *glmS*^{wt} and control *glmS*^{M9} were treated with and without 15 mM GlcN and cell density was measured by cell counting. Data from two independent biological replicates are shown. Line generation of endogenous *ERV* fused to *glmS*^{wt} and control *glmS*^{M9}, sequencing and measurement of parasitemia was performed by Luzia Schneider. Modified from Turra *et al.*[152].

CHAPTER

DISCUSSION 4

4.1 Genetic and proteomic studies on *Leishmania tarentolae* Erv

Protein disulphide bond formation within cellular compartments is usually catalysed by two essential components. An oxidoreductase that oxidizes substrates inducing folding and a second component that re-oxidizes the first adapter [68, 81, 162]. This common mechanistic feature exists in prokaryotes and eukaryotes. Examples of the latter include, the endoplasmic reticulum, the thylakoid lumen of chloroplasts, and the mitochondrial intermembrane space (IMS) [57, 163-165]. This two-components paradigm certainly influenced how we try to understand other models. For example, the incomplete MIA pathway in the IMS of *Plasmodium* and Kinetoplastida, where only Erv was found, and Mia40 is not present [20, 70, 72, 73]. Thus, it was reasonable to hypothesize, that a missing component awaiting to be described, replaces Mia40 and fulfils its function. The alternative hypothesis proposes that Erv can oxidize substrates alone. However, this would need the description of an unusual and not understood mechanism that appears to be absent in Erv-homologues. Either way, whether a Mia40 homolog exist or not, Kinetoplastida Erv presents unique properties. This protein differs in its protein structure and has unknown structure-function relationships [72, 73, 131]. Our group took the kinetoplastid parasite *L. tarentolae* as a model [20] to gain, for example, an understanding of *L. tarentolae* Erv (*Lt*Erv) function. Understanding how LtErv works would lead to a still missing oxidative folding model in the IMS of Kinetoplastida, supporting or dismissing the presence of a Mia40 adapter. Studies on yeast provided the first hints that underpin the function of *Lt*Erv domains and cysteines [131]. Our group also searched extensively for a putative Mia40 homolog without finding interactions between Erv and other proteins [166].

In this work, I used the CRISPR/Cas9 system to give examples that dissect the structure and function of *Lt*Erv at the chromosomal level. The main conclusions from the genetic experiments in this fashion are simple to interpret. The main premise is that *Lt*Erv is essential for viability. Therefore, any generation of a mutant *LtERV* that did not lead to cell death, provided experimental evidence that the essential function was still conserved. This strategy was used to give answers to the following questions: what is the function of Cysteine 17? Are the intervening sequences in the redox pair important for viability? What is the function of the KISS domain? Is the import of Erv dependent on its C-terminal KISS domain? Can we detect a specific mixed disulphide bond between Erv and other cysteine-containing proteins?

4.1.1 Erv mutants and structure-function studies in *L. tarentolae*

A multiple sequence alignment between Erv-homologues evidenced that a long C-terminal domain is only present in Kinetoplastida [72]. This sequence encodes for almost 50% of the polypeptide chain that separates the FAD-binding domain from the C-terminal redox pair. The domain has no apparent homologues in other eukaryotes, is predicted to be unstructured, and its function is still unknown [72, 73]. Considering that it is present in a model where no Mia40 was found, it has been suggested that the KISS domain could exert Mia40 functions. This long unstructured domain lacks cysteine residues but could mimic the hydrophobic cleft found in Mia40 that binds substrates [73, 131]. In Mia40, this domain forms a substrate-binding pocket that recruits unfolded substrates via hydrophobic interactions mediating the interaction with the shuttle disulphide [167, 168]. Recently, it was also proposed that the KISS domain might be an adaptation for binding another oxidoreductase that replaces Mia40 [78]. Analysis of the protein sequence of the KISS domain shows no obvious hydrophobic sequence. However, experimental evidence suggests that *Lt*Erv seems to associate with the mitochondrial membrane in *Leishmania* and yeast [20, 72, 131]. The relationship of *Lt*Erv with the outer or inner membrane of the mitochondria could be related to its function. The import of *Lt*Erv itself is not yet understood and it is not known if this mechanism is interconnected with its membrane association. To propose that the KISS domain (i) plays a role as a substrate-binding cleft, (ii) that interacts with the Mia40-replacement, or (iii) mediates the import of *Lt*Erv through membrane association, leads to the hypothesis that all these functions should be essential. Therefore, to address the latter hypothesis and to understand the relevance of

the KISS domain *in vivo,* I constructed *LtERV*$^{\Delta KISS}$. This consisted of an endogenous chimera lacking a great portion of the C-terminal domain but conserving the shuttle redox pair. The successful generation of viable mutants dismisses an essential function of the KISS domain in kinetoplastid parasites. This is in line with the results described by Specht *et. al* where fusing the KISS domain to *Sc*Erv did not affect its function by a negative dominant effect [131]. The hypothesis that the KISS domain could be the adapter for substrate binding and therefore replace Mia40 in Kinetoplastida was also dismissed in the yeast model. In this experiment, a targeted *Lt*Erv (containing a transmembrane segment of ScMia40) could not replace *Sc*Mia40. The hypothesis that the KISS domain could serve as a substrate-binding site, was also not supported, in previous experiments by Peleh *et al.* [169]. In the latter study, cross-species experiments in yeast revealed that *Arabidopsis thaliana* Erv (*At*Erv) (which lacks the KISS domain) could mediate the import and oxidative folding of some essential substrates without the redox assistance of *Sc*Mia40. Of note, *Plasmodium falciparum* also lacks the KISS domain and no Mia40 has been found [70, 72, 170]. Although I could not detect any growth impairment in mutants lacking the KISS domain and I dismiss the KISS structure-function relationship hypothesis explained earlier, I do not propose that it lacks function. Or more specifically, that the KISS domain might be the result of a neutral genetic drift, where no specific function was selected through evolution. Although not essential for survival in laboratory conditions, little evidence was given in this work as to how this mutant behaves in the whole parasite life cycle, including the vector and its host. The growth curve analysis presented in this study could also be extended to include new media formulations. For example, I also tested a defined medium without glucose (data not shown). In these conditions, parasites are forced to rely on mitochondrial oxidative phosphorylation to survive as the main source of ATP, since they cannot obtain it by substrate-level phosphorylation [171]. This preliminary experiment failed, probably due to incorrect media formulation for *L. tarentolae*. Nevertheless, this principle is used to unmask previous unnoticed phenotypes for protein mutants involved in mitochondria protein import in *Trypanosomes* [172] and yeast [58, 79, 131]. The correct defined formulation in a medium lacking glucose would be of utility to underpin mitochondrial-related functions in *Lt*Erv substrates and *Lt*Erv mutants. The hypothesis that the KISS domain is playing a role in membrane association could also be addressed in future studies by carbonate or digitonin extraction using the line Erv$^{\Delta KISS}$

generated in this work. The most remarkable difference between the Erv$^{\Delta KISS}$ and full-length Erv in this work was the presence of higher molecular disulphide complexes as revealed by western blots. This was rather unusual and never found for wild-type Erv. The disulphide nature of these interactions was corroborated by treating samples with a reducing agent revealing the expected mutant monomer size. Nonetheless, it is difficult to provide evidence of the identity of the protein/s involved. The observed bands in western blots could be the result of Erv itself or the involvement of other proteins. The former could be explained as a result of a complete alteration of the oligomer state of Erv due to the deletion of the KISS domain. The latter could be indeed the long-awaited substrate/s interacting with *Lt*Erv. These two explanations are not necessarily excluding each other. Nevertheless, the formation of upper bands in western blots provided evidence that the deletion of the KISS domain altered the protein kinetics in a process not yet understood. Much remains to be studied to define which step or steps are involved. The flexible KISS domain could play a role in substrate oxidation or/and homodimerization. This would mean that the KISS domain could recruit substrates to effectively oxidize them. Alternatively, the KISS domain could be needed to stabilize the Erv dimer when the monomer is not interacting with the other subunit, during electron transfer to the flavin. To point specifically which step or steps are altered could provide new insights on Erv mechanisms of catalysis or to further support its actual mechanism [72]. An altered disulphide pattern was also observed in Erv$^{\Delta KISS}$ when expressed heterologously in *E. coli.* Although this is not evidence that this also occurs *in vivo*, the result could suggest that the upper bands observed in western blots are complexes of *Lt*Erv itself. We should also note that the molecular mass of the mutant is around 20 kDa and the main band detected in western blots matches the dimmer. The range of the monomer size also includes many interesting candidates (e.g Mic20). This creates difficulty in non-reducing western blot analysis. It is not possible to distinguish between an Erv homodimer and a mixed disulphide between Erv and a candidate Mia40 homolog around 20 kDa. For this, we would require specific antibodies against the putative candidate.

Complementation cross-species studies in yeast showed that *Lt*Erv was not able to replace *Sc*Mia40 and that only a mutant *Lt*Erv could replace *Sc*Erv [131]. This latter mutant was *Lt*Erv^{C17S}, where the cysteine 17 was eliciting a dominant-negative effect in complementation experiments. The main study's hypothesis was to find which property in *Lt*Erv structure was

responsible for the failure in the replacement of *Sc*Erv as previously reported by Eckers et. al [72]. Considering these complementation experiments, I decided to ask if this conserved cysteine residue has a physiological function and if it is essential for *L. tarentolae* survival. It was proposed that the structural function of this cysteine residue located at the N-termini of *Lt*Erv could be related to the equivalent residue in Erv present in mammals (ARL). In this homologue, the cysteine residue serves as a clamp disulphide bond locking the Erv polypeptides to form a homodimer [57, 72, 82, 168, 173]. Although the absence of this residue was essential for complementation studies in yeast, I found in the current study that a cysteine 17 mutant can be generated. This would remove its contribution to the essential function of *Lt*Erv and suggest a rather specific structural artificial relationship between *Lt*Erv and *Sc*Mia40 in the yeast model. This is further supported by the experiment where *Lt*Erv^{C17S} could also not complement the loss of Mia40 [131]. This could suggest that the reason for *Lt*Erv^{C17S} ability to replace *Sc*Erv was not an essential function of cysteine 17, but rather an incompatibility with *Sc*Mia40. This cysteine could have been oxidized by *Sc*Mia40 and poisoned its catalysis. A similar observation was done by Peleh *et al*. for *At*Erv where a negative dominant effect was also observed, as *At*Erv appeared to arrest the normal *Sc*Mia40 function while it was not able to oxidize it [169]. This type of analysis in combination with the mutant phenotypes generated by CRISPR/Cas9 in *L. tarentola*e could highlight a word of caution regarding chimeric/cross-species complementation assays. Although these preliminary studies could give us important thermodynamic, structural, and evolutionary lessons [79, 169], they should not be considered to be complete. The ultimate experiment that attempts to dissect structure from function should be carried out in the specific organism we aim to understand. Again, the generation of these mutants does not imply a lack of structural function, rather points out that the essential function/s features of *Lt*Erv in Kinetoplastida could dispense them. A possible explanation is that an ancestor of *Lt*Erv used this cysteine to form a disulphide bond with the other subunit to stabilize a homodimer (like in *Rn*Erv) (Figure 1.6). However, this function might have been lost when the equivalent cysteine from ARL in *Lt*Erv was mutated at its C-terminal end.

The biogenesis of *Sc*Erv has been characterized in yeast. It is known that *Sc*Erv is dependent on *Sc*Mia40 for its correct sorting into the IMS [174]. How *Lt*Erv is imported into the mitochondria in kinetoplastid parasites is still an unrevealed mechanism. However, the

studies of this work suggest that the import of *Lt*Erv into the IMS could be independent of the KISS domain or the conserved clamp cysteine 17.

We have looked into Erv mutant phenotypes generated by CRISPR considering that this protein is essential (section 3.4.4). Therefore, if the gene or a sequence is refractory to intervention, it is negative evidence that the element is essential. Before the wide use of CRISPR/Cas9 system, conclusions derived from negative results were viewed with askance [28]. This is because the genetic experiment cannot be separated from the transfection. Although today improved, transfections can be a source of technical failure [28, 117]. Another source of misinterpretation can be encountered when an essential gene is targeted. Because of aneuploidy in *Leishmania,* both chromosomes could be targeted but cells could eventually alter the karyotype to survive, generating extra copies of the essential gene [28, 95]. For this reason, it is important to show the absence of extra gene copies in the genotyping analysis of knockout mutants. Nowadays, CRISPR/Cas9 induces a high frequency of recombination and sgRNAs can be tested in independent experiments. Therefore, the impossibility to target a gene can provide evidence of the gene's essentiality. For example, the technique known as genome-wide CRISPR/Cas9 knockout screens has been used to classify genes as indispensable or essential based on the principle of negative results [175, 176]. In the case of *ERV*, I attempted to knockout the gene, in two independent experiments. The failure to obtain viable cells supports that it is essential as it was previously reported for *Leishmania infantum* using classic genetics [131]. Of note, the sgRNA used to target downstream from the coding sequence of *ERV* was already tested and used satisfactorily to fuse His-tag, Strep-Tag, $glmS^{wt}$ and $glmS^{M9}$. This dismisses the explanation that Cas9 was not active during the knockout experiment. Moreover, attempts to fuse mCherry to the N- and C-terminal domains of Erv also failed to produce viable cells. Similar results were obtained when I tried to fuse Erv to a 3xMyc-tag. This highlights, as in many studies ignored, that tags can impair protein function [177]. The use of CRISPR/Cas9 to generate endogenous mutants also provides advantages related to this matter. When working with an essential protein, any modification that survives guarantees correct localization and function. This cannot be easily shown when expressing the protein of interest transiently in episomes. Experiments that involve the tagging of IMS proteins may cause disruption of the function (e.g. because of mislocalization) or could be a property of Erv itself. The difficulty to tag Erv with proteins such as mCherry limited the

proteomic strategies in this work since I did not consider to fuse APEX2 to Erv. This technique was used with excellent results to generate an IMS proteome in mammals by proximity labelling using biotin [178].

I also could not modify the redox pair CQVYC-motif at residue 303 that is part of the intervening sequences [83]. The identity of amino acids adjacent to the cysteine has been pointed out to play a role in the redox potential of sulfhydryl oxidases [83] and kinetics of thioredoxins [179]. Indeed, modifications in this region led to the generation of trapping mutants that identified novel substrates in PDI [180]. The observation that the mutation was skipped (but not the cassette) provided evidence that this position was essential for survival. The recodonisation of *ERV* to avoid such mechanism and the impossibility to get mutants after transfection further supported this observation. Therefore, the experiments suggest that the tyrosine in position 303 (conserved in Kinetoplastida) may be essential in *L. tarentolae.* This would also further support the idea that the redox pair motif could play an essential role in the function that distinguishes Erv-homologues from each other [83, 181]. Interestingly, the CQVYC-motif of *L. trarentolae* could replace the CRSC-motif of *Sc*Erv1 but not *vice versa* in the mitochondrial yeast model [131].

In summary, the endogenous Erv mutants indicate that the KISS domain and cysteine 17 are not essential for survival. Apparently, the essential operations of Erv do not require the presence of these two components. However, the lack of the KISS domain seems to alter the normal function of Erv in a mechanism that remains to be elucidated. Furthermore, Erv and the intervening sequences of the redox pair were refractory to deletion, providing evidence that are essential.

4.1.2 Life without Mia40

Since the first description of Mia40 in the yeast model [65] and an incomplete MIA pathway in Kinetoplastida and apicomplexans [20, 57, 70], an exciting debate is held in the field. The hypotheses diverge. However, all models have in common the necessity to explain how Erv functions in a context where Mia40 was not found [20, 21, 70, 73, 78, 169, 182]. The most prominent hypothesis suggests that the organisms that lack Mia40 are exceptional examples of free unicellular early branched eukaryotes. Instead of being the rule, in these cells, a streamlined and “minimalistic” import system is sufficient [70]. In this context, the model

proposes a stepwise evolution where an ancestral simplified pathway contained only Erv. Later, during the diversification of the eukaryotes, Mia40 was added. Extensive research was done in this matter and indirect evidence was offered as a support of this model. To name a few, Mia40 is dispensable in plants [183], and Erv from the same organism can partially replace the yeast homolog and oxidize substrates alone. However, the replacement impaired growth and could not exclude the residual activity of Mia40 (temperature-sensitive mutant). Furthermore, in shuffling experiments the complementation was only possible with a redox inactive Mia40, possible highlighting the essential function of the Mia40 hydrophobic cleft to bind substrates [169]. The fact that Erv can oxidize substrates alone is plausible if is considered that QSOX, a related flavoenzyme in the secretory pathway of yeast, introduces disulfide bonds alone [184]. However, studies done by Specht *et al*. using *Lt*Erv could not complement the loss of Mia40 [131]. This could suggest that a non-identified Mia40 homologue could be present in this model. But what do we know about the function of Erv in kinetoplastids concerning the mitochondrial import pathway into the IMS? A breakthrough study done by Peikert *et al.* [160] revealed that the ablation of Erv by RNAi also reduces the import level of known and novel IMS substrates. It is also known that *Leishmanian* Erv and a substrate are essential for survival [131]. Therefore, we can think of a model in which Erv mediates the import of substrates into the IMS of Kinetoplastida in a yet not fully-understood but essential pathway. We reasoned that the experiments done by Peikert *et al.* did not offer a solution as to whether Erv operates alone or with a Mia40 homolog. Similar results would have been obtained for Erv even if a Mia40 homologue is present or not. Therefore, I decided to fuse a His-tag to Erv and Erv$^{\Delta KISS}$ using CRISPR/Cas9 to obtain endogenously tagged baits and generate an Erv interactome. This would test the models explained earlier. The principle, in theory, was simple: If Erv operates alone, IMS substrates should be detected in denaturing conditions evidencing covalent interactions. A similar approach was used previously to describe the Mia40 interactome [144]. In contrast, if Erv operates together with a Mia40 homolog, the main interaction detected should correspond to this protein. It is worth mentioning that the experiments carried out in this work would have missed a non-redox adapter that would act as a receptor not interacting via disulphide bonds.

In the case of the Erv-His$_8$ interactome, SILAC-based affinity purification coupled with mass spectrometry did not reveal the presence of any known substrates with the CX3C or CX9C-

motifs. Instead, I detected proteins in which a putative function in the disulphide relay system is not easily understandable. I detected a mitochondrial tryparedoxin peroxidase (LtaP15.1080) which is essential in *L. infantum* [185], and its homolog in *T. brucei* (Tb927.8.1990) was also found in the Erv depletome [160]. This protein was also found in the same model organism in a pull-down experiment using the IMS substrate Tim10 as a bait [186]. It is not the first time that a tryparedoxin peroxidase was found in a pull-down using Erv of Kinetoplastida [182]. Previously, this tryparedoxin peroxidase was proposed as a possible candidate replacement for Mia40 [131]. However, is difficult to discern if this is rather a false positive due to the high abundance of peroxiredoxins [127, 135]. Whether this is a false hit or indeed this candidate is somehow related to Erv function (assisting in folding, import, or as a chaperone) remains to be elucidated.

Regarding the Erv$^{\Delta KISS}$-His$_8$ interactome, I could detect putative mixed disulphides bonds that were enriched in the mutant bait. These results are consistent with the observations done in western blot analysis. Similar to the Erv-His$_8$ interactome, this experiment did not found known substrates with the CX3C or CX9C-motifs. However, this time a high portion of the proteins in the enriched quadrant were also found in the *Trypanosome brucei* Erv (*Tb*Erv) depletome [160] and some of their predicted functions are unknown. Among the candidates is worth to mention “unknown function 2” (UF2) (LtaP32.0380). The homolog of this protein (Tb11.v5.0420) was found in the *Tb*Erv depletome, and is reported to be essential in a high-throughput screening in the same organism [187]. This candidate has a CHAYCRCSYC-motif located at the N-terminal that is conserved in Kinetoplastida. The cysteines in this motif could potentially have redox-active functions. Another hit to mention is a protein annotated in *T. brucei* as a riboflavin transmembrane transporter activity (LtaP05.0490). If this interaction can be linked somehow with the binding of the FAD cofactor during Erv biogenesis could be an interesting hypothesis to analyse. Furthermore, it is worth to mention the enrichment of the cytochrome *c* oxidase subunit iv (LtaP12.0690) also present in *Tb*Erv depletome (Tb927.1.4100) and the detection of an import receptor from the ATOM40 complex (LtaP35.4850). To assess the physiological role and the relevance of the substrates, a first line of experiments could be designed. The candidates could be tested for essentiality and localization. The interaction should also be further validated with an orthogonal experiment (e.g reverse pulldown using the substrate as a bait showing that Erv is detected). To assess if

one of the candidates could be the Mia40-replacement, the selection should be based first in its essentiality. Some of the substrates imported by the MIA pathway are essential, therefore, is expected that the Mia40-replacement is essential as well. A rapid cross-species complementation in yeast could also be interesting to test. For example, shuffling experiments could test if the *Leishmania* candidate could replace Mia40 in yeast. Similarly, an experiment in *L. tarentolae* could test if Mia40 could complement the loss of the essential candidate. The successful complementation would link the function of the candidate with Mia40 operations and provide evidence that is the Mia40-reaplecement in kinetoplastid parasites.

Recently, a thioredoxin-like protein Mic20 was proposed to be the Mia40 homolog replacement in kinetoplastid parasites in studies done in *T. brucei* [78]. This would suppose that Mic20 and Erv can interact *via* disulphide bonds. The evidence offered by the authors is that the knockdown regulation of Mic20 phenocopied the effect seen for the RNAi ablation of Erv [160]. *L. tarentolae* has two homologues for Mic20 (LtaP33.1480, LtaP26.1460) and we could detect one of them in whole-cell lysates controls (LtaP26.1460). We could not detect Mic20 in our interactomes using Erv and Erv$^{\Delta KISS}$. However, lack of detection might simply be owing to experimental limitations to produce a high quality Erv interactome and not the absence of Mic20 or substrates. Similar results were reported before where a *T. brucei* Erv intaractome did not show canonical substrates or Mic20 [182]. A possible explanation (although not easy to refute) is that Erv has very fast catalysis. This is the opposite of the system well described in yeast. Is it known that Mia40 can form stable interactions detectable in western blots [66] because it forms long-lived mixed disulphides [188]. In contrast, Erv in Kinetoplastida may resolve its interaction/s faster similar to PDI and DbsA [155]. This would make mixed disulphides short-lived and a challenge to detect in affinity purification experiments. For this reason, I attempted to generate mutants that could slow down this reaction but failing to obtain viable cells for the intervening sequences of the CQVYC-motif. Another explanation for the proteome detected in this work is related to the special features of the IMS. The IMS proteome is difficult to separate from the cytosol, since probably most of them are dual localized [59]. A well-studied case for example, is the superoxide dismutase SOD1 [189]. If Erv could not be shown to interact with the broad IMS substrate repertoire or a Mia40 replacement by affinity purification what other alternatives remain to study this

system? A reverse pull-down showing that Mic20 or similar candidates could bind Erv will be essential to corroborate the model. Reverse pull-downs using substrates as baits are currently conducted by the PhD student Luzia Schneider and will provide essential evidence if the baits interact with Erv, Mic20 or a distinct candidate. However, technical difficulties might arise in visualizing this type of interaction in Kinetoplastida. Improvements in affinity purification protocols and the development of new alternatives will certainly be essential to address this issue. An *in vitro* reconstituted assay would also be of value. This way, a recombinant produced candidate and Erv could be shown to effectively oxidize substrates together and reduce cytochrome *c*. Experiments of this type have been done in other MIA pathway models [73, 169, 190, 191].

The report of Mic20 as a candidate supported by its phenotype [78] is an important advance to the field. Proposing Mic20 as Mia40 replacement in Kinetoplastida leads to a model that can be experimentally challenged and maybe refuted. However, although the description of Mic20 proposes the first two-components-model where IMS substrates are imported without Mia40, still do not explain how other eukaryotic systems work without Mia40. For example, it is still unclear how Erv operates in apicomplexan parasites since Mic20 seems to be restricted to Kinetoplastida and related euglenids [9, 192]. For this reason, previous PhD students in the lab have not considered Mic20 to be a replacement of Mia40, since it was first described in the Erv depletome [160]. Mic20 has two isoforms in *L. tarentolae* (LtaP33.1480, LtaP26.1460), to test if it is the Mia40 replacement, a knockout experiment targeting both loci should be conducted to assess if it is essential. If Mic20 is essential and is responsible of the oxidation of substrates, mutations in the proposed redox pair, the CIPC-motif should be essential (e.g. cysteine-to-serine substitution). In contrast the generation of mutant viable cells would dismiss Mic20 as the Mia40 replacement. Furthermore, a cross-species complementation assay could also be conducted. If Mic20 is essential in *L. tarentolae*, an experiment could test if Mia40 could complement the loss of Mic20 in this model.

In summary, much remains to be studied in the cryptic and incomplete MIA pathway from kinetoplastid parasites. Is probably worth to test which are the physiological roles of the substrates imported by this pathway in kinetoplastid parasites. In comparison with the extensive studies describing the pathway components (Mia40 and/or Erv), is striking the asymmetry of knowledge we have concerning the functions of the substrates itself. Here I

offer new candidates whose functions are unknown and could potentially be related to Erv and provide a new understanding of the function of the IMS compartment.

4.2 Investigation of general genetic techniques for the study of essential genes

An extended strategy to unravel fundamental gene-function relationships in the cell is to study genes that are essential for cell viability. Efforts of trying to gain an understanding of the gene-function of essential pathways is a subject of intense study on parasites of relevance to human health. The target-based drug development strategy for essential pathways has been one of the leading principles to treat infectious parasitic diseases [28, 193-196]. In this context, a first step consists of determining if the target gene encodes for a protein that is essential for survival by a knockout experiment [175, 197-200]. Although this preliminary genetic experiment sets the necessary correlation between the expression of the target gene and cell survival, it does not proportionate any evidence as to why this gene is essential, nor does it necessarily explain what is the function or contribution within the studied pathway. For this reason, the generation of a conditional mutant where cell viability is dependent on growth conditions is key for studying gene function *in vivo* [27, 130, 143, 201]. We chose to study two of these tools in *L. tarentolae*: the function of a negative marker system to perform shuffling experiments and the *glmS* riboswitch to generate knockdown mutants.

4.2.1 Negative selection systems in *Leishmania*

There are currently two markers tested in *Leishmania* encoding enzymes that convert non-toxic compounds to ones that are toxic for wild-type cells: The thymidine kinase (*TK*) [202] and the yeast cytosine deaminase (*CD*) [203]. The usefulness of the counter-selection in shuffling experiments depends on the “cidal” activity and potency of the drug. This is especially relevant when the plasmid rescues an essential gene or a lethal mutation is being tested. In the case of *TK*, Papee *et al.* reported that ganciclovir could not force the loss of a plasmid rescuing an essential gene [27] as is commonly done in yeast [204]. Therefore, we decided to test first, an improved construct based on the *CD* gene, the cytosine deaminase and uridyl phosphoribosyl transferase (*yFCU*) [205, 206]. The latter gene was reported as a

counter-selectable marker in the protozoan parasite *Plasmodium falciparum* [207, 208]. In this work, the exogenous addition of 5-FC to pX-*NEO*-*T2A*-*yFCU*-transfected parasites failed to induce any growth defect. Although evidence of *yFCU* expression was not given, the *NEO* expression served as a control since both genes were fused with skip peptide T2A. Therefore, the results described in this work indicate that the yFCU-5FC system is presumably incompatible or inefficient in *L. tarentolae.*

Owing to the negative results obtained for *yFCU* and the lack of efficiency reported for *TK,* we considered that an efficient negative marker was missing. However, the results obtained in the characterization of Δ*APRT* in this work set the perfect frame to find if *APRT* could be used as a novel negative selection tool in *Leishmania*. I described the high toxicity of APP for wild-type cells and the non-toxicity for the same drug in Δ*APRT* lines. To test if the expression of *APRT* was efficient to induce plasmid segregation, I generated a transgenic line. This mutant parasite contained a deletion for *APRT* (Δ*APRT)* on the chromosome and a wild-type copy of *APRT* on a plasmid. The presence of the plasmid was evaluated after exposing cells to APP, measuring its effect on parasite growth and fluorescence. These experiments showed that the plasmid was being excluded upon APP counter-selection. The most important requirement in an efficient negative marker is that the toxicity generated by the gene expression is so high that the plasmid is lost even if the selected genotype is also lethal. To imitate this situation, I selected for the loss of the pPLOT-*BLAST*-*T2A*-*APRT* plasmid with APP in the presence of blasticidin. The selection led to cell death, obtaining virtually no viable cells as compared to controls. These results were expected because of the potency of APP (EC_{50} =4 µM) and the strong resistance found in Δ*APRT* (tested up to 400 µM). In comparison, in a similar experiment for the *TK* system, ganciclovir was used at 200 µM without selecting for the loss of the *TK*-expressing plasmid even after 5 passages [27]. To date, and to the best of our knowledge, it is not feasible to do classical plasmid shuffling experiments in *Leishmania.*

The results shown in this work provide a first line of experiments for a new method for counter-selection of the *APRT* gene in future shuffling experiments as proposed in Figure 4.1. The Δ*APRT* strain described earlier was obtained using a targeting cassette expressing puromycin *N*-acetyltransferase. Instead, and for sake of sparing a valuable antibiotic resistance cassette, a marker-free Δ*APRT* line could be also constructed. Similarly, as described in this work, CRISPR/Cas9 could be used to generate a knockout cassette expressing

the fluorescent protein mNeonGreen (without an antibiotic resistance cassette) and select the Δ*APRT* mutant with 400 µM APP.

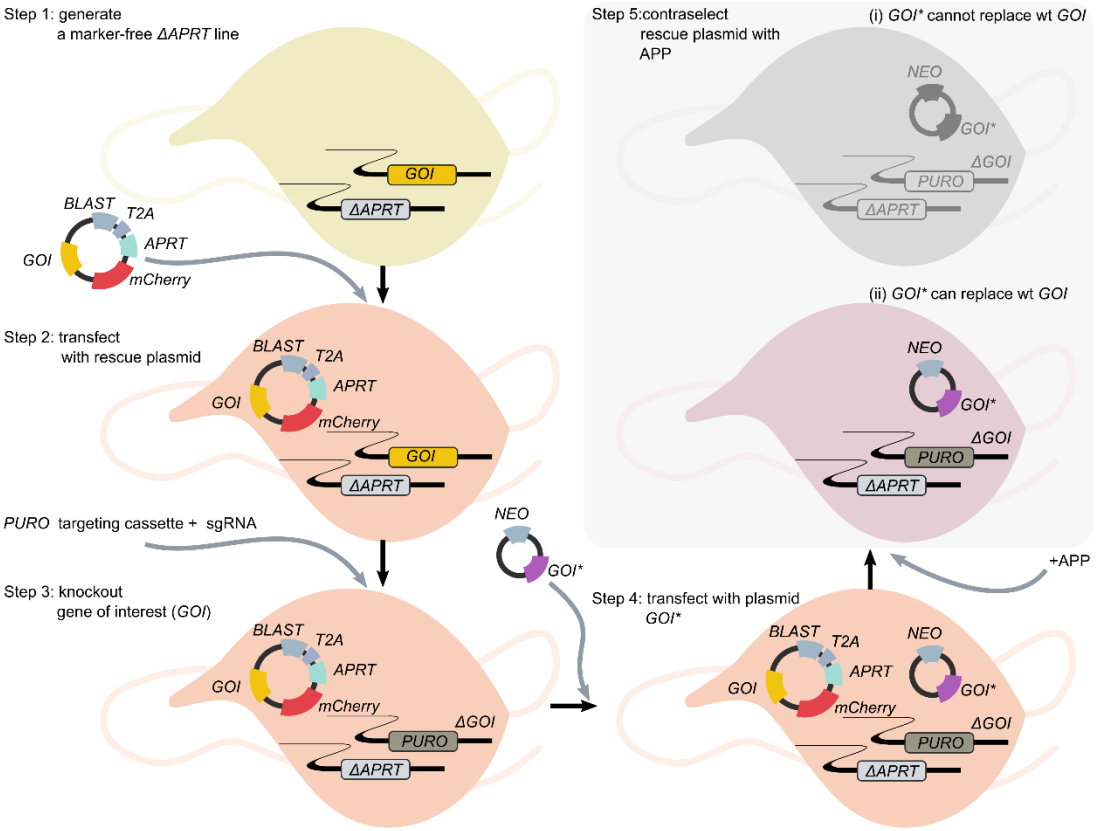


Figure 4.1 Proposed shuffling experiments using the APRT/APP system.
Proposed steps to perform shuffling experiments. **Step 1**, a marker-free Δ*APRT* line should be generated. **Step 2**, the Δ*APRT* line should be used to transfect the rescue pPLOT-*BLAST-T2A-APRT-GOI* plasmid. **Step 3**, the gene of interest (*GOI*) should be subsequently deleted the using, for example, a targeting cassette expressing puromycin *N*-acetyltransferase. **Step 4**, the strain generated in the former step is then transformed by a second plasmid containing a mutation, a construct chimera, or a gene encoding for a homolog or functional replacement of the *GOI* (e.g. pX-*NEO-GOI**). **Step 5**, selection with APP. Whether the GOI* gene can complement the loss of the wild type GOI copy present in pPLOT-BLAST-T2A-APRT-GOI (violet cell) or not (gray cell), is revealed upon APP selection. The counter-selection of the wild-type *GOI* copy by APP will render cells exclusively dependable on the mutant *GOI**. If cells survive this experiment, it could be interpreted as evidence that the mutant or homolog etc. can replace the function of the essential gene studied.

In summary, I could not find evidence that pX-*NEO-T2A-yFCU* transfectants render cells sensitive to 5FC, highlighting the necessity of alternative efficient negative markers. In contrast, the APRT/APP system seems to provide the ability to segregate a plasmid even if the genotype selected is lethal. Therefore, the capacity of the system suggests that a novel and efficient negative marker could be used in *Leishmania.*

4.2.2 The *glmS* riboswitch in *Leishmania tarentolae*

The generation of an efficient knockdown system in *Leishmania* is a technical problem not yet resolved since Robinson *et al.* found that RNAi was incompatible in *Leishmania* major and *Leishmania donovani* [119]. Due to the former reason and the additional transcriptional nature of *Leishmania* [15] we chose to examine if the *glmS* riboswitch could be activated in *L. tarentolae* in analogy to other eukaryotic models [126, 127, 209]. The major concern was to find if this activation has the magnitude to induce a knockdown phenotype at the chromosomal level. The first episomal reporter experiment provided evidence that the knockdown effect was effectively taking place without the addition of glucosamine. Additionally, this effect could also be further enhanced if glucosamine was exogenously provided in the medium. The orders of magnitude in the downregulation were not as extensive when compared to other knockdown regulation reports in *Trypanosoma* [130, 197, 209]. However, we decided to finally corroborate if there was any effect in fusing *glmS* endogenously in genes whose phenotype and function are known or predicted. The experiments did not find an apparent knockdown phenotype. We chose *APRT*, *PF16* and *ERV* because I characterized the corresponding knockout phenotypes. Furthermore, the knockdown phenotypes generated by RNAi were also described for *PF16* and *ERV* in *T. brucei*, showing no motility and a defect in cell growth [73, 153]. Proteomic and transcriptomic evidence verifying the presence or absence of downregulation of each of the genes tested is missing in this work (western blot analysis and RT-qPCR respectively). Nonetheless, the correct fusion of an already tested sequence [209] was provided by PCR and Sanger sequencing. Since the execution of these experiments will unlikely provide a functional tool, an explanation as to why the *glmS* riboswitch did not efficiently knockdown gene expression in *L. tarentolae* is missing. A possible explanation is that, although a gene downregulation was taking place with or without the addition of glucosamine, such regulation was not sufficient to render cells with a measurable phenotype. For example, the turnover of proteins depends on numerous factors among which stability plays an important role. For example, the reporter mCherrry is considered a stable protein. In contrast, in nature, the glmS protein (controlled by the *glmS* riboswitch) is rather unstable [210]. Another further explanation (not excluding the first one) is that certain phenotypes can be elusive, and the correct differentiation between wild type and mutant phenotype could require specific growth conditions not known

beforehand [171, 172]. The mCherry reporter-transfectants characterized in this work can be thought of as sensors of glucosamine-6-phosphate (GlcN6P) during post-transcriptional gene regulation. The glucosamine metabolism in *Leishmania* has been previously described highlighting its essential role in protein glycosylation [211]. The main source of GlcN6P is catalysed by fructose-6-phosphate aminotransferase (*GFAT*) from endogenous fructose-6-phosphate. The ablation of the latter gene was only possible if mutants were grown on 2.8 mM glucosamine. These results are in line with the findings of this work, where the *glmS* riboswitch can be activated endogenously without further addition of glucosamine. A similar phenomenon was reported by Lander *et al.* in *Trypanosome cruzi* [130]. The generation of a *GFAT* knockout and the glucosamine requirement also provided evidence that glucosamine can enter the cell and be converted to GlcN6P [211]. Another putative source of the *glmS*-inducer GlcN6P is the activity of the *N*-acetyl glucosamine 6-phosphate deacetylase (*GNAT*). Recently, it was suggested to be one key factor for the *glmS* system to be functional and is also present in *Leishmania* [130]. In summary, the results outlined in this work suggest that despite the evidence that *Leishmania* can synthesize the *glmS*-inducer GlcN6P, somehow this synthesis is not sufficient to efficiently generate a knockdown phenotype at the chromosomal level. Whether a successful knockdown system could be further developed by a correct media culture condition, the generation of a mutant strain where the GlcN6P is kept at higher levels, or if the system is incompatible with intrinsic features in *Leishmania* biology, is subject to further research. The lack of an efficient inducible knockdown system represents an important bottleneck in the study of essential genes in *Leishmania*. An alternative solution to regulate the amount of protein could be based in the shuffling experiments with *APRT* described in section 4.2.1. A pPLOT-*BLAST-T2A-APRT-GOI*-containing *L. tarentolae* line could be obtained to subsequently generate a *ΔGOI* line. Thereafter, instead of a complete counter-selection with 400 µM APP a potential “knockdown” phenotype could be assessed by regulating the amount of plasmid-*GOI* using APP in a dose-dependent manner.

4.3 The CRISPR/Cas9 system in *Leishmania tarentolae*

The introduction of the CRISPR/Cas9 system as a novel genetic editing tool provoked a revolution in the parasitology field [96, 212, 213]. In kinetoplastid parasites, many strategies were designed to offer solutions in terms of the delivery of the components for this system to function: To name a few, systems differ in how the Cas9 is expressed (from a plasmid or integrated into the chromosome), which promoter is used to transcribe the sgRNA, and the nature of the donor DNA cassette (contained in a plasmid or fragment and the length of the homology arms). In the first year of my doctoral studies, I attempted to improve a strategy that in my opinion was the best at that time [107]. To spare an antibiotic resistance cassette and simplify transfection procedures, I constructed a single plasmid (pLPC) containing the needed components Cas9 and sgRNA cassettes. The original Cas9 reported by Zhang *et al.* was also replaced by an improved version fused with GFP. This intended to reduce the probability of off-target events and facilitate the identification of Cas9-expressing transfectants by GFP. Finally, the use of a suicidal cassette eliminated the necessity of using a negative marker as is required when the donor DNA is transfected as a plasmid [100]. The principle of combining the CRISPR/Cas9 components in one single vector is not new. Indeed, years later the same authors reported their improved version using one single plasmid [214]. Unfortunately, the system I constructed failed to generate a proof-of-concept knockout of *APRT*. Given that the sgRNA was functional, a possible explanation is that the improved Cas9 used was not efficient. Similar results were obtained when this Cas9 was used in *T. brucei* (Roberto Docampo, personal communication). The selection of GFP as a reporter was also a mistake in the design since I could not identify fluorescent colonies in a BHI agar plate containing pLPC as expected. This could be explained because of the high autofluorescence of *Leishmania* as previously reported [215]. Using a fluorescent reporter to detect expression in *L. tarentolae* colonies grown on BHI agar is possible, as exemplified by mCherry for the LEXSY expression system [42], even though the LEXSY system also uses the strong T7 promoter. I also attempted to generate an *APRT* knockout using classic genetics, without successful results. This provided an already known lesson for molecular parasitologists. Although the generation of mutants in *Leishmania* was possible from the early 1990s [94, 95, 216], classic genetic techniques are inefficient. The low probabilities of success make this

procedure cumbersome involving projects that could take years in the generation of the desired mutant (Helena Castro, personal communication).

These negative results were contemporaneous with the report by Beneke *et al.* for a new CRISPR/Cas9 strategy in *Leishmania mexicana* [109]. Therefore, we decided to adopt this system and test it in *L. tarentolae.* As the first step, I generated and validate a TB005 line expressing Cas9 and T7. A growth curve was also analysed dismissing toxic effects caused by Cas9 or T7RNA polymerase. Similar results were reported previously [109, 217], although growth-impairments caused by Cas9 have been described in other protists [106, 218]. The system was then used to generate more than 17 mutants in this work with unparalleled simplicity and efficiency. One of the reasons is that no cloning steps are involved and primers are interchangeable for knockout or knockin experiments. Moreover, the transfection protocol is compatible with PCR products that do not need further purification. The system was optimized to generate knockouts and fuse proteins with diverse tags. However, because of the good design of the targeting cassettes, I easily adapted it to generate single-base substitutions as described for the ERV^{C17S} mutant. The skipping of the homology arm was used to improve the cassette with a recodonisied version of *ERV*. This principle was employed to generate the single-base substation described earlier, the $ERV^{\Delta KISS}$, and the putative lethal effects in the intervening sequence present in the CQVYC-motif.

Further improvements of the technique could also be tested in the future. I explored the possibility of site saturation mutagenesis in the codon coding for tyrosine at position 303 of Erv without the generation of viable cells (data not shown). The technique consists of the use of degenerate primers that are synthesized commercially as wobble bases. Wobble bases or mixed bases contain equimolar mixtures of two or more different bases within the primer sequence and can be part of a codon. Primers that include mixed bases could be used in the generation of targeting cassettes (equivalent to the strategy shown in Figure 3.9). Thus, the generation of a targeting cassette would contain a small library that not only includes a tyrosine-to-aspartate substitution for example, but up to eight distinct codons coding for distinct amino acids. Therefore, the technique would offer a rapid, cost-effective technique to screen the effect of several substitutions by using a single degenerate primer. For example, wobble primers can be purchased without extra cost on Metabion (Planegg, Germany). The

design of degenerate primers that include wobble bases can be easily designed with an online tool [219].

In summary, the CRISPR/Cas9 system was successfully tested in *L. tarentolae* showing its great utility for the rapid generation of a collection of mutants. The system was slightly modified in this work to allow *glmS,* His and Strep tagging. Other modification in this study also allowed the generation of a truncated and a single base substitution mutant of Erv. The latter could serve as examples to generate similar constructs in genes that are also essential.

4.3.1 Conclusion

Any scientific project that aims to use the state of the art techniques, would greatly benefit from having beforehand an accurate knowledge of which systems work properly. In this thesis, I did the necessary groundwork to provide such knowledge. I tested novel genetic techniques in *Leishmania tarentolae* and streamlined protocols that could be useful in future projects in this model. I showed the simplicity of the CRISPR/Cas9 system to generate endogenous mutants, covering all the examples possible (knockout, tagging, single base substitution, and truncation). It is now possible to modify genes at the chromosomal level where the limit is virtually the elegance and imagination of the experiment. I also tested the *glmS* system, where I found no evidence of a knockdown phenotype, highlighting the necessity of alternative solutions to generate these mutants. I showed that *APRT* could be an excellent novel negative selection marker, that could allow us to perform cross-species complementation studies and efficient shuffling experiments for the first time in *Leishmania.* I streamlined and tested a protocol to perform SILAC experiments that could be used for several proteomic approaches. Concerning the genetics and proteomic studies on Erv, I showed with a few examples the significance of the conclusions when mutants are obtained at the chromosomal level. These experiments could offer hints about the function of the KISS domain, cysteine 17, and intervening sequences in the redox pair of *Leishmania* Erv. Finally, the Erv interactome generated in this work could lead to new questions and hypotheses. For this, it will be essential to understand which are the functions of these candidates and which is the relationship (if any) with the oxidoreductase Erv.

CHAPTER

SUMMARY 5

Compared to canonical model organisms, the genetic toolbox of Kinetoplastid parasites have a considerable gap in the transgenic techniques available. The implementation of the CRISPR/Cas9 technology is poised to transform the way we perform genetic manipulations and offers a new and exciting horizon for molecular parasitology. In this study, we use the Kinetoplastid parasite *Leishmania tarentolae* as a model organism. This unicellular eukaryote is an attractive model for both basic and applied research. Understanding *Leishmania's* basic biology is valuable to underpin differences to the host that might help to treat infectious diseases. Furthermore, it also provides new examples of non-conserved mechanisms that will help to understand the fundamental principles of the biology of eukaryotes and their evolution. In this work, the CRISPR/Cas9 system was used to study mitochondrial protein import.

Here I show the efficacy of CRISPR/Cas9 to generate knockout and knockin mutants. Proof- of- concept gene *PF16* was used to generate knockout immotile parasites and knockin fluorescent mutants fused with mCherry. The *APRT* gene was also knocked out showing resistance to APP.

In addition, I generated endogenous mutants of a constituent of the mitochondrial import machineries, the sulfhydryl oxidoreductase Erv. I showed that the KISS domain and cysteine 17 are dispensable for survival dismissing that their functions correlate with the essential operation/s of Erv. I report that the *ERV* gene and the intervening sequences of its shuttle pair cysteines are refractory to ablation and modification, respectively, indicating that they are essential for survival. I also generated Erv interactomes using full-length and mutant ($Erv^{\Delta KISS}$)

baits showing candidates with hitherto unknown functions that might be related to Erv function.

I also tested the *glmS* riboswitch and generate endogenous mutants with CRISPR/Cas9. We asked if it was possible in *Leishmania* to obtain knockdown mutants with this technique. The evidence of this study indicates that the system is inefficient in provoking a knockdown phenotype for the genes characterized.

An alternative negative marker was also developed in this work. I propose the *APRT* gene as a novel and efficient counter-selectable marker as compared to the current *yFCU* and *TK* genes. The implementation of this system could lead to first shuffling experiments that are not feasible in *Leishmania* further highlighting the value of this model organism.

REFERECES 6

1. Caron, D.A., et al., *Protists are microbes too: a perspective.* Isme j, 2009. **3**(1): p. 4-12.
2. Adl, S.M., et al., *The revised classification of eukaryotes.* J Eukaryot Microbiol, 2012. **59**(5): p. 429-93.
3. Scamardella, J.M., *Not plants or animals: a brief history of the origin of Kingdoms Protozoa, Protista and Protoctista.* Int Microbiol, 1999. **2**(4): p. 207-16.
4. Vickerman, K., *The evolutionary expansion of the trypanosomatid flagellates.* Int J Parasitol, 1994. **24**(8): p. 1317-31.
5. Simpson, A.G., Y. Inagaki, and A.J. Roger, *Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of "primitive" eukaryotes.* Mol Biol Evol, 2006. **23**(3): p. 615-25.
6. World Health, O., U.N.U.W.B.W.S.P.f. Research, and D. Training in Tropical, *Global report for research on infectious diseases of poverty 2012*. 2012, World Health Organization: Geneva.
7. Organization, W.H., *Investing to Overcome the Global Impact of Neglected Tropical Diseases: Third WHO Report on Neglected Tropical Diseases 2015*. 2015: World Health Organization.
8. Barrett, M.P., et al., *The trypanosomiases.* Lancet, 2003. **362**(9394): p. 1469-80.
9. Hashimi, H., *A parasite's take on the evolutionary cell biology of MICOS.* PLoS Pathog, 2019. **15**(12): p. e1008166.
10. Simpson, A.G., J.R. Stevens, and J. Lukes, *The evolution and diversity of kinetoplastid flagellates.* Trends Parasitol, 2006. **22**(4): p. 168-74.
11. Harmer, J., et al., *Farming, slaving and enslavement: histories of endosymbioses during kinetoplastid evolution.* Parasitology, 2018. **145**(10): p. 1311-1323.
12. Simpson, A.G.B., J. Lukeš, and A.J. Roger, *The Evolutionary History of Kinetoplastids and Their Kinetoplasts.* Molecular Biology and Evolution, 2002. **19**(12): p. 2071-2083.
13. Rodrigues, J.C., J.L. Godinho, and W. de Souza, *Biology of human pathogenic trypanosomatids: epidemiology, lifecycle and ultrastructure.* Subcell Biochem, 2014. **74**: p. 1-42.
14. Langousis, G. and K.L. Hill, *Motility and more: the flagellum of Trypanosoma brucei.* Nat Rev Microbiol, 2014. **12**(7): p. 505-18.
15. Clayton, C.E., *Life without transcriptional control? From fly to man and back again.* Embo j, 2002. **21**(8): p. 1881-8.
16. Clayton, C.E., *Networks of gene expression regulation in Trypanosoma brucei.* Mol Biochem Parasitol, 2014. **195**(2): p. 96-106.

17. Simpson, L. and J. Shaw, *RNA editing and the mitochondrial cryptogenes of kinetoplastid protozoa.* Cell, 1989. **57**(3): p. 355-66.
18. Haanstra, J.R., et al., *Biogenesis, maintenance and dynamics of glycosomes in trypanosomatid parasites.* Biochim Biophys Acta, 2016. **1863**(5): p. 1038-48.
19. Field, M.C. and M. Carrington, *The trypanosome flagellar pocket.* Nat Rev Microbiol, 2009. **7**(11): p. 775-86.
20. Eckers, E., et al., *Mitochondrial protein import pathways are functionally conserved among eukaryotes despite compositional diversity of the import machineries.* Biol Chem, 2012. **393**(6): p. 513-24.
21. Schneider, A., *Evolution of mitochondrial protein import - lessons from trypanosomes.* Biol Chem, 2020. **401**(6-7): p. 663-676.
22. Harsman, A. and A. Schneider, *Mitochondrial protein import in trypanosomes: Expect the unexpected.* Traffic, 2017. **18**(2): p. 96-109.
23. Akhoundi, M., et al., *A Historical Overview of the Classification, Evolution, and Dispersion of Leishmania Parasites and Sandflies.* PLoS Negl Trop Dis, 2016. **10**(3): p. e0004349.
24. Torres-Guerrero, E., et al., *Leishmaniasis: a review.* F1000Res, 2017. **6**: p. 750.
25. Sacks, D. and N. Noben-Trauth, *The immunology of susceptibility and resistance to Leishmania major in mice.* Nat Rev Immunol, 2002. **2**(11): p. 845-58.
26. Saunders, E.C., et al., *Induction of a stringent metabolic response in intracellular stages of Leishmania mexicana leads to increased dependence on mitochondrial metabolism.* PLoS Pathog, 2014. **10**(1): p. e1003888.
27. Paape, D., et al., *Genetic validation of Leishmania genes essential for amastigote survival in vivo using N-myristoyltransferase as a model.* Parasit Vectors, 2020. **13**(1): p. 132.
28. Jones, N.G., et al., *Genetically Validated Drug Targets in Leishmania: Current Knowledge and Future Prospects.* ACS Infect Dis, 2018. **4**(4): p. 467-477.
29. Klatt, S., et al., *Leishmania tarentolae: Taxonomic classification and its application as a promising biotechnological expression host.* PLoS Negl Trop Dis, 2019. **13**(7): p. e0007424.
30. Contreras, I., et al., *Leishmania-induced inactivation of the macrophage transcription factor AP-1 is mediated by the parasite metalloprotease GP63.* PLoS Pathog, 2010. **6**(10): p. e1001148.
31. Fritsche, C., et al., *Characterization of the growth behavior of Leishmania tarentolae: a new expression system for recombinant proteins.* J Basic Microbiol, 2007. **47**(5): p. 384-93.
32. Raymond, F., et al., *Genome sequencing of the lizard parasite Leishmania tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species.* Nucleic Acids Res, 2012. **40**(3): p. 1131-47.
33. Goto, Y., et al., *Draft Genome Sequence of Leishmania tarentolae Parrot Tar II, Obtained by Single-Molecule Real-Time Sequencing.* Microbiol Resour Announc, 2020. **9**(21).
34. Blum, B., N. Bakalara, and L. Simpson, *A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information.* Cell, 1990. **60**(2): p. 189-98.
35. Ouellette, M., et al., *Direct and inverted DNA repeats associated with P-glycoprotein gene amplification in drug resistant Leishmania.* Embo j, 1991. **10**(4): p. 1009-16.

36. Papadopoulou, B., et al., *Gene disruption of the P-glycoprotein related gene pgpa of Leishmania tarentolae.* Biochem Biophys Res Commun, 1996. **224**(3): p. 772-8.
37. Kündig, C., et al., *Increased transport of pteridines compensates for mutations in the high affinity folate transporter and contributes to methotrexate resistance in the protozoan parasite Leishmania tarentolae.* Embo j, 1999. **18**(9): p. 2342-51.
38. Ouellette, M., et al., *Pterin transport and metabolism in Leishmania and related trypanosomatid parasites.* Int J Parasitol, 2002. **32**(4): p. 385-98.
39. Saito, R.M., M.G. Elgort, and D.A. Campbell, *A conserved upstream element is essential for transcription of the Leishmania tarentolae mini-exon gene.* Embo j, 1994. **13**(22): p. 5460-9.
40. Lukes, J., et al., *Translational initiation in Leishmania tarentolae and Phytomonas serpens (Kinetoplastida) is strongly influenced by pre-ATG triplet and its 5' sequence context.* Mol Biochem Parasitol, 2006. **148**(2): p. 125-32.
41. Breitling, R., et al., *Non-pathogenic trypanosomatid protozoa as a platform for protein research and production.* Protein Expr Purif, 2002. **25**(2): p. 209-18.
42. Kushnir, S., et al., *Artificial linear episome-based protein expression system for protozoon Leishmania tarentolae.* Mol Biochem Parasitol, 2011. **176**(2): p. 69-79.
43. Basile, G. and M. Peticca, *Recombinant protein expression in Leishmania tarentolae.* Mol Biotechnol, 2009. **43**(3): p. 273-8.
44. Langer, T., et al., *Expression and purification of the extracellular domains of human glycoprotein VI (GPVI) and the receptor for advanced glycation end products (RAGE) from Rattus norvegicus in Leishmania tarentolae.* Prep Biochem Biotechnol, 2017. **47**(10): p. 1008-1015.
45. Mannaert, A., et al., *Adaptive mechanisms in pathogens: universal aneuploidy in Leishmania.* Trends Parasitol, 2012. **28**(9): p. 370-6.
46. Prieto Barja, P., et al., *Haplotype selection as an adaptive mechanism in the protozoan pathogen Leishmania donovani.* Nat Ecol Evol, 2017. **1**(12): p. 1961-1969.
47. Gray, M.W., G. Burger, and B.F. Lang, *Mitochondrial evolution.* Science, 1999. **283**(5407): p. 1476-81.
48. Roger, A.J., S.A. Muñoz-Gómez, and R. Kamikawa, *The Origin and Diversification of Mitochondria.* Curr Biol, 2017. **27**(21): p. R1177-r1192.
49. Timmis, J.N., et al., *Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes.* Nat Rev Genet, 2004. **5**(2): p. 123-35.
50. Tokatlidis, K., *Shaping the import system of mitochondria.* Elife, 2018. **7**.
51. Mannella, C.A., *Structure and dynamics of the mitochondrial inner membrane cristae.* Biochim Biophys Acta, 2006. **1763**(5-6): p. 542-8.
52. Baker, K.P. and G. Schatz, *Mitochondrial proteins essential for viability mediate protein import into yeast mitochondria.* Nature, 1991. **349**(6306): p. 205-8.
53. Chacinska, A., et al., *Importing mitochondrial proteins: machineries and mechanisms.* Cell, 2009. **138**(4): p. 628-44.
54. Neupert, W. and J.M. Herrmann, *Translocation of proteins into mitochondria.* Annu Rev Biochem, 2007. **76**: p. 723-49.
55. Herrmann, J.M. and K. Hell, *Chopped, trapped or tacked--protein translocation into the IMS of mitochondria.* Trends Biochem Sci, 2005. **30**(4): p. 205-11.
56. Frey, T.G. and C.A. Mannella, *The internal structure of mitochondria.* Trends Biochem Sci, 2000. **25**(7): p. 319-24.

57. Deponte, M. and K. Hell, *Disulphide bond formation in the intermembrane space of mitochondria.* J Biochem, 2009. **146**(5): p. 599-608.
58. Longen, S., et al., *Systematic analysis of the twin cx(9)c protein family.* J Mol Biol, 2009. **393**(2): p. 356-68.
59. Finger, Y. and J. Riemer, *Protein import by the mitochondrial disulfide relay in higher eukaryotes.* Biol Chem, 2020. **401**(6-7): p. 749-763.
60. Herrmann, J.M. and R. Köhl, *Catch me if you can! Oxidative protein trapping in the intermembrane space of mitochondria.* J Cell Biol, 2007. **176**(5): p. 559-63.
61. Vögtle, F.N., et al., *Intermembrane space proteome of yeast mitochondria.* Mol Cell Proteomics, 2012. **11**(12): p. 1840-52.
62. Jiang, X. and X. Wang, *Cytochrome C-mediated apoptosis.* Annu Rev Biochem, 2004. **73**: p. 87-106.
63. Swenson, S.A., et al., *From Synthesis to Utilization: The Ins and Outs of Mitochondrial Heme.* Cells, 2020. **9**(3).
64. Habich, M., S.L. Salscheider, and J. Riemer, *Cysteine residues in mitochondrial intermembrane space proteins: more than just import.* Br J Pharmacol, 2019. **176**(4): p. 514-531.
65. Chacinska, A., et al., *Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins.* Embo j, 2004. **23**(19): p. 3735-46.
66. Mesecke, N., et al., *A disulfide relay system in the intermembrane space of mitochondria that mediates protein import.* Cell, 2005. **121**(7): p. 1059-69.
67. Peleh, V., E. Cordat, and J.M. Herrmann, *Mia40 is a trans-site receptor that drives protein import into the mitochondrial intermembrane space by hydrophobic substrate binding.* Elife, 2016. **5**.
68. Sevier, C.S. and C.A. Kaiser, *Formation and transfer of disulphide bonds in living cells.* Nat Rev Mol Cell Biol, 2002. **3**(11): p. 836-47.
69. Lutz, T., W. Neupert, and J.M. Herrmann, *Import of small Tim proteins into the mitochondrial intermembrane space.* Embo j, 2003. **22**(17): p. 4400-8.
70. Allen, J.W., S.J. Ferguson, and M.L. Ginger, *Distinctive biochemistry in the trypanosome mitochondrial intermembrane space suggests a model for stepwise evolution of the MIA pathway for import of cysteine-rich proteins.* FEBS Lett, 2008. **582**(19): p. 2817-25.
71. Wiedemann, N., A.E. Frazier, and N. Pfanner, *The protein import machinery of mitochondria.* J Biol Chem, 2004. **279**(15): p. 14473-6.
72. Eckers, E., et al., *Divergent molecular evolution of the mitochondrial sulfhydryl:cytochrome C oxidoreductase Erv in opisthokonts and parasitic protists.* J Biol Chem, 2013. **288**(4): p. 2676-88.
73. Basu, S., et al., *Divergence of Erv1-associated mitochondrial import and export pathways in trypanosomes and anaerobic protists.* Eukaryot Cell, 2013. **12**(2): p. 343-55.
74. Dolezal, P., et al., *The essentials of protein import in the degenerate mitochondrion of Entamoeba histolytica.* PLoS Pathog, 2010. **6**(3): p. e1000812.
75. van Dooren, G.G., et al., *The Import of Proteins into the Mitochondrion of Toxoplasma gondii.* J Biol Chem, 2016. **291**(37): p. 19335-50.
76. de Souza, W., M. Attias, and J.C. Rodrigues, *Particularities of mitochondrial structure in parasitic protists (Apicomplexa and Kinetoplastida).* Int J Biochem Cell Biol, 2009. **41**(10): p. 2069-80.

77. Delage, L., et al., *In silico survey of the mitochondrial protein uptake and maturation systems in the brown alga Ectocarpus siliculosus.* PLoS One, 2011. **6**(5): p. e19540.
78. Kaurov, I., et al., *The Diverged Trypanosome MICOS Complex as a Hub for Mitochondrial Cristae Shaping and Protein Import.* Curr Biol, 2018. **28**(21): p. 3393-3407.e5.
79. Backes, S., et al., *Development of the Mitochondrial Intermembrane Space Disulfide Relay Represents a Critical Step in Eukaryotic Evolution.* Mol Biol Evol, 2019. **36**(4): p. 742-756.
80. Lisowsky, T., *Dual function of a new nuclear gene for oxidative phosphorylation and vegetative growth in yeast.* Mol Gen Genet, 1992. **232**(1): p. 58-64.
81. Riemer, J., N. Bulleid, and J.M. Herrmann, *Disulfide formation in the ER and mitochondria: two solutions to a common process.* Science, 2009. **324**(5932): p. 1284-7.
82. Wu, C.K., et al., *The crystal structure of augmenter of liver regeneration: A mammalian FAD-dependent sulfhydryl oxidase.* Protein Sci, 2003. **12**(5): p. 1109-18.
83. Fass, D., *The Erv family of sulfhydryl oxidases.* Biochim Biophys Acta, 2008. **1783**(4): p. 557-66.
84. Vitu, E., et al., *Gain of function in an ERV/ALR sulfhydryl oxidase by molecular engineering of the shuttle disulfide.* J Mol Biol, 2006. **362**(1): p. 89-101.
85. Thorpe, C., et al., *Sulfhydryl oxidases: emerging catalysts of protein disulfide bond formation in eukaryotes.* Arch Biochem Biophys, 2002. **405**(1): p. 1-12.
86. Ang, S.K., et al., *Mitochondrial thiol oxidase Erv1: both shuttle cysteine residues are required for its function with distinct roles.* Biochem J, 2014. **460**(2): p. 199-210.
87. Guo, P.C., et al., *Structure of yeast sulfhydryl oxidase erv1 reveals electron transfer of the disulfide relay system in the mitochondrial intermembrane space.* J Biol Chem, 2012. **287**(42): p. 34961-9.
88. Levitan, A., A. Danon, and T. Lisowsky, *Unique features of plant mitochondrial sulfhydryl oxidase.* J Biol Chem, 2004. **279**(19): p. 20002-8.
89. Bach, R.D., O. Dmitrenko, and C. Thorpe, *Mechanism of thiolate-disulfide interchange reactions in biochemistry.* J Org Chem, 2008. **73**(1): p. 12-21.
90. Kodali, V.K. and C. Thorpe, *Oxidative protein folding and the Quiescin-sulfhydryl oxidase family of flavoproteins.* Antioxid Redox Signal, 2010. **13**(8): p. 1217-30.
91. Chivers, P.T., K.E. Prehoda, and R.T. Raines, *The CXXC motif: a rheostat in the active site.* Biochemistry, 1997. **36**(14): p. 4061-6.
92. Crooks, G.E., et al., *WebLogo: a sequence logo generator.* Genome Res, 2004. **14**(6): p. 1188-90.
93. Clayton, C.E., *Genetic manipulation of kinetoplastida.* Parasitol Today, 1999. **15**(9): p. 372-8.
94. Cruz, A. and S.M. Beverley, *Gene replacement in parasitic protozoa.* Nature, 1990. **348**(6297): p. 171-3.
95. Cruz, A., C.M. Coburn, and S.M. Beverley, *Double targeted gene replacement for creating null mutants.* Proc Natl Acad Sci U S A, 1991. **88**(16): p. 7170-4.
96. Bryant, J.M., et al., *CRISPR in Parasitology: Not Exactly Cut and Dried!* Trends Parasitol, 2019. **35**(6): p. 409-422.
97. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.* Trends Biotechnol, 2013. **31**(7): p. 397-405.

98. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
99. Mali, P., K.M. Esvelt, and G.M. Church, *Cas9 as a versatile tool for engineering biology.* Nat Methods, 2013. **10**(10): p. 957-63.
100. Ghorbal, M., et al., *Genome editing in the human malaria parasite Plasmodium falciparum using the CRISPR-Cas9 system.* Nat Biotechnol, 2014. **32**(8): p. 819-21.
101. Sander, J.D. and J.K. Joung, *CRISPR-Cas systems for editing, regulating and targeting genomes.* Nat Biotechnol, 2014. **32**(4): p. 347-55.
102. Horvath, P. and R. Barrangou, *CRISPR/Cas, the immune system of bacteria and archaea.* Science, 2010. **327**(5962): p. 167-70.
103. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
104. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.
105. Qi, L.S., et al., *Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.* Cell, 2013. **152**(5): p. 1173-83.
106. Peng, D., et al., *CRISPR-Cas9-mediated single-gene and gene family disruption in Trypanosoma cruzi.* mBio, 2014. **6**(1): p. e02097-14.
107. Zhang, W.W. and G. Matlashewski, *CRISPR-Cas9-Mediated Genome Editing in Leishmania donovani.* MBio, 2015. **6**(4): p. e00861.
108. Sollelis, L., et al., *First efficient CRISPR-Cas9-mediated genome editing in Leishmania parasites.* Cell Microbiol, 2015. **17**(10): p. 1405-12.
109. Beneke, T., et al., *A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids.* R Soc Open Sci, 2017. **4**(5): p. 170095.
110. Martel, D., et al., *Characterisation of Casein Kinase 1.1 in Leishmania donovani Using the CRISPR Cas9 Toolkit.* Biomed Res Int, 2017. **2017**: p. 4635605.
111. Costa, F.C., et al., *Expanding the toolbox for Trypanosoma cruzi: A parasite line incorporating a bioluminescence-fluorescence dual reporter and streamlined CRISPR/Cas9 functionality for rapid in vivo localisation and phenotyping.* PLoS Negl Trop Dis, 2018. **12**(4): p. e0006388.
112. Hartwell, L.H., et al., *Genetic Control of the Cell Division Cycle in Yeast: V. Genetic Analysis of cdc Mutants.* Genetics, 1973. **74**(2): p. 267-86.
113. Ben-Aroya, S., et al., *Toward a comprehensive temperature-sensitive mutant repository of the essential genes of Saccharomyces cerevisiae.* Mol Cell, 2008. **30**(2): p. 248-58.
114. Kraeva, N., et al., *Tetracycline-inducible gene expression system in Leishmania mexicana.* Mol Biochem Parasitol, 2014. **198**(1): p. 11-3.
115. Madeira da Silva, L., et al., *Regulated expression of the Leishmania major surface virulence factor lipophosphoglycan using conditionally destabilized fusion proteins.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7583-8.
116. Wheeler, R.J., E. Gluenz, and K. Gull, *Basal body multipotency and axonemal remodelling are two pathways to a 9+0 flagellum.* Nat Commun, 2015. **6**: p. 8964.
117. Murta, S.M., et al., *Methylene tetrahydrofolate dehydrogenase/cyclohydrolase and the synthesis of 10-CHO-THF are essential in Leishmania major.* Mol Microbiol, 2009. **71**(6): p. 1386-401.

118. McCall, L.I., et al., *Targeting Ergosterol biosynthesis in Leishmania donovani: essentiality of sterol 14 alpha-demethylase.* PLoS Negl Trop Dis, 2015. **9**(3): p. e0003588.
119. Robinson, K.A. and S.M. Beverley, *Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite Leishmania.* Mol Biochem Parasitol, 2003. **128**(2): p. 217-28.
120. Kolev, N.G., C. Tschudi, and E. Ullu, *RNA interference in protozoan parasites: achievements and challenges.* Eukaryot Cell, 2011. **10**(9): p. 1156-63.
121. Lye, L.F., et al., *Retention and loss of RNA interference pathways in trypanosomatid protozoans.* PLoS Pathog, 2010. **6**(10): p. e1001161.
122. Drinnenberg, I.A., et al., *RNAi in budding yeast.* Science, 2009. **326**(5952): p. 544-550.
123. Winkler, W.C., et al., *Control of gene expression by a natural metabolite-responsive ribozyme.* Nature, 2004. **428**(6980): p. 281-6.
124. Milewski, S., *Glucosamine-6-phosphate synthase--the multi-facets enzyme.* Biochim Biophys Acta, 2002. **1597**(2): p. 173-92.
125. Collins, J.A., et al., *Mechanism of mRNA destabilization by the glmS ribozyme.* Genes Dev, 2007. **21**(24): p. 3356-68.
126. Watson, P.Y. and M.J. Fedor, *The glmS riboswitch integrates signals from activating and inhibitory metabolites in vivo.* Nat Struct Mol Biol, 2011. **18**(3): p. 359-63.
127. Prommana, P., et al., *Inducible knockdown of Plasmodium gene expression using the glmS ribozyme.* PLoS One, 2013. **8**(8): p. e73783.
128. Jankowska-Döllken, M., et al., *Overexpression of the HECT ubiquitin ligase PfUT prolongs the intraerythrocytic cycle and reduces invasion efficiency of Plasmodium falciparum.* Sci Rep, 2019. **9**(1): p. 18333.
129. Cruz-Bustos, T., et al., *A Riboswitch-based Inducible Gene Expression System for Trypanosoma brucei.* J Eukaryot Microbiol, 2017.
130. Lander, N., T. Cruz-Bustos, and R. Docampo, *A CRISPR/Cas9-riboswitch-Based Method for Downregulation of Gene Expression in Trypanosoma cruzi.* Front Cell Infect Microbiol, 2020. **10**: p. 68.
131. Specht, S., et al., *A single-cysteine mutant and chimeras of essential Leishmania Erv can complement the loss of Erv1 but not of Mia40 in yeast.* Redox Biol, 2018. **15**: p. 363-374.
132. Birnboim, H.C. and J. Doly, *A rapid alkaline extraction procedure for screening recombinant plasmid DNA.* Nucleic Acids Res, 1979. **7**(6): p. 1513-23.
133. Slaymaker, I.M., et al., *Rationally engineered Cas9 nucleases with improved specificity.* Science, 2016. **351**(6268): p. 84-8.
134. LeBowitz, J.H., et al., *Development of a stable Leishmania expression vector and application to the study of parasite surface antigen genes.* Proc Natl Acad Sci U S A, 1990. **87**(24): p. 9736-40.
135. Jeacock, L., J. Faria, and D. Horn, *Codon usage bias controls mRNA and protein abundance in trypanosomatids.* Elife, 2018. **7**.
136. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
137. Link, A.J. and J. LaBaer, *Trichloroacetic acid (TCA) precipitation of proteins.* Cold Spring Harb Protoc, 2011. **2011**(8): p. 993-4.
138. Studier, F.W., *Protein production by auto-induction in high density shaking cultures.* Protein Expr Purif, 2005. **41**(1): p. 207-34.

139. Yao, M., et al., *An efficient strategy for the expression of Jingzhaotoxin-III in Escherichia coli.* Biotechnology & Biotechnological Equipment, 2017. **31**(4): p. 821-827.
140. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.
141. Dyballa, N. and S. Metzger, *Fast and sensitive colloidal coomassie G-250 staining for proteins in polyacrylamide gels.* J Vis Exp, 2009(30).
142. Chevallet, M., S. Luche, and T. Rabilloud, *Silver staining of proteins in polyacrylamide gels.* Nat Protoc, 2006. **1**(4): p. 1852-8.
143. Schumann Burkard, G., P. Jutzi, and I. Roditi, *Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters.* Mol Biochem Parasitol, 2011. **175**(1): p. 91-4.
144. Petrungaro, C., et al., *The Ca(2+)-Dependent Release of the Mia40-Induced MICU1-MICU2 Dimer from MCU Regulates Mitochondrial Ca(2+) Uptake.* Cell Metab, 2015. **22**(4): p. 721-33.
145. Kadokura, H., et al., *Identification of the redox partners of ERdj5/JPDI, a PDI family member, from an animal tissue.* Biochem Biophys Res Commun, 2013. **440**(2): p. 245-50.
146. Antoniou, A.N. and S.J. Powis, *Characterization of the ERp57-Tapasin complex by rapid cellular acidification and thiol modification.* Antioxid Redox Signal, 2003. **5**(4): p. 375-9.
147. Silverman, J.M., et al., *Proteomic analysis of the secretome of Leishmania donovani.* Genome Biol, 2008. **9**(2): p. R35.
148. Brigotti, M., et al., *4-Aminopyrazolo[3,4-d]pyrimidine (4-APP) as a novel inhibitor of the RNA and DNA depurination induced by Shiga toxin 1.* Nucleic Acids Res, 2000. **28**(12): p. 2383-8.
149. Iovannisci, D.M., et al., *Genetic analysis of adenine metabolism in Leishmania donovani promastigotes. Evidence for diploidy at the adenine phosphoribosyltransferase locus.* J Biol Chem, 1984. **259**(23): p. 14617-23.
150. Hwang, H.Y. and B. Ullman, *Genetic analysis of purine metabolism in Leishmania donovani.* J Biol Chem, 1997. **272**(31): p. 19488-96.
151. Hwang, H.Y., et al., *Creation of homozygous mutants of Leishmania donovani with single targeting constructs.* J Biol Chem, 1996. **271**(48): p. 30840-6.
152. Turra, G.L., et al., *Testing the CRISPR-Cas9 and glmS ribozyme systems in Leishmania tarentolae.* Preprint deposited in bioRxiv, 2020: p. 2020.07.23.218073.
153. Ralston, K.S., et al., *Flagellar motility contributes to cytokinesis in Trypanosoma brucei and is modulated by an evolutionarily conserved dynein regulatory system.* Eukaryot Cell, 2006. **5**(4): p. 696-711.
154. Straschil, U., et al., *The Armadillo repeat protein PF16 is essential for flagellar structure and function in Plasmodium male gametes.* PLoS One, 2010. **5**(9): p. e12901.
155. Koch, J.R. and F.X. Schmid, *Mia40 targets cysteines in a hydrophobic environment to direct oxidative protein folding in the mitochondria.* Nat Commun, 2014. **5**: p. 3041.
156. Warscheid, B., *Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC): Methods and Protocols*. 2014.
157. Vickers, T.J. and S.M. Beverley, *Folate metabolic pathways in Leishmania.* Essays Biochem, 2011. **51**: p. 63-80.

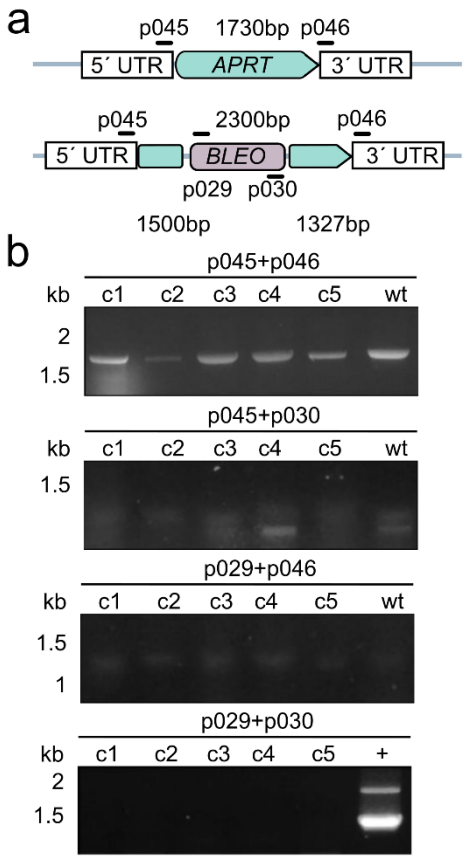
158. Carter, N.S., et al., *Purine and pyrimidine metabolism in Leishmania.* Adv Exp Med Biol, 2008. **625**: p. 141-54.
159. Brotherton, M.C., et al., *Quantitative proteomic analysis of amphotericin B resistance in Leishmania infantum.* Int J Parasitol Drugs Drug Resist, 2014. **4**(2): p. 126-32.
160. Peikert, C.D., et al., *Charting organellar importomes by quantitative mass spectrometry.* Nat Commun, 2017. **8**: p. 15272.
161. Fukasawa, Y., et al., *MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites.* Mol Cell Proteomics, 2015. **14**(4): p. 1113-26.
162. Depuydt, M., J. Messens, and J.F. Collet, *How proteins form disulfide bonds.* Antioxid Redox Signal, 2011. **15**(1): p. 49-66.
163. Denoncin, K. and J.F. Collet, *Disulfide bond formation in the bacterial periplasm: major achievements and challenges ahead.* Antioxid Redox Signal, 2013. **19**(1): p. 63-71.
164. Frand, A.R., J.W. Cuozzo, and C.A. Kaiser, *Pathways for protein disulphide bond formation.* Trends Cell Biol, 2000. **10**(5): p. 203-10.
165. Meyer, A.J., J. Riemer, and N. Rouhier, *Oxidative protein folding: state-of-the-art and current avenues of research in plants.* New Phytol, 2019. **221**(3): p. 1230-1246.
166. Liedgens, L., *Investigation of oxidative protein folding in protist mitochondria and elucidation of the catalytic mechanism of glutaredoxins*. 2019, Dissertation, Ruperto Carola University of Heidelberg, 2018: Heidelberg. p. 1 Online-Ressource (XI, 156 Seiten).
167. Sideris, D.P., et al., *A novel intermembrane space-targeting signal docks cysteines onto Mia40 during mitochondrial oxidative folding.* J Cell Biol, 2009. **187**(7): p. 1007-22.
168. Banci, L., et al., *Molecular recognition and substrate mimicry drive the electron-transfer process between MIA40 and ALR.* Proc Natl Acad Sci U S A, 2011. **108**(12): p. 4811-6.
169. Peleh, V., et al., *Erv1 of Arabidopsis thaliana can directly oxidize mitochondrial intermembrane space proteins in the absence of redox-active Mia40.* BMC Biol, 2017. **15**(1): p. 106.
170. Mallo, N., et al., *Protein Import into the Endosymbiotic Organelles of Apicomplexan Parasites.* Genes (Basel), 2018. **9**(8).
171. Lamour, N., et al., *Proline metabolism in procyclic Trypanosoma brucei is down-regulated in the presence of glucose.* J Biol Chem, 2005. **280**(12): p. 11902-10.
172. Eichenberger, C., et al., *The highly diverged trypanosomal MICOS complex is organized in a nonessential integral membrane and an essential peripheral module.* Mol Microbiol, 2019. **112**(6): p. 1731-1743.
173. Farrell, S.R. and C. Thorpe, *Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome c reductase activity.* Biochemistry, 2005. **44**(5): p. 1532-41.
174. Kallergi, E., et al., *Targeting and maturation of Erv1/ALR in the mitochondrial intermembrane space.* ACS Chem Biol, 2012. **7**(4): p. 707-14.
175. Sidik, S.M., et al., *A Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes.* Cell, 2016. **166**(6): p. 1423-1435.e12.
176. Shalem, O., et al., *Genome-scale CRISPR-Cas9 knockout screening in human cells.* Science, 2014. **343**(6166): p. 84-87.
177. Chen, X., J.L. Zaro, and W.C. Shen, *Fusion protein linkers: property, design and functionality.* Adv Drug Deliv Rev, 2013. **65**(10): p. 1357-69.

178. Hung, V., et al., *Proteomic mapping of the human mitochondrial intermembrane space in live cells via ratiometric APEX tagging.* Mol Cell, 2014. **55**(2): p. 332-41.
179. Gleason, F.K., *Mutation of conserved residues in Escherichia coli thioredoxin: effects on stability and function.* Protein Sci, 1992. **1**(5): p. 609-16.
180. Stopa, J.D., et al., *Kinetic-based trapping by intervening sequence variants of the active sites of protein-disulfide isomerase identifies platelet protein substrates.* J Biol Chem, 2017. **292**(22): p. 9063-9074.
181. Carrie, C. and J. Soll, *To Mia or not to Mia: stepwise evolution of the mitochondrial intermembrane space disulfide relay.* BMC Biol, 2017. **15**(1): p. 119.
182. Haindrich, A.C., et al., *The intermembrane space protein Erv1 of Trypanosoma brucei is essential for mitochondrial Fe-S cluster assembly and operates alone.* Mol Biochem Parasitol, 2017. **214**: p. 47-51.
183. Carrie, C., et al., *Conserved and novel functions for Arabidopsis thaliana MIA40 in assembly of proteins in mitochondria and peroxisomes.* J Biol Chem, 2010. **285**(46): p. 36138-48.
184. Ilani, T., et al., *A secreted disulfide catalyst controls extracellular matrix composition and function.* Science, 2013. **341**(6141): p. 74-6.
185. Castro, H., et al., *Leishmania mitochondrial peroxiredoxin plays a crucial peroxidase-unrelated role during infection: insight into its novel chaperone activity.* PLoS Pathog, 2011. **7**(10): p. e1002325.
186. Smith, J.T., Jr., et al., *Divergent Small Tim Homologues Are Associated with TbTim17 and Critical for the Biogenesis of TbTim17 Protein Complexes in Trypanosoma brucei.* mSphere, 2018. **3**(3).
187. Alsford, S., et al., *High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome.* Genome Res, 2011. **21**(6): p. 915-24.
188. Koch, J.R. and F.X. Schmid, *Mia40 is optimized for function in mitochondrial oxidative protein folding and import.* ACS Chem Biol, 2014. **9**(9): p. 2049-57.
189. Kawamata, H. and G. Manfredi, *Import, maturation, and function of SOD1 and its copper chaperone CCS in the mitochondrial intermembrane space.* Antioxid Redox Signal, 2010. **13**(9): p. 1375-84.
190. Bien, M., et al., *Mitochondrial disulfide bond formation is driven by intersubunit electron transfer in Erv1 and proofread by glutathione.* Mol Cell, 2010. **37**(4): p. 516-28.
191. Ceh-Pavia, E., et al., *Redox characterisation of Erv1, a key component for protein import and folding in yeast mitochondria.* Febs j, 2020. **287**(11): p. 2281-2291.
192. Hammond, M.J., et al., *A Uniquely Complex Mitochondrial Proteome from Euglena gracilis.* Mol Biol Evol, 2020. **37**(8): p. 2173-2191.
193. Alcântara, L.M., et al., *Challenges in drug discovery targeting TriTryp diseases with an emphasis on leishmaniasis.* Int J Parasitol Drugs Drug Resist, 2018. **8**(3): p. 430-439.
194. Osorio-Méndez, J.F. and A.M. Cevallos, *Discovery and Genetic Validation of Chemotherapeutic Targets for Chagas' Disease.* Front Cell Infect Microbiol, 2018. **8**: p. 439.
195. Altamura, F., et al., *The current drug discovery landscape for trypanosomiasis and leishmaniasis: Challenges and strategies to identify drug targets.* Drug Dev Res, 2020.
196. Barrett, M.P., J.C. Mottram, and G.H. Coombs, *Recent advances in identifying and validating drug targets in trypanosomes and leishmanias.* Trends Microbiol, 1999. **7**(2): p. 82-8.

197. Comini, M.A., et al., *Validation of Trypanosoma brucei trypanothione synthetase as drug target.* Free Radic Biol Med, 2004. **36**(10): p. 1289-302.
198. Djuika, C.F., et al., *Knockout of the peroxiredoxin 5 homologue PFAOP does not affect the artemisinin susceptibility of Plasmodium falciparum.* Sci Rep, 2017. **7**(1): p. 4410.
199. Wezena, C.A., et al., *The cytosolic glyoxalases of Plasmodium falciparum are dispensable during asexual blood-stage development.* Microb Cell, 2017. **5**(1): p. 32-41.
200. Zhang, M., et al., *Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis.* Science, 2018. **360**(6388).
201. Duncan, S.M., N.G. Jones, and J.C. Mottram, *Recent advances in Leishmania reverse genetics: Manipulating a manipulative parasite.* Mol Biochem Parasitol, 2017. **216**: p. 30-38.
202. LeBowitz, J.H., A. Cruz, and S.M. Beverley, *Thymidine kinase as a negative selectable marker in Leishmania major.* Mol Biochem Parasitol, 1992. **51**(2): p. 321-5.
203. Davoudi, N., et al., *Development of a recombinant Leishmania major strain sensitive to ganciclovir and 5-fluorocytosine for use as a live vaccine challenge in clinical trials.* Vaccine, 2005. **23**(9): p. 1170-7.
204. Boeke, J.D., et al., *5-Fluoroorotic acid as a selective agent in yeast molecular genetics.* Methods Enzymol, 1987. **154**: p. 164-75.
205. Erbs, P., et al., *In vivo cancer gene therapy by adenovirus-mediated transfer of a bifunctional yeast cytosine deaminase/uracil phosphoribosyltransferase fusion gene.* Cancer Res, 2000. **60**(14): p. 3813-22.
206. Chung-Faye, G.A., et al., *In vivo gene therapy for colon cancer using adenovirus-mediated, transfer of the fusion gene cytosine deaminase and uracil phosphoribosyltransferase.* Gene Ther, 2001. **8**(20): p. 1547-54.
207. Braks, J.A., et al., *Development and application of a positive-negative selectable marker system for use in reverse genetics in Plasmodium.* Nucleic Acids Res, 2006. **34**(5): p. e39.
208. Duraisingh, M.T., T. Triglia, and A.F. Cowman, *Negative selection of Plasmodium falciparum reveals targeted gene deletion by double crossover recombination.* Int J Parasitol, 2002. **32**(1): p. 81-9.
209. Cruz-Bustos, T., et al., *A Riboswitch-based Inducible Gene Expression System for Trypanosoma brucei.* J Eukaryot Microbiol, 2018. **65**(3): p. 412-421.
210. Ferré-D'Amaré, A.R., *The glmS ribozyme: use of a small molecule coenzyme by a gene-regulatory RNA.* Q Rev Biophys, 2010. **43**(4): p. 423-47.
211. Naderer, T., E. Wee, and M.J. McConville, *Role of hexosamine biosynthesis in Leishmania growth and virulence.* Mol Microbiol, 2008. **69**(4): p. 858-69.
212. Lander, N., M.A. Chiurillo, and R. Docampo, *Genome Editing by CRISPR/Cas9: A Game Change in the Genetic Manipulation of Protists.* J Eukaryot Microbiol, 2016. **63**(5): p. 679-90.
213. Ren, B. and N. Gupta, *Taming Parasites by Tailoring Them.* Front Cell Infect Microbiol, 2017. **7**: p. 292.
214. Zhang, W.W., P. Lypaczewski, and G. Matlashewski, *Optimized CRISPR-Cas9 Genome Editing for Leishmania and Its Use To Target a Multigene Family, Induce Chromosomal Translocation, and Study DNA Break Repair Mechanisms.* mSphere, 2017. **2**(1).

215. Eckers, E. and M. Deponte, *No need for labels: the autofluorescence of Leishmania tarentolae mitochondria and the necessity of negative controls.* PLoS One, 2012. **7**(10): p. e47641.
216. Cruz, A.K., R. Titus, and S.M. Beverley, *Plasticity in chromosome number and testing of essential genes in Leishmania by targeting.* Proc Natl Acad Sci U S A, 1993. **90**(4): p. 1599-603.
217. Lander, N., et al., *CRISPR/Cas9-Induced Disruption of Paraflagellar Rod Protein 1 and 2 Genes in Trypanosoma cruzi Reveals Their Role in Flagellar Attachment.* mBio, 2015. **6**(4): p. e01012.
218. Janssen, B.D., et al., *CRISPR/Cas9-mediated gene modification and gene knock out in the human-infective parasite Trichomonas vaginalis.* Sci Rep, 2018. **8**(1): p. 270.
219. Pines, G. and R.T. Gill, *Dynamic Management of Codon Compression for Saturation Mutagenesis.* Methods Mol Biol, 2018. **1772**: p. 171-189.

7 Supplementary information



Supplementary figure 7.1 Attempts to generate *APRT* knockout with CRISPR/Cas9 plasmid-encoded strategy.
(a) Schematic of the wild-type *APRT* locus and expected knockout mutation by targeting the gene with a bleomycin resistant cassette. Primer binding sites are depicted in black lines with expected PCR product sizes **(b)** Analytical PCR visualizing the presence of the wild type locus in bleomycin resistant lines despite the failure of amplification of the bleomycin resistance gene (PCR p029+p030). + symbolizes the use of a plasmid bearing the bleomycin resistance gene as a positive control.