Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of the Ruperto Carola University Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

Presented by

M.Sc. Kathrin Bajak

Born in: Heilbronn, Germany

Oral examination: 12th of July 2019

Networks of gene expression control in *Trypanosoma brucei*

Referees: Prof. Dr. Christine Clayton

Prof. Dr. Dr. Georg Stoecklin

Acknowledgment

I would like to kindly thank

Prof. Dr. Christine Clayton for giving me the opportunity to work in her laboratory, for great scientific input, thoughtful guidance and many inspiring discussions.

Dr. Esteban Erben for his supervision, numerous fruitful, critical as well as inspiring discussions, for technical help, for proofreading this thesis, for your belief in my work and all your support.

Prof. Dr. Nina Papavasiliou and **Dr. Erec Stebbins** for giving me the opportunity to work in their lab for the last year of my PhD.

Prof. Dr. Luise Krauth-Siegel and **Prof. Dr. Dr. Georg Stoecklin** for scientific support and inspiring discussions in the TAC meetings.

Prof. Dr. Dr. Georg Stoecklin and **Dr. Susanne Kramer** for providing cell lines, reagents and equipment.

Dr. Esteban Erben and **Kevin Leiss** for their research on the ZC3H5 project on which I could built up my work.

Pia Hartwig for her assistance and contribution in the BFR1L project during her lab rotation.

The whole **Papavasiliou/Stebbins lab** for the friendly atmosphere and the great time that we shared in and outside the lab.

Dr. Thomas Ruppert and **Sabine Merker** from the Core Facility for Mass Spectrometry & Proteomics at the ZMBH and **David Ibberson** from the Cell Networks Deep Sequencing Core Facility at the University of Heidelberg for useful discussions and analysis of our samples.

Dr. Holger Lorenz and **Christian Hoerth** from the Imaging facility at the ZMBH for their help with microscopy and image analysis and **Tobias Rubner** from the FACS facility at the DKFZ for his help and support.

Ute and **Claudia** from the Clayton lab for the technical support, which made my live in the lab much easier. **Larissa Melo Do Nascimento** from the Clayton lab for all the scientific and non-scientific discussions. And **Daniela Begolo** for your great friendship, it was a pleasure having you around in the lab and spending fun leisure time together.

Katharina Haneke for your friendship, endless support and advise during the PhD time.

To all my friends for their support, motivation and the great time we spent. Girls from riding stable for distracting me from the lab work and spending time with the horses; friends from university that we are still in contact, even if we are spread around the world; friends from home for the amazing weekend times.

My parents and **my sister** for their patience, care, love and support. For all the lovely family time we spent together. For their belief in me and my work; for encouraging and motivating me.

Table of contents

List of abbreviations

abbreviation	description
3'UTR	3' untranslated region
5'TOP	5’ Terminal Oligo Pyrimidine
5'UTR	5' untranslated region
AD	acitvating domain
ALBA	acetylation lowers binding affinity
ALPH1	ApaH-like phosphatase 1
BD	binding-domain
BF	bloodstream form
BFA	Brefeldin A
BFR1L	Bfr1-like
Bfr1p	Brefeldin A resistance protein
BiP	binding immunoglobulin protein
BSD	Blasticidin-S deaminase
CAT	chloramphenicol acetyltransferase
CBC	cap-binding complex
CDK	cyclin-dependent kinase
cDKO	conditional double knockout
CDS	coding sequence
cSKO	conditional single knockout
DKO	double knockout
DNA	deoxyribonucleic acid
ER	Endoplasmic reticulum
ES	Expression site
ESAG8	expression site-associated gene 8
FACS	fluorescent activated cell sorting
FAZ	flagellar attachment zone
FISH	fluorescence *in situ* hybridization
gDNA	genomic DNA
G-protein	guanine nucleotide-binding protein
IF microscopy	immunofluorescence microscopy
kDNA	kinetoplast DNA
LC/MS/MS	Liquid chromatography-tandem mass spectrometry
LipDH	Lipoamide dehydrogenase
M	Mitosis
MAPK	mitogen-activated kinase
mRNA	messenger RNA
mRNP	messenger ribonucleoprotein
MS	mass spectrometry
ORF	open reading frame
PABP	poly(A)-binding protein
PAC	Puromycin N-acetyltransferase
PAG	procyclin associated gene
PAG	Procyclin-associated gene
PC	procyclic form
PCR	polymerase chain reaction

PLK	polo-like kinase
qPCR	quantitative PCR
RACK1	receptor for activated C kinase 1
RBP	RNA-binding protein
RIP-Seq	RNA affinity purification followed by next-generation sequencing
RIT-Seq	RNA interference target sequencing
RNA	ribonucleid acid
RNAi	RNA interference
RRM	RNA recognition motif
rRNA	ribosomal RNA
RT-qPCR	reverse transcription quantitative PCR
SG	stress granule
Sk	Kinetoplast S-phase
SKO	single knockout
SL	spliced-leader
Sn	Nuclear S-phase
SSR	strand switch region
Suppl.	supplementary
TAC	tripartite attachment complex
TAP	Tandem affinity purification
TCA	Trichloroacetic acid
Tet	Tetracycline
TR	Trypanothione Reductase
TRACK	trypanosome receptor for activated C kinase
VSG	variant surface glycoprotein
WT	wild type
Y2H	yeast-two-hybrid
YFP	yellow fluorescent protein
ZPFM	Zimmerman's Post Fusion Medium

List of figures

List of tables

Summary

Since in trypanosomes most protein-coding genes are constitutively transcribed by RNA polymerase II in a polycistronic manner, gene expression is mainly regulated at the post-transcriptional level. It is therefore interesting to investigate the relevant regulatory factors. In a previous genome-wide tethering screen hundreds of putative mRNA-fate regulators were found, including the proteins BFR1L, an up-regulator of gene expression, and ZC3H5, a down-regulator of gene expression.

BFR1L has some similarities to yeast Bfr1p, an ER- and polysome-associated protein. BFR1L displays *in vivo* mRNA binding although it lacks canonical RNA-binding domains. Double-knockout bloodstream form trypanosomes displayed a slight growth defect. By immunofluorescence microscopy, a tagged version was located in the cytoplasm and overlapped partially with an ER marker. RNA pull-down analysis suggested that most of the BFR1L-bound mRNAs encode ribosomal proteins, but no common RNA motif could be found by *in silico* analysis. The mRNAs encoding ribosomal proteins are known not to sequester in granules upon starvation stress or heat shock. Similarly, BFR1L protein did not go to granules under starvation stress. It is tempting to speculate that the interaction remains active during stress, and targeting of the mRNAs to the ER could prevent sequestration into granules and could keep the mRNAs in ribosomes. The attachment of BFR1L to the ER could be mediated via the putative interaction partner Tb927.9.9550, which has a transmembrane domain.

ZC3H5 knock-down led to a fast growth defect, killing the cells after 48 h of RNAi induction. We therefore analyzed the RNAi effect on growth kinetics, protein levels, nuclei/kinetoplasts ratios and the transcriptome at different time points. After RNAi induction, the proportion of 2N2K cells increased rapidly. In addition, the ZC3H5 RNAi cells often possessed abnormal and higher numbers of nuclei and kinetoplast. While short-term down-regulation of ZC3H5 showed only a minor effect with respect to the transcriptome, an increase of mRNAs encoding ribosomal proteins, the increase of the monosomal peak without an increase of mRNAs in this fraction, the occurrence of half-mers as well as an increase of mRNAs encoding ribosomal proteins in the free fraction were observed with respect to the polysomal profiles. This suggest that the ribosome assembly is disturbed upon knock-down of ZC3H5. However, this seems to be a secondary effect. RNA pull-down analysis suggested that ZC3H5 binds mRNAs encode cytoskeleton proteins. Tandem Affinity Purification of ZC3H5 followed by MS analysis revealed three putative interaction partners which were validated by co-immunoprecipitation (Tb927.8.1500, Tb927.7.3040 and Tb927.11.4900). In addition, tethering of ZC3H5 and its interaction partners to a CAT reporter showed that the proteins are repressors; thus, we have identified a novel repressor complex that may regulate genes required for cell cycle progression. The exact mechanism of action is not known at the moment, but it is tempting to speculate that the function of Tb927.11.4900 as a G protein is responsible for the association and dissociation of the complex. This could be cell cycle dependent, because the target mRNAs peak in S-phase. Maybe the ZC3H5 complex represses its targets during the rest of the cell cycle and targets are de-repressed in S-phase to produce the proteins needed for cytokinesis.

Zusammenfassung

In Trypanosomen werden die meisten Protein-kodierenden Gene konstitutiv durch die RNA Polymerase II transkribiert. Deswegen wird die Genexpression hauptsächlich durch post-transkriptionelle Mechanismen reguliert, was die Untersuchung relevanter regulatorischer Faktoren interessant macht. In einem vorangegangenen „tethering screen“ wurden hunderte mögliche Regulatoren identifiziert, einschließlich BFR1L, welches die Gen-expression hochreguliert, und ZC3H5, welches die Geneexpression herunterreguliert.

BFR1L hat Ähnlichkeiten zu dem Hefeprotein Bfr1p, welches mit dem ER und Polysomen assoziiert. BFR1L kann *in vivo* an mRNA binden, obwohl es keine kanonische RNA-bindende Domäne besitzt. Das Ausschalten von BFR1 in der Blutstromform von Trypanosomen führt zu einem leichten Wachstumsdefekt. Per Immunfluoreszenz-mikroskopie konnte gezeigt werden, dass sich BFR1L im Zytoplasma befindet und teilweise mit einem ER-Marker überlappt. RNA Pull-Down Analysen deuten darauf hin, dass BFR1L hauptsächlich mit mRNAs, welche für ribosomale Proteine kodieren, interagiert. Mittels *in silico*-Analyse konnte jedoch kein übereinstimmendes RNA-Motiv unter den präferenziell gebundenen mRNAs gefunden werden. mRNAs, die für ribosomale Proteine kodieren, sind bekannt dafür, dass sie nicht in Stress- oder Hitzeschock-induzierten Granula akkumulieren. Ebenso befindet sich BFR1L unter Stressbedingungen nicht in diesen Granula. Daher wird vermutet, dass die Interaktion zwischen BFR1L und seinen interagierenden mRNAs auch unter Stressbedingungen aktiv bleibt, BFR1L die mRNAs zum ER rekrutiert und somit die Akkumulierung der mRNAs in Stressgranula verhindern könnte. Die Bindung von BFR1L zum ER könnte durch den möglichen Interaktionspartner Tb927.9.9550, welcher eine Transmembrandomäne hat, vermittelt werden.

Das Herunterregulieren von ZC3H5 führte zu einem rapiden Wachstumsdefekt, was die Zellen innerhalb von 48 Stunden tötet. Es konnte festgestellt werden, dass die Menge an Zellen mit 2 Nuklei und 2 Kinetoplasten nach der Herunterregulierung rapide anstieg und die Zellen oft generell eine erhöhte Anzahl an Nuklei und Kinetoplasten besaßen. Das Transkriptom zeigte bei kurzzeitiger Herunterregulierung von ZC3H5 einen Anstieg von mRNAs, welche für ribosomale Protein kodieren, und wurde ansonsten kaum beeinflusst. Hinsichtlich des Polysomen Profils konnte ein Anstieg der Monosomendichte, jedoch ohne gleichzeitigen Anstieg der mRNA Menge in dieser Fraktion, das Auftreten von „half-mers“ und dem Anstieg von mRNAs, welche für ribosomale Proteine kodieren, in der freien Fraktion beobachtet werden. Dies lässt vermuten, dass es ich um einen Defekt in der Zu-sammensetzung von Ribosomen handelt, was vermutlich ein sekundärer Effekt ist. Eine RNA-Pull-Down Analyse zeigte, dass ZC3H5 mit mRNAs interagiert, welche für Proteine des Zytoskeletts kodieren. Tandem-Affinitätsaufreinigung gefolgt von massenspektro-metrischer Analyse konnte eine Interaktion von ZC3H5 mit drei potentiellen Proteinen zeigen (Tb927.8.1500, Tb927.7.3040 und Tb927.11.4900). Außerdem konnte gezeigt werden, dass alle vier Proteine die Genexpression eines Reporters herunterregulieren. Dies zeigt, dass ein neuer repressiver Komplex identifiziert wurde, welcher eventuell Gene reguliert, die eine Rolle im Fortschreiten des Zellzyklus spielen. Es wird vermutet, dass die Funktion von Tb927.11.4900 als G-Protein für die Assoziation und Dissoziation des Komplexes verantwortlich ist. Dies könnte abhängig vom Zellzyklus passieren, da die mRNAs, welche mit ZC3H5 interagieren, vermehrt in der S-Phase zu finden sind. Eventuell hemmt der ZC3H5-Komplex die mRNAs während des restlichen Zellzyklus, doch während der S-Phase, wenn der Komplex dissoziiert, können die mRNAs translatiert werden und die Proteine, welche für die Zytokinese benötigt werden, werden produziert.

1. Introduction

1.1. Trypanosomes- disease and model organism

Trypanosomes are single-celled eukaryotes, which belong to the order *Kinetoplastida*. They are spindle-shaped cells with a size from 8 to over 50 µm and a diameter of 1.5-3.5 µm depending on the life cycle stage (Uilenberg, 1998). The procyclic form found in the midgut of the Tsetse fly vector and the bloodstream form multiplies in mammalian blood and tissue fluids. The African trypanosomes, *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense*, can cause the human disease African sleeping sickness (trypanosomiasis), which is transmitted by the Tsetse fly. It is estimated that 55 million people are at risk of the disease in sub-Saharan Africa, since the Tsetse fly infests 350000 km^2 of the landmass of Africa (Franco et al., 2017). However, the number of reported annually HAT cases decreased rapidly in the last decade with only 3797 new cases in 2014 (Franco et al., 2017). If sleeping sickness is not or inadequately treated, it is usually lethal and is thereby a cause of morbidity and mortality in sub-Saharan Africa. However, for treating the disease only a few drugs are known. Some of them suffer from poor efficacy or contain toxic arsenic derivates leading to sever side effects (Field et al., 2017). In cattle, trypanosomiasis leads to a decrease in meat and milk production. Furthermore, the animals cannot act as draught animals in crop cultivation anymore.

Besides the relevance of Trypanosomes as pathogenic parasites, they are an interesting model organism. They diverged early from the Opisthokonts, sharing the common features of eukaryotes, but also having special features (Adl et al., 2019). They have all the conventional organelles of a eukaryotic cell, like a nucleus, lysosomes, endo- and exocytosis system, Golgi and endoplasmic reticulum (ER) (Figure 1.1) (Clayton et al., 1995). One peculiarity is the single mitochondrion with the mitochondrial genome that has a disc like structure composed of maxi- and minicircles (Shapiro and Englund, 1995; Simpson, 1987). This structure is called the kinetoplast, which gave the name to the Kinetoplastida. Interestingly, the mitochondrial RNA encoded in the maxicircles is edited by insertions and deletions of uridines by the editosome (Liu et al., 2005; Panigrahi et al., 2006). However, the function of the mitochondrion differs in the two life cycle stages. In the bloodstream form, where the cells are in an environment with high glucose, the function of the mitochondrion is repressed, because the energy is mostly obtained by glycolysis. Most of the glycolytic enzymes can be found in another specific feature: the glycosomes (Michels et al., 2006; Opperdoes et al., 1984). In contrast to that, the procyclic form obtains the energy by oxidative phosphorylation of amino acids (Bringaud et al., 2006). Another peculiarity is the single flagellum that runs along the trypanosome. It exits near the posterior end from the flagellar pocket, which is in addition the place of endo- and exocytosis. The trypanosomes can move with the help of the flagellum, which is crucial for their viability

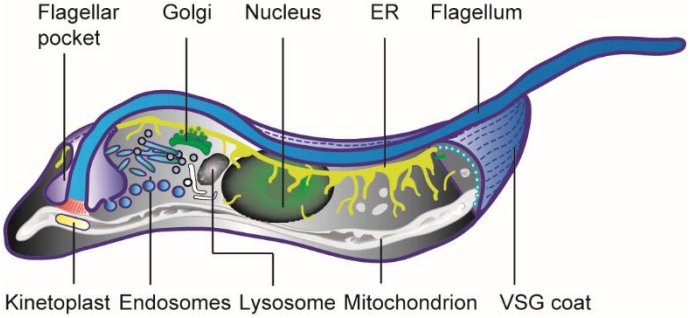


Figure 1.1: The structure and morphology of *T. brucei*. Generalized cellular structure of the bloodstream form. Figure taken from: (Overath and Engstler, 2004).

(Broadhead et al., 2006). In addition, the flagellum plays a role in the attachment to the host surface and in the morphogenesis and cytokinesis (Kohl et al., 2003). Trypanosomes major surface proteins make up approximately 10% of the total protein content of the cell and are GPI-anchored. In the bloodstream form the variant surface glycoprotein (VSG) is expressed. The expression of VSGs undergoes antigenic variation, which ensures that only one VSG isoform is expressed at one time. This switching helps the cells to evade from the host immune response (Hovel-Miner et al., 2015).

1.2. The life cycle of *Trypanosoma brucei*

The whole life cycle is illustrated in figure 1.2 (Langousis and Hill, 2014). By having a blood meal on an infected host, the Tsetse fly can take up trypanosomes. The non-dividing short stumpy form is pre-adapted for the survival in the midgut of the fly, where it differentiates into the procyclic form (Matthews, 1999). This transformation includes many changes: in cell metabolism from glycolysis to oxidative phosphorylation (Matthews, 2005), in cytoskeleton architecture, functioning of the organelles and switching of the surface proteins from VSG to procyclins (Ziegelbauer and Overath, 1990). The procyclic trypanosomes can then migrate along the foregut to the proventriculus where they divide asymmetrically to make long and short epimastigotes (Figure 1.2 B). In this stage the parasites express the surface protein BARP (Ziegelbauer and Overath, 1990). The short epimastigotes will then migrate to the salivary glands and differentiate into non-dividing metacyclic trypomastigotes with a VSG coat on their surface. This form can then by transmitted to the mammalian host by another bloodmeal of the Tsetse fly (Matthews, 2005). In the tissue of the mammalian host, the metacyclic trypomastigotes differentiate into long slender trypomastigotes, which multiply and start to invade the lymph nodes through the lymphatic system. From there they establish the bloodstream infection. As soon as the levels of parasitemia increase a quorum sensing mechanism triggers the differentiation of long slender bloodstream form to the short stumpy form (Matthews, 2005).

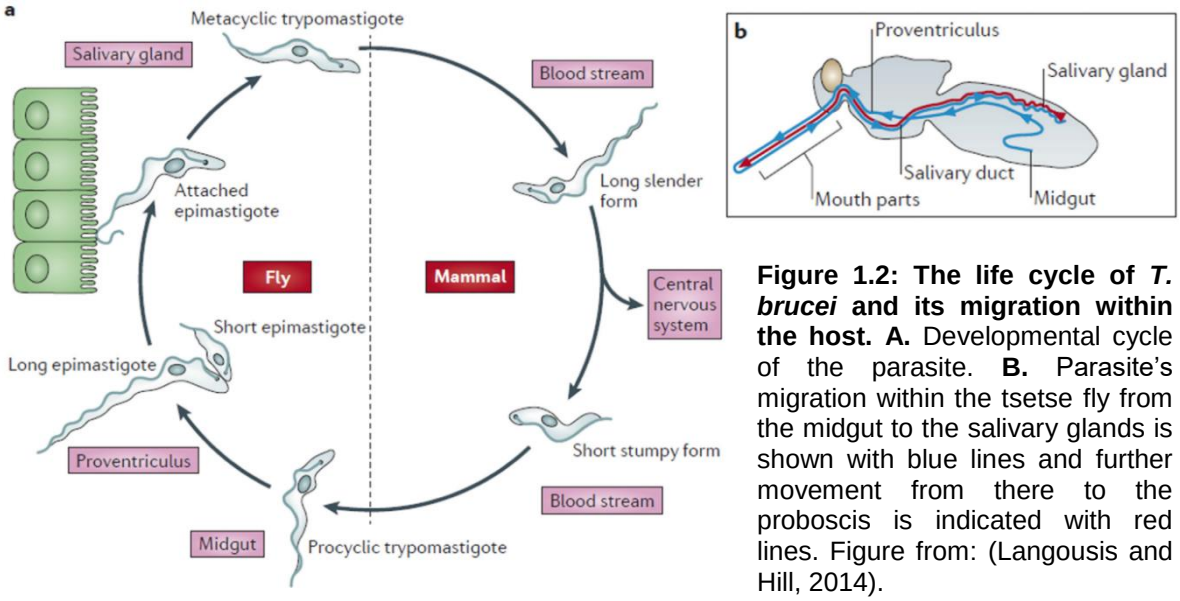


Figure 1.2: The life cycle of *T. brucei* and its migration within the host. A. Developmental cycle of the parasite. **B.** Parasite's migration within the tsetse fly from the midgut to the salivary glands is shown with blue lines and further movement from there to the proboscis is indicated with red lines. Figure from: (Langousis and Hill, 2014).

1.3. The cell cycle in *T. brucei* and its regulation

A typical cell cycle of eukaryotes consists of four phases: G_0/G_1, S, G_2 and M. In the first gap phase (G_0/G_1) the cells are getting prepared for entry into a new round of replication and cell division. In the S-phase the DNA is replicated and after the second gap-phase (G_2), the cells are dividing during the M-phase. The cell cycle of Trypanosomes follows this scheme, but it has some unique features and requirements (McKean, 2003). Trypanosomes have distinctive structures and organelles, which must be duplicated and segregated accurately during cell division, which happens in a precise order. At first, the basal body is elongated and maturated and the new flagellum is nucleated, which depends on the *de novo* recruitment of γ-tubulin (McKean et al., 2003). In the S-phase, the DNAs of the nuclear and kinetoplast genomes are replicated separately (Woodward and Gull, 1990). Kinetoplast S-phase (Sk) occurs before the nuclear S-phase (Sn) and the segregation of the kinetoplast (D) is completed before onset of mitosis (M) (Figure 1.3) (McKean, 2003). These separate events of kinetoplast and nuclear S-phase make it possible to determine the defined cell cycle stage of trypanosomes by staining of the DNA (e.g. with DAPI). At the beginning of

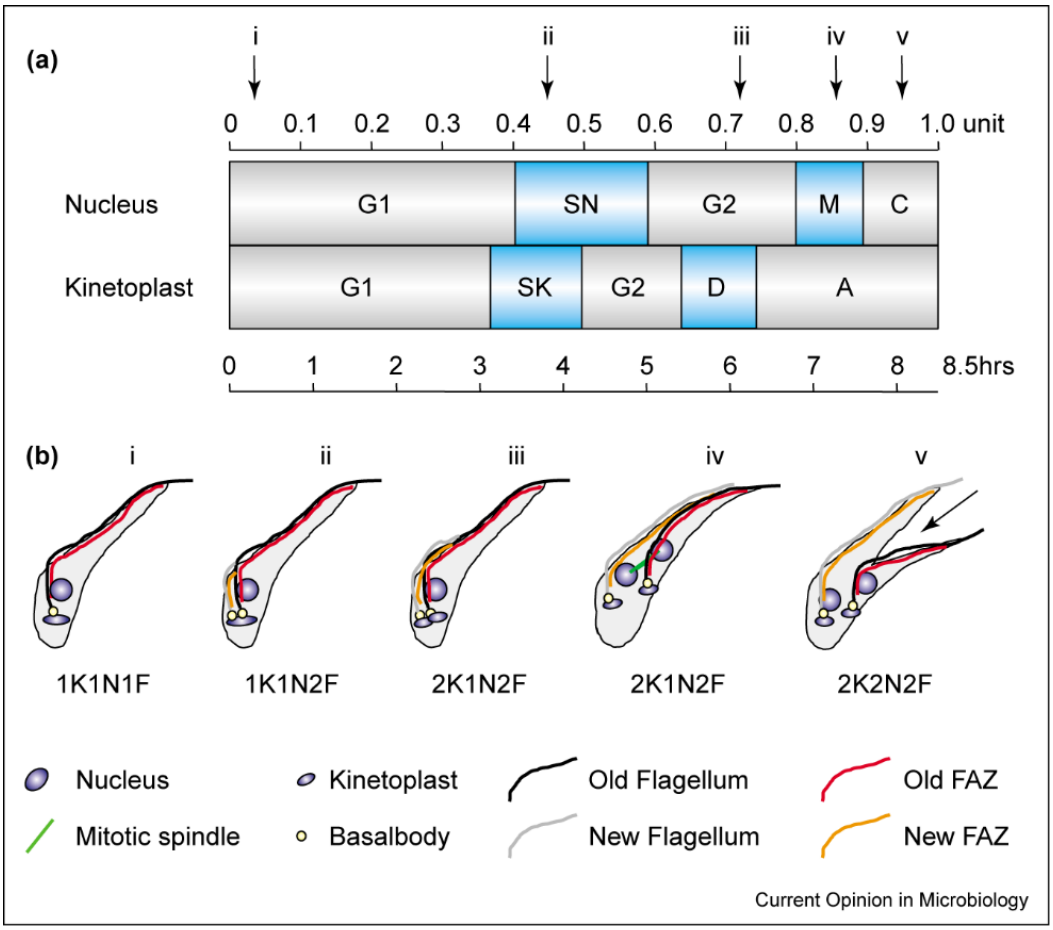


Figure 1.3: Cell Cycle of *T. brucei*. A. The trypanosome cell cycle is separated into nuclear and kinetoplast components. Cell cycle duration for exponentially growing procyclic trypanosomes is 8.5 h. Kinetoplast replication (S) initiates before nuclear S phase, but is considerably shorter and consequently kinetoplast segregation (D) occurs before the onset of nuclear mitosis (M). The phase annotated on the kinetoplast cycle as 'A' refers to the 'apportioning' phase during which basal bodies continue to move apart. **B.** Schematic representations of trypanosome cells taken from various time points through the cell cycle. The black arrow indicates the direction and position of the cleavage furrow. Figure and description from (McKean, 2003).

the G_2-phase of the nuclear cycle, the basal bodies separate in a microtubule-dependent manner (Robinson and Gull, 1991). The kinetoplast is connected with the proximal end of the basal body by a tripartite attachment complex (TAC), which ensures that the movement of the basal body results in the segregation of the replicated kinetoplast DNA (kDNA) (Ogbadoyi et al., 2003). The mitosis in trypanosomes is a closed process meaning that the nuclear envelope is not disrupted and the spindle is formed intra-nuclearly (Ogbadoyi et al., 2000). Finally, the trypanosome is divided into daughter cells along the longitudinal axis, where the cleavage furrow is formed. The ingression proceeds unidirectionally from the anterior to the posterior pole between the two flagella to form daughter cells. It has been proposed that the structural information, which is required for positioning of the cleavage furrow is provided by the flagellum attachment zone (FAZ) (Robinson et al., 1995).

All steps of the cell cycle in *Trypanosoma* are regulated by proteins, that include cyclin-dependent kinases (CDKs), mitogen-activated kinases (MAPK), aurora kinases and polo-like kinases (PLKs). However, the kinase functions are often divergent to the mammalian system (Hammarton et al., 2003; Hammarton et al., 2005; Kumar and Wang, 2006). An example is the trypanosome receptor for activated C kinase (TRACK), which regulates cytokinesis. TRACK contains WD40 repeats and is a homologue of the conserved scaffold protein receptor for activated C kinase 1 (RACK1). RACK1 regulates a variety of cell activities, like protein translation, cell growth and cell shape (McCahill et al., 2002).

1.4. Regulation of gene expression in *T. brucei*

All organisms have to regulate gene expression and usually controlled steps include transcription, mRNA processing, export of mRNAs from the nucleus, nuclear and cytoplasmic degradation of mRNAs as well as regulation of translation. Due to the existence of different forms, the stage-specific regulation of gene expression is essential for the survival of trypanosomes. The wrong expression of stage-specific proteins (e.g. procyclic-form proteins expressed in the bloodstream form) will lead to cell death (Blattner et al., 1998; Wurst et al., 2012).

1.4.1. Transcription

Kinetoplastid genes are constitutively transcribed as groups by RNA polymerase II in a polycistronic manner. These pre-mRNAs can consist of up to 100 different open reading frames (ORFs), that do not encode functionally related products (Clayton et al., 2008). The transcription is initiated and terminated at the so-called strand switch regions (SSRs), which are marked by modified histones (acetylation and methylation) and histone variants (Siegel et al., 2009). In general, the gene expression is not controlled by regulating transcription initiation. However, there is some specificity. A recent study showed that transcription can be driven by a GT-rich promoter, which can deposit the histone variant H2A.Z. The transcription by RNA pol II is then initiated at the 5' end of H2A.Z peaks (Wedel et al., 2017). The rRNA genes as well as stage-specific surface proteins (VSGs and procyclins) are transcribed by RNA polymerase I (Hernandez and Cevallos, 2014). In addition, RNA polymerase III transcribes tRNAs, the 7SL RNA (component of the signal recognition particle) and all U-rich snRNA genes (Vanhamme and Pays, 1995). Polycistronic transcription results in a loss of transcriptional regulation, which means that regulation of gene expression can only be regulated post-transcriptionally. Different kinds of post-transcriptional regulation are polyadenylation, *trans*-splicing, degradation of the mRNAs in the nucleus or cytosol, export from the nucleus or control of translation initiation or elongation (Clayton, 2014). During mRNA-processing the individual mRNAs are generated

by *trans*-splicing of the spliced-leader (SL) mini-exon at the 5'end and the addition of a poly(A) tail at the 3'end (Michaeli, 2011). As a result of this *trans*-splicing each mRNA contains the same 39 nucleotide long SL RNA sequence at its 5'end, which acquires the hypermethylated cap4 structure (Zamudio et al., 2009). Polyadenylation of the mRNA is coupled to the splicing reaction and occurs 100-300nt upstream of the splice site, which is marked with the polypyrimidine tract. Signals within the 5'untranslated region (5'UTR) and the polypyrimidine tract can regulate the efficiency of splicing and thereby influence the abundance of the mature mRNAs (Michaeli, 2011). The hypermethylated cap of the processed transcripts is then bound by the cap-binding complex (CBC), which consists of five subunits: CBP20 (binds directly to the cap), importin-α (might play a role in shuttling between nucleus and cytoplasm) and three uncharacterized proteins (Li and Tschudi, 2005).

1.4.2. Export of mRNA from the nucleus

As soon as the mRNAs are transcribed, the transcripts are bound by various RNA-binding proteins to form messenger ribonucleoproteins (mRNPs) and are exported from the nucleus. This process occurs by a RanGTP-dependent mechanism. In Trypanosomes some components of this complex are conserved, while others are specific for trypanosomes. One example is the heterodimeric nuclear export receptor Mex67-Mtr2, which is conserved in all eukaryotes, but has some special features in trypanosomes (Dostalova et al., 2013). In contrast to yeast, the trypanosome Mex67 contains a N-terminal CCCH-domain, which might be involved in the recruitment of the heterodimeric complex to the SL sequence at the 5'-end of the mRNAs. Downregulation of Mex67 and Mtr2 is essential and leads to a retention of mRNAs in the nucleus (Dostalova et al., 2013). However, there are also other proteins included in the export from the nucleus, like NMD3, which plays a role in the export of procyclin associated genes (PAG) transcripts that are transcribed by RNA pol I (Buhlmann et al., 2015).

1.4.3. mRNA decay in Trypanosomes

Recently, the decay of trypanosome mRNAs was analyzed by a transcriptome-wide study in procyclic and bloodstream forms using transcription inhibition and RNA sequencing (Fadda et al., 2014). The authors found that in the bloodstream form the mRNAs have on average half-lives between 10 and 20 min and are present at 1-4 copies per cell (Fadda et al., 2014; Manful et al., 2011). In general, low abundant mRNAs are unstable, while abundant mRNAs are more stable than the average and/or encoded by more gene copies per cell. Furthermore, mRNAs, which are developmentally regulated, frequently show regulated decay rates. It seems like steady-state mRNA levels are influenced by mRNA splicing and polyadenylation. Nevertheless, this appears to be dependent on the rates of RNA processing and co-transcriptional mRNA precursor destruction (Fadda et al., 2014).

Decay of mRNAs usually starts with the removal of the Poly(A) tail (deadenylation) and is followed by removal of the cap (decapping) and degradation. Like in other eukaryotes, deadenylation is carried out by the CAF1/NOT complex (Erben et al., 2014a; Fadda et al., 2013; Farber et al., 2013) and by the PAN2/PAN3 complex, whose role is not so clear (Fadda et al., 2013; Schwede et al., 2009). After the removal of the poly(A) tail, the mRNA is degraded in the 3'-5' direction by the exosome (Clayton and Estevez, 2010; Fadda et al., 2013) and in the 5'-3'direction by the exoribonuclease XRNA (Manful et al., 2011). In contrast to the deadenylation complex, Trypanosomes lack all homologues of proteins known to be responsible for the decapping of mRNAs. However, Susanne Kramer recently

identified an ApaH-like phosphatase 1 (ALPH1) as major mRNA decapping enzyme. It co-localizes with XRNA at the posterior end of the cell. Knock-down of ALPH1 leads to an increase of total mRNA, which are deadenylated but not yet degraded by 5'-3' decay. This suggests that ALPH1 operates after deadenylation happened, but before mRNA degradation occurs (Kramer, 2017).

1.4.4. Localization of mRNAs in Trypanosomes

An additional way to regulate gene expression on a post-transcriptional way is the storage or decay of mRNA in RNP granules, which are distinct non-membranous structures (Cassola, 2011). Different types of RNP granules exist in trypanosomes. One type are P-bodies, which contain enzymes of the 5'-3' RNA degradation pathway and proteins involved in translational repression, like Caf1 (Kruger et al., 2013). P-bodies are constitutively present in the cells and they contain mRNA in equilibrium with the translating polysomes. If the global translation is repressed, the amount of P-bodies increases, whereas the amount decreases when the dissociation of polysomes is blocked by cycloheximide (Kramer, 2014).

Another type of RNP granule are stress granules, which are larger than P-bodies. They are formed upon stress, like starvation or heat shock, and contain components of the translation initiation machinery (Kruger et al., 2013). It has been shown that starvation stress granules contain proteins like the DEAD box RNA helicase DHH1, the Xrn1 homologue XRNA, the scaffold protein SCD6, the Poly(A) binding proteins PABP1 and PABP2, the U-rich RNA binding protein, translation initiation factors, like eIF4E1, 2 and 3 as well as several RBPs (Kramer, 2014). In contrast to mammalian stress granules, neither ribosomal subunits nor the eIF4G scaffold proteins were detected in stress granules (Cassola, 2011). Stress granules are suggested to serve as temporal storage compartment of mRNAs during stress, because they contain polyadenylated mRNAs (Cassola et al., 2007). As soon as the starvation stress is over, the stress granules disassemble and the mRNAs can be translated. A novel method, which allows the purification of stress granules in trypanosomes, gave new insights in the components of starvation stress granules. The stress granule purification consists of two steps. At first, the plasma membrane of the trypanosomes is lysed. Due to the cage-like subpellicular microtubule array of the cytoskeleton (Angelopoulos, 1970; Lacomble et al., 2009; Sherwin and Gull, 1989), the stress granules will be kept inside this cage while soluble proteins can be washed out. In the second step, the granules are released by depolymerization of the microtubules. Mass spectrometry analysis of the granules enriched fraction identified 463 putative stress granule proteins and RNA seq showed that mRNAs encoding ribosomal proteins are excluded from stress granules (Fritz et al., 2015). This is different to the mammalian system, where most mRNAs encoding ribosomal proteins contain the 5' Terminal Oligo Pyrimidine (5'TOP) motif (Yoshihama et al., 2002) and they are recruited to stress granules in a TIA-dependent manner (Ivanov et al., 2011). This *cis*-regulatory element functions in translational repression of the mRNAs and is also conserved in vertebrates. The 5'TOP mRNAs have an invariable C residue at the cap followed by a stretch of 4 to 15 pyrimidines. Most members have a similar number of C and U residues within this pyrimidine stretch and a CG-rich region downstream of the 5'TOP motif (Meyuhas and Kahan, 2015). Under stress conditions, like nutritional stress, the translation of the 5'TOP mRNAs is repressed. It is known that mRNAs encoding ribosomal proteins are not particularly well translated although they are abundant and very stable in Trypanosoma (Antwi et al., 2016). The regulation of mRNAs encoding ribosomal proteins after stress is very different from that of most RNAs,

since they are neither found in starvation stress granules (Fritz et al., 2015) nor in heat-shock granules (Minia et al., 2016).

In addition to the mentioned types of granules, there are nuclear periphery granules, posterior pole granules and tRNA half granules (Kramer, 2014).

1.4.5. Translation

In general, during translation initiation of eukaryotes eIF4E binds to the cap at the 5'UTR of the mRNA. eIF4E can then interact with the scaffold protein eIF4G, which on its part recruits eIF4A. eIF4A is an RNA helicase, which is involved in the unwinding of secondary structures of the target mRNA. eIF4A thereby facilitates, together with eIF4B, the scanning of the 40S subunit on the target mRNA. The complex consisting of eIF4E, eIF4G and eIF4A is called eIF4F complex. It is known that eIF4G and eIF4B interact with the poly(A)-binding protein (PABP), which binds to the poly(A) tail at the 3' end of the mRNA target. This leads to the circularization of the mRNA, which is suggested to increase translation efficiency and ribosome recycling (Sonenberg and Hinnebusch, 2009). In trypanosomes six eIF4Es, five eIF4Gs and two eIF4As are known. Not all these factors can bind the same proteins and thereby also differ in function (Freire et al., 2017). The six eIF4Es, for example, can be divided in 3 different groups. Group one contains eIF4E1 and eIF4E2, which cannot form eIF4F-like complexes and thereby would not function in general translation (Freire et al., 2017). eIF4E1 is suggested to repress translation by interaction with 4E-IP, which is needed for normal differentiation (Terrao et al., 2018). In addition, eIF4E1 is suggested to stimulate translation by interaction with the eIF3 complex. However, this interaction could not be confirmed by our lab (unpublished data). Not much is known about the function of eIF4E2. Group 2 contains eIF4E3 and eIF4E4, which can form eIF4F complexes and therefore are responsible for general translation. Group 3 contains eIF4E5 and eIF4E6, which can bind to eIF4G and form eIF4F(-like) complexes. However, these complexes are suggested to not be involved in general translation (Freire et al., 2017).

Studies comparing the transcriptomes and proteomes of procyclic and bloodstream forms (Butter et al., 2013; Gunasekera et al., 2012; Urbaniak et al., 2012) showed that mRNA and protein levels correlate comparatively poorly meaning that there has to be regulation of protein stability or translation. The enormous regulation of translation could be shown by ribosome profiling data. The translation efficiency varies heavily between the different life cycle stages and up to 100-fold between different genes. This suggests that the gene expression is regulated by translation efficiency as much as by mRNA stability (Vasquez et al., 2014). As described above: upon stress conditions the translation is generally suppressed and mRNA sequester in granules (Kruger et al., 2013; Zinoviev et al., 2012). Not much is known about the mechanism of translational regulation in trypanosomes. Only some mechanisms for translational regulation of specific transcripts by RBPs are known for Opisthokont cells, like the pumilio domain proteins that lead to the sequestration in granules (Kotani et al., 2013) or the inhibition of translation elongation by interaction with eEF1A (Friend et al., 2012).

1.4.6. Influence of codon usage on the regulation of gene expression

in the last years, it has been shown that codon usage can determine mRNA levels in a variety of organisms. These data show that codons, which are translated into the same amino acid, are used with a different abundance dependent on the different species, but also within one organism and even between different transcripts. It is thought that the

abundance of the tRNA at least partly determines the codon optimality (how fast and efficient a codon is translated) and that transcripts with higher codon optimality are prone to be more stable (Hanson and Coller, 2018). More recently two studies also showed the influence of codon usage on translation in Trypanosomes. Both groups compared the mRNA abundances, determined by RNA-Seq, with the codon usage (de Freitas Nascimento et al., 2018; Jeacock et al., 2018). Jeacock *et al.* could show that highly abundant mRNAs encoding abundant proteins have a higher codon optimality, whereas it decreases in poorly expressed genes. In addition, they could show that it is more likely for genes encoding proteins of one protein complex to have a similar codon composition (Jeacock et al., 2018). de Freitas Nascimento *et al.* observed quite similar results showing that the codon optimality correlates with the reporter protein and mRNA levels. In addition, they could show that the codon usage influences the mRNA half-life and that translation is needed for that process (de Freitas Nascimento et al., 2018). Taken together, these two studies suggest an influence of the codon usage on the half-life of the mRNAs in *Trypanosoma*. However, it is not known if and how the codon usage influences mRNA decay rates.

1.5. RNA-binding proteins and their role in regulation of gene expression in *T. brucei*

RNA-binding proteins (RBPs) play an important role in the post-transcriptional regulation of gene expression (Kramer and Carrington, 2011). RBPs are classified according to their structural domains that are involved in the interaction with RNA. The genome of *T. brucei* encodes about 48 CCCH-type zinc finer domain proteins (Kramer et al., 2010), over 75 RNA recognition motif (RRM) proteins (De Gaudenzi et al., 2005), 4 ‘acetylation lowers binding affinity’ (ALBA) domain proteins (Mani et al., 2011) and at least 12 PUF domain proteins (Caro et al., 2006). The RBPs can bind to specific sequences within each transcript and can regulate the localization, stability or translation of the target. The majority of regulatory *cis*-elements are located within the 3’UTR of target mRNAs (Clayton and Shapira, 2007). In addition, each RBP can have multiple targets.

1.5.1. RRM domain containing proteins

The RRM domain is one of the most abundant protein domains in eukaryotes (Clery et al., 2008). A single RRM domain consists of two α-helices and a four-stranded β-sheet and can interact with a 2-8 nucleotide long sequences on single stranded RNA (Lunde et al., 2007). Approximately half of the RRM domain-containing proteins are essential in at least one life cycle stage of *T. brucei* (Alsford et al., 2011) and many are known to regulate mRNA degradation, splicing or translation. However, only some RRM-domain containing proteins in *Trypanosoma* have orthologues in other eukaryotes, like PABP1 and PABP2, the translation initiation factor eIF3B and the splicing factor U2AF35 (Kramer and Carrington, 2011). The uridine-binding proteins UBP1 and UBP2, which are two related proteins with a single RRM domain, affect the abundances of F-box protein mRNAs involved in cell cycle control (Hartmann et al., 2007). The cytosolic RBP42 binds to the coding sequence (CDS) of mRNAs encoding proteins involved in energy metabolism (Das et al., 2012). In addition, two RRM domain containing proteins, RPB6 and RBP10 are involved in the life-stage specific gene expression (Kolev et al., 2012; Mugo and Clayton, 2017; Wurst et al., 2012). RBP10 is a cytoplasmic protein specifically expressed in the bloodstream from cells. While knock-down of RBP10 leads to a decrease of mRNAs specific for the bloodstream form, expression of RBP10 in the procyclic form leads to an increase of mRNAs specific for the bloodstream form (Wurst et al., 2012). In the bloodstream form RBP10 binds to an

UAUUUUUU motif of procyclic-specific mRNAs leading to their translational repression and destruction. On one hand, knock-down of RBP10 in the bloodstream form enabled the differentiation to procyclic cells, when the cells were transferred to 27°C and procyclic medium, even without cis-aconitate, which usually primes the cells for differentiation. On the other hand, expression of RBP10 in the procyclic form leads to cells with higher expression of VSGs and reduced expression of procyclin and culturing of these cells in bloodstream form medium and at 37°C leads to differentiation to bloodstream form cells (Mugo and Clayton, 2017). Procyclic cells expressing RBP10 do not form epimastigotes, although they become infective for mice, which suggests an alternative differentiation pathway (Mugo and Clayton, 2017; Mugo et al., 2017). Expression of RBP6 in the procyclic form initiates the differentiation to epimastigotes and metacyclic cells and after transfer to mice the cells differentiate to bloodstream form cells. However, *in vitro* the metacyclic cells could not differentiate to the bloodstream form (Kolev et al., 2012). A more recent study could show that expression of RBP6 containing a single point mutation (Q109K) enables the procyclic cells to differentiate into metacyclic cells and then progress to the bloodstream form *in vitro*. In addition, the expression of the mutated RBP6 skips the intermediate epimastigotes form (Shi et al., 2018) as it was shown for RBP10 (Mugo and Clayton, 2017).

1.5.2. ALBA domain containing proteins

In *T. brucei* only four ALBA domain proteins are known: ALBA 1-4. All four are can form homo- and heterodimeric complexes and are located in the cytoplasm. In addition, knock-down of all four ALBA domain proteins leads do a decrease of translation of a reporter mRNA (Mani et al., 2011). Dimers of ALBA1/2 and ALBA 3/4 can be found in starvation stress granules together with poly(A) RNA, whereas dimers of ALBA2/3 interact with eIF4E4 and partially associate with polysomes. The association of the ALBA proteins with stress granules and the translation machinery suggests a role in the control of translation (Mani et al., 2011). Subota *et al.* published a study about the characterization of ALBA3/4, where they could show that ALBA3/4 co-localizes with DHH1 and poly(A) RNA in stress granules (Subota et al., 2011), which coincide with the localization studies of Mani *et al.* (Mani et al., 2011). However, they went further and could show that ALBA3/4 is expressed in all developmental stages of Trypanosomes in the Tsetse fly, except for the transition states in the proventriculus region. In this phase the nucleus migrates towards the posterior end of the cell, which can be disturbed by the expression of ALBA3 (Subota et al., 2011).

1.5.3. PUF domain containing proteins

A typical PUF domain includes eight tandemly-repeated α-helices, which separately interact with eight bases of the RNA binding sequence (Wang et al., 2002). PUF9, for example, binds to the putative recognition motif UUGUAC of a small number of target mRNAs and thereby stabilizes them during S-phase (Archer et al., 2009). PUF7 and PUF10 are both involved in rRNA maturation and localize to the nucleolus (Droll et al., 2010; Schumann Burkard et al., 2013). PUF7 interacts with a nuclear cyclophilin-like protein and knock-down of PUF7 results in reduced cell growth and inhibition of ribosomal RNA processing (Droll et al., 2010). Another study showed that PUF7, as well as PUF10, interact with BOP1, which is a protein involved in rRNA processing, and all of them interact with the nucleolar regulator of GPEET 1 (NRG1). Knock-down of any of these proteins leads to a reduction of the 5.8S rRNA level and its immediate precursor, and to increased GPEET expression. These data suggest that the proteins of the rRNA maturation complex can in addition regulate mRNAs with origin in the nucleolus (Schumann Burkard et al., 2013). PUF1, interacts with the expression site-associated gene 8 (ESAG8) protein as well as mRNA. Expression sites (ES)

promote the expression and switching of VSGs and the ESAGs are co-transcribed from the upstream promoter. Overexpression of PUF1 protein leads to an increase of ESAG8 mRNA, which suggests a positive feedback loop: high PUF1 levels will lead to an increase of the levels of ESAG8 mRNA and protein. In addition, overexpression of PUF1 leads to a reduced parasite virulence. However, it is not clear, if this is a direct effect of PUF1 overexpression and its interaction with ESAG8 (Hoek et al., 2002). In contrast to these results, depletion of PUF1 did not affect cell growth or the abundances of any other mRNAs in trypanosomes. This suggests a functional redundancy of the PUF proteins in *Trypanosoma* (Luu et al., 2006).

1.5.4. Zinc finger domain containing proteins

Proteins containing a zinc finger domain are defined by the presence of a C-X_{4-15}-C-X_{4-6}-C-X_3-H, which mostly binds single stranded RNA. These proteins have widespread functions in RNA metabolism. Around 65% of the zinc finger proteins have only one CCCH motif, while the rest have additional domains (Kolev et al., 2014; Kramer and Carrington, 2011). The three zinc finger proteins ZFP1, ZFP2 and ZFP3 play a role in differentiation from the bloodstream form to the procyclic form (Hendriks and Matthews, 2005; Hendriks et al., 2001). ZC3H11 is essential in the bloodstream form and procyclic cells, and is required to survive a heat-shock response at 41°C. ZC3H11 binds to mRNAs encoding chaperones and stabilizes these mRNAs. While the zinc finger domain at the N-terminus of ZC3H11 recognizes UAU repeats in the 3'-UTR of the mRNA, the C-terminal domain stabilizes the bound mRNA (Droll et al., 2013). ZC3H11 then recruits MKT1 and PBP1. PBP1 in turn recruits LSM12 and PABP. This could protect the poly(A) tail against deadenylation and in addition might facilitate the circularization with the cap-binding complex (Singh et al., 2014). In addition, ZC3H20, ZC3H18 and ZC3H13 are needed for differentiation and cell growth (Benz et al., 2011; Ling et al., 2011; Ouna et al., 2012).

1.5.5. Non-canonical RBPs

A major feature of proteins that regulate post-transcriptional gene expression is their ability to bind to mRNA. To identify these RBPs, the mRNA-bound proteome of bloodstream cells was analyzed (Lueong et al., 2016). Proteins were crosslinked with the mRNAs *in vivo* and the polyadenylated mRNAs were enriched by oligo (dT) magnetic beads. The mRNA interactome was then determined by quantitative mass spectrometry (Castello et al., 2013). This screen identified many putative RBPs, of which many were not known to bind to mRNA before. They include non-canonical RBPs (Lueong et al., 2016), which can bind to RNAs even without a known RBP (Gerstberger et al., 2014; Hogan et al., 2008; Lueong et al., 2016). A trypanosome example is Tb927.10.14150, which was one of the strongest hits in the *T. brucei* mRNA interactome (Lueong et al., 2016).

One non-canonical RBP in yeast is Bfr1p (Brefeldin A resistance protein). It was discovered in the 90s in a genetic screen for high-copy suppressors of Brefeldin A (BFA) in *Saccharomyces cerevisiae*. BFA, which is a fungal toxin, influences the function and structure of the organelles of the secretory pathway (Klausner et al., 1992; Lippincott-Schwartz, 1993). It leads to tubulation of the lysosome, to fusion of the trans-Golgi network with the endosomes and to tubulation and fusion of the Golgi with the ER (Hunziker et al., 1992; Klausner et al., 1992; Pelham, 1991). Bfr1p can partially suppress the phenotypes caused by BFA (Jackson and Kepes, 1994). Bfr1p is not essential (Jackson and Kepes, 1994) and localizes to the outside of the ER (Lang et al., 2001). Lang and colleagues could show that Bfr1p is part of a polyribosome-associated mRNP complex, which contains

Scp160p, Pap1p, additional unidentified proteins and polyadenylated mRNA (Lang and Fridovich-Keil, 2000; Lang et al., 2001). A more recent study suggests that Bfr1p together with Scd160 inhibits P-body formation under normal growth conditions and thereby protects RNAs at ribosomes (Figure 1.4 A). By an interaction with RNA and P-body components, the P-bodies are left in a 'waiting position' and P-body formation is inhibited (Simpson et al., 2014; Weidner et al., 2014). In Bfr1p-depleted cells (Figure 1.4 B), Scp160p cannot be recruited to the polysomes efficiently and thereby cannot protect the RNA, which leads to the formation of P-bodies. In contrast, if Scd160 is depleted (Figure 1.4 C), Bfr1p can, to an extent, still be recruited to the polysomes.

However, P-bodies cannot be assembled properly without Scp160, which leads to the formation of pseudo P-bodies. This model suggests that under stress conditions, Scp160 and Bfr1 cannot protect the polysomes anymore, which gives the P-body components access to the mRNAs and leads to P-body formation (Weidner et al., 2014). Scd160p as well as Bfr1p were identified to bind to more than a thousand mRNA targets by RIP-Chip and the targets were enriched for mRNAs encoding proteins that localize to the nucleolus and that are involved in ribosome biogenesis and RNA processing (Hogan et al., 2008). A more recent study identified the mRNA targets of Bfr1p by RNA tagging. The targets were enriched for mRNAs encoding for proteins involved in cytoplasmic translation, containing ribosomal proteins, and membrane-associated functions (Lapointe et al., 2015). Taken together, the localization of Bfr1p at the outside of the ER (Lang et al., 2001), the association of Bfr1p with polysomes (Weidner et al., 2014) and the enrichment of its mRNA targets for

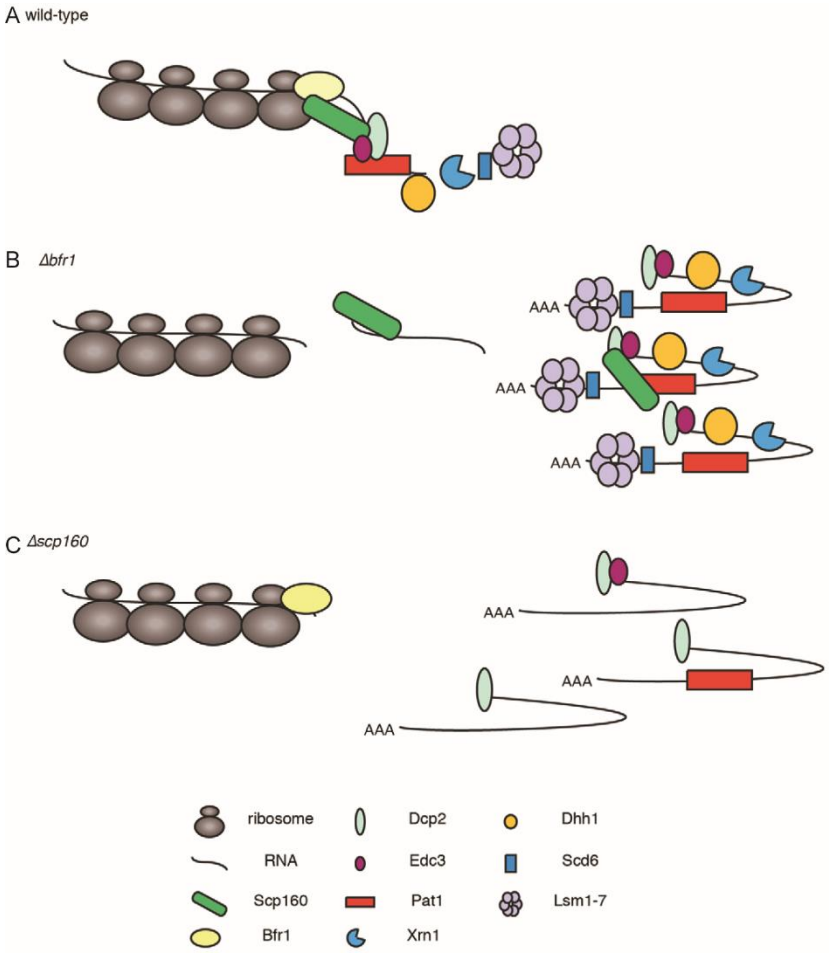


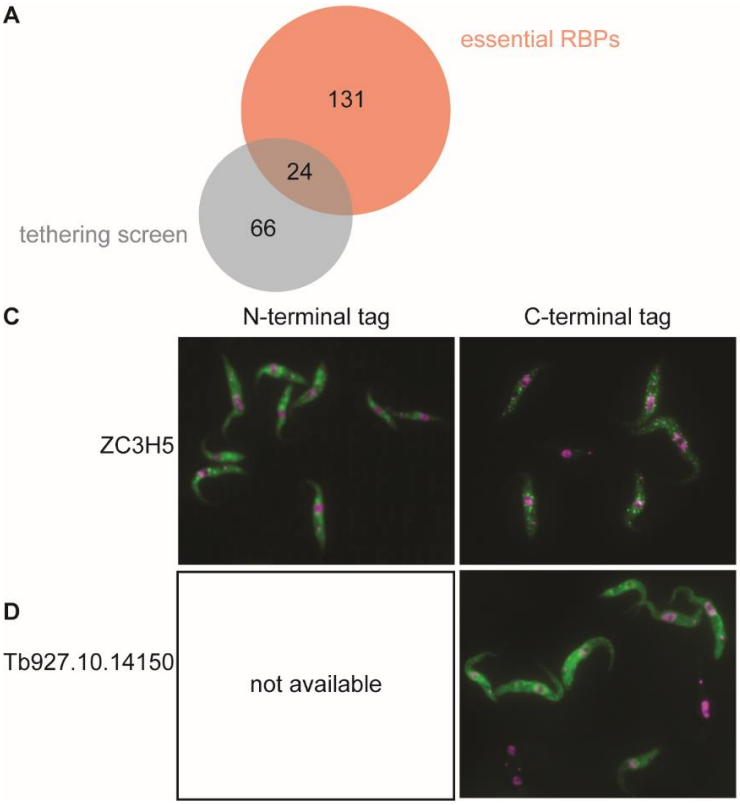
Figure 1.4: Scp160 and Bfr1 inhibit P body formation under normal growth conditions. A. In wild-type cells under normal growth conditions, Scp160 and Bfr1 protect RNA at polysomes. Scp160 interacts with RNA and P body components and leaves P bodies in a 'waiting' position while still inhibiting P body formation. **B.** In Δbfr1 cells, Scp160 is not efficiently recruited to the polysomes and cannot properly protect the RNA, thus, allowing P body components to access RNAs. **C.** In Δscp160, Bfr1 can, to an extent, be recruited to polysomes, but without Scp160, proper P body assembly cannot occur, leading to the formation of pseudo P bodies. Figure and description from: (Weidner et al., 2014).

membrane-related proteins (Lapointe et al., 2015) suggests that the Bfr1p targets are directly translated to the ER.

1.6. Identification of new putative post-transcriptional regulators

A tethering assay was performed to identify putative trypanosome post-transcriptional regulators. The basis of this method is the interaction of the lambda-N peptide with the boxB RNA sequence. Proteins of interest were fused N-terminally to the lambda-N peptide and inducibly expressed in trypanosomes. About 300 proteins that potentially regulate mRNA fate were identified by the screen. Many proteins with no previously known function could be identified as activators, and repressors, as well as 25 RBPs and known translation initiation factors as activators and 16 RBPs and 4E-IP as repressors (Erben et al., 2014b). In the tethering screen only random genomic fragments were used leading to some false-negatives due to the missing of full-length proteins. There were also false-positives, which could arise from wrong folding of the protein fragments. A mini-library of 384 full-length ORF fragments was created containing proteins with RNA-binding domains, some translation factors, components of the RNA degradation machinery as well as proteins with previously unknown function. This mini-library identified 90 proteins that were activators or repressors (Lueong et al., 2016). This was the basis of my study. I wanted to investigate the function of putative post-transcriptional regulators in more detail. To choose the candidates I set 4 premises: proteins that 1. were shown to be repressors or activators in the tethering screen (Erben et al., 2014b; Lueong et al., 2016), 2. Were shown to be essential by RNA interference target sequencing (RIT-Seq) (Alsford et al., 2011), 3. Were shown to bind to mRNA by interactome capture (Lueong et al., 2016) and 4. Are located in the cytosol according to TrypTag (Dean et al., 2017) (Figure 1.5). 155 of the putative RBPs identified by interactome capture (Lueong et al., 2016) were also shown to be essential by RIT-Seq (Alsford et al., 2011) (Figure 1.5 A). 24 of these candidates also showed an effect as repressor or activator in the tethering screen (Figure 1.5 B). Next, the localization of these proteins according to TrypTag was investigated (Dean et al., 2017). I will focus on proteins that are located outside the nucleus, because post-transcriptional regulation of gene expression, like regulation of translation, mRNA degradation, etc. occurs in the cytoplasm.

The two proteins, which were chosen as candidates are one activator, Tb927.10.14150, and one suppressor, ZC3H5 (Erben et al., 2014b). (Figure 1.5 B). Tb927.10.14150 is a protein of unknown function that is conserved in all Kinetoplastids, but no other organisms. It showed *in vivo* mRNA binding (Lueong et al., 2016), although it lacks a canonical RNA-binding domain. Tb927.10.14150 shows the same relative protein abundance in both stages during differentiation (Dejung et al., 2016). In addition, it has some similarities to yeast Bfr1p (TriTrypDataBase). Tb927.10.14150 was shown to be essential in a high-throughput screen in bloodstream form cells (Alsford et al., 2011). According to TrypTag, a C-terminally tagged version is located in the cytoplasm and at the ER (Figure 1.5 D) (Dean et al., 2017). ZC3H5 is an RNA-binding protein that is conserved in Trypanosomatids and contains a single C3H1-type zinc finger domain. It was shown to be essential in a high-throughput screen in bloodstream form cells (Alsford et al., 2011). According to TrypTag, a N-terminally tagged version is located in the cytosol and a C-terminally tagged version is located in the cytosol as patchy structures (Figure 1.5 C) (Dean et al., 2017).



B

gene ID	protein
Tb927.10.12660	PUF2
Tb927.10.14150	hypothetical protein
Tb927.10.14930	ZC3H39
Tb927.10.14950	ZC3H40
Tb927.10.1540	ZC3H30
Tb927.10.310	PUF3
Tb927.11.2300	ERF1
Tb927.11.6160	ERF3
Tb927.3.1920	NOT5
Tb927.3.3940	DRBD11
Tb927.3.5250	ZC3H8
Tb927.3.740	ZC3H5
Tb927.4.400	DRBD7
Tb927.5.1580	ZC3H13
Tb927.5.4320	FIP1
Tb927.6.3480	DRBD5
Tb927.6.4440	RBP42
Tb927.6.4770	MKT1
Tb927.6.600	CAF1
Tb927.6.850	NOT2
Tb927.8.2780	RBP10
Tb927.8.910	hypothetical protein
Tb927.9.13990	DRBD2
Tb927.9.7110	hypothetical protein

Figure 1.5: How to choose candidate RBPs? A. Activators or repressors of the tethering screen (n=90) (Lueong et al., 2016) were compared with essential RBPs (n=155). Essential RBPs are proteins, which were identified in the mRNA interactome capture (Lueong et al., 2016) and showed reduced growth fitness by RNA interference target sequencing (Alsford et al., 2011). **B.** Table of the overlapping candidates form figure A. Green=activator; red=repressor in the tethering screen. Grey box: candidates that were chosen for closer analysis. **C.** TrypTag images of N-terminally and C-terminally GFP-tagged ZC3H5 (green). Nucleus and Kinetoplast were stained with Hoechst (cyan) (Dean et al., 2017). **D.** TrypTag images of C-terminally GFP-tagged Tb927.10.14150 (green). Protein is not tagged N-terminally, yet. Nucleus and Kinetoplast were stained with Hoechst (cyan) (Dean et al., 2017).

1.7. Aim of the study

In this thesis, I will investigate the function of the two candidates Tb927.10.14150 and ZC3H5 in more detail.

The specific aims were

- localization of proteins
- Analysis of knock-down and effect on mRNA activity/stability
- Identification of the target mRNAs
- Identification of the protein interaction partners

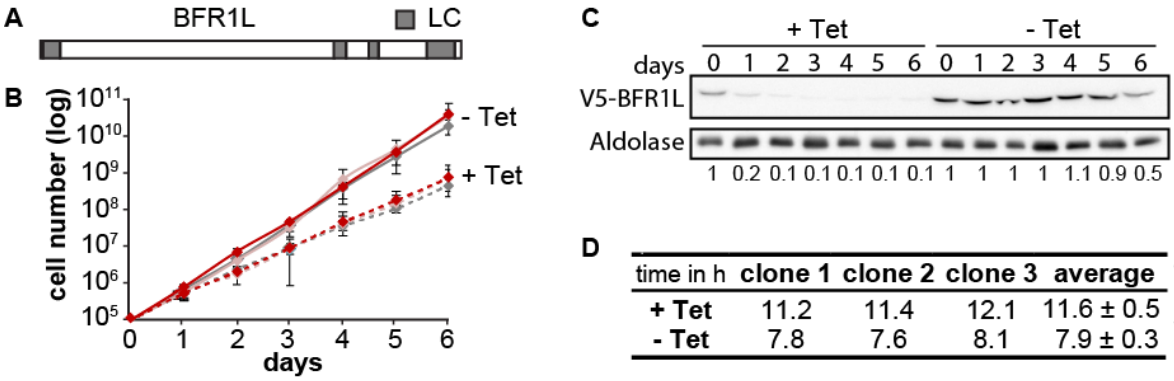
Dependent on these aims a deeper knowledge on the role of the two candidate proteins in the post-transcriptional regulation of gene expression should be investigated.

2. mRNAs encoding ribosomal proteins might be regulated by BFR1L

2.1. Results

2.1.1. Knock-down of Tb927.10.14150 affects cell growth in the bloodstream form

In this part of the thesis, I will report on the characterization of Tb927.10.14150, which is a putative RNA-binding protein without any canonical RNA-binding domain that is conserved in Kinetoplastids and has also some orthologues in oomycetes and brown algae. Sequence identity in Kinetoplastida is equally spread through the protein sequence (Suppl. Figure 1). However, sequence identity to the oomycetes and brown algae is not very high. Tb927.10.14150 was an up-regulator in the tethering screen (Erben et al., 2014b) and the relative protein abundance of the protein is similar in both stages during differentiation (Dejung et al., 2016). Tb927.10.14150 protein displays *in vivo* mRNA binding (Lueong et al., 2016) although it lacks canonical RNA-binding domains. It has some similarities to yeast Bfr1p (Brefeldin A resistance protein), an ER- and polysome-associated protein (Lang et al., 2001; Weidner et al., 2014), which is not essential (Jackson and Kepes, 1994). Bfr1p interacts with RNA, although it lacks canonical RNA-binding domains. Bfr1p together with Scd160 inhibits P-body formation under normal growth conditions and thereby protects RNAs at ribosomes (Weidner et al., 2014). Since Tb927.10.14150 has some similarities to the yeast Bfr1 protein, even if the sequence identity is very low (Suppl. Figure 2.1), I will rename Tb927.1014150 as BFR1-like (BFR1L) protein. BFR1L consists of 479 amino acids. It has a predicted molecular weight of 55.3 kDa and contains four low complexity regions, of which one is located in the N-terminal region and the other three are located in the C-terminal region (Figure 2.1 A). To investigate whether BFR1L is essential in bloodstream form trypanosomes (BF), cell growth after RNAi-mediated knock-down of BFR1L was analyzed. I could observe a growth defect upon RNAi (+ Tet) in comparison to the uninduced cells (- Tet), but it was not lethal (Figure 2.1 B). The decrease of the *in situ* V5-tagged protein after RNAi was monitored over time and analyzed by Western Blotting (Figure 2.1 C). The protein level decreased already after 1 day of Tetracycline induction and was not detectable at later points, which showed successful knock-down. The division



time in h	clone 1	clone 2	clone 3	average
+ Tet	11.2	11.4	12.1	11.6 ± 0.5
- Tet	7.8	7.6	8.1	7.9 ± 0.3

Figure 2.1: Knockdown of BFR1L expression in BF V5-BFR1L cells. A. Domains of BFR1L. LC: Low complexity region. **B.** Growth curve for dsRNAi of BFR1L in BF V5-BFR1L cells of three independent clones. Error bars indicate standard deviation. +Tet (dashed lines) and -Tet (solid lines). **C.** 3x10^6 cells were collected each day after counting and Western Blot of V5-BFR1L was performed. Aldolase was used as loading control. Numbers below the blot indicate quantification of the V5-BFR1L signal. Day 0 was set as 1. **D.** Division times (in h) of +/- Tet calculated from 3 independent clones time ± standard deviation.

time after knock-down (+Tet) was 11.6 ± 0.5 h, whereas the division time of the control cells (-Tet) was 7.9 ± 0.3 h (Figure 2.1 D).

To investigate whether BFR1L is essential in the bloodstream form, both ORFs of *BFR1L* were replaced with two different genes encoding for antibiotic resistances (Figure 2.2 A). At first, a single knockout (SKO) was created in which one ORF of BFR1L was replaced by the gene encoding for Blasticidin-S deaminase (BSD). To create a double knockout (DKO) the second ORF of *BFR1L* was replaced by the gene encoding for Puromycin N-acetyltransferase (PAC). To confirm that both genes of *BFR1L* were knocked out in this cell

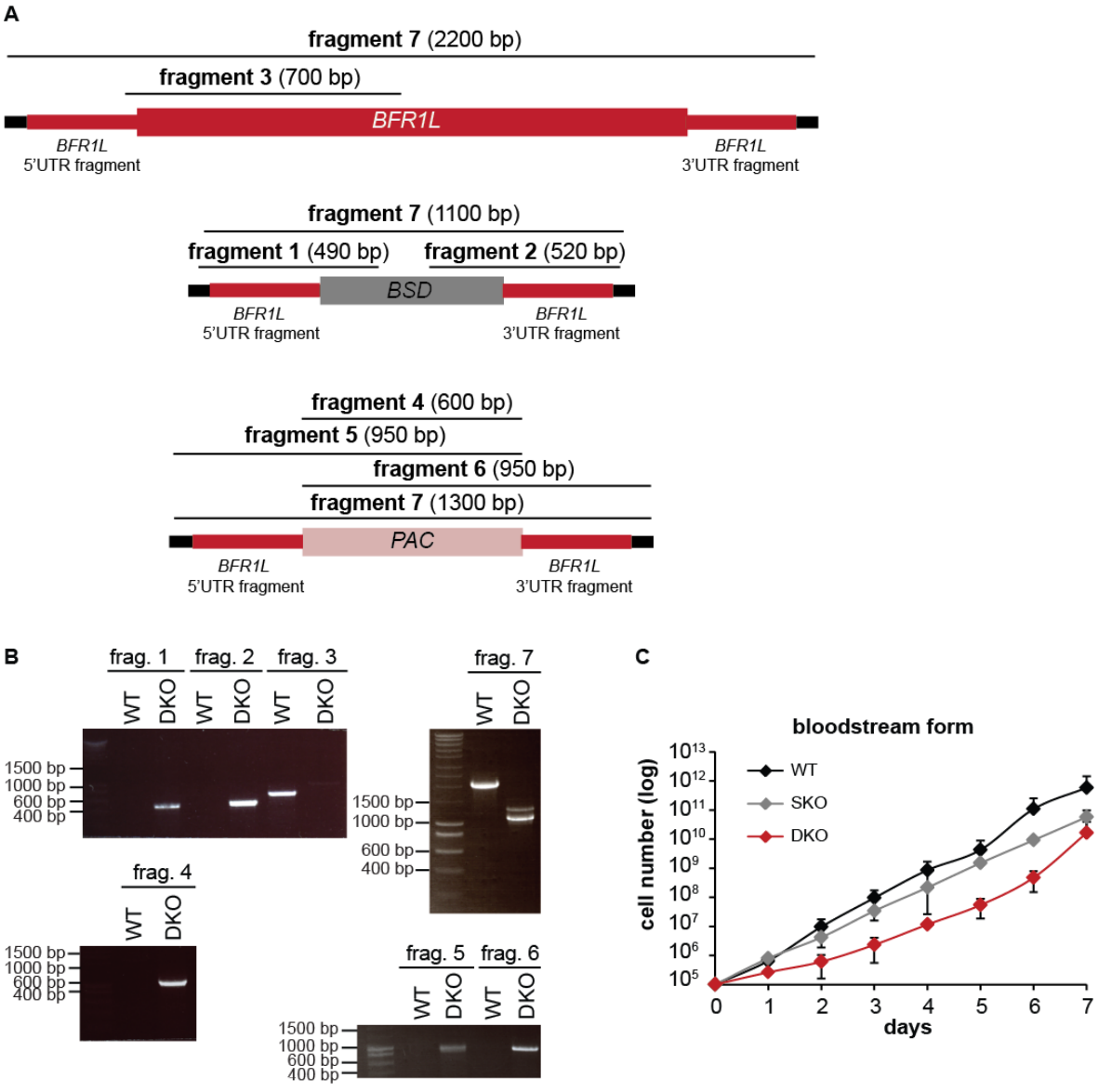


Figure 2.2: Growth of bloodstream form cells without *BFR1L*. A. Schematic representation of the different fragments that were amplified by PCR. BSD: Blasticidin-S deaminase; PAC: Puromycin N-acetyltransferase. **B.** gDNA of wild type BF cells (WT) and BF *BFR1L*$^{-/-}$ (DKO) was extracted and amplified using different primer pairs to amplify the different fragments described in A. **C.** Cumulative growth curve of 3 independent experiments, which were done at different times but with the same clones, for BF (WT), BF *BFR1L*$^{+/-}$ (SKO) and BF *BFR1L*$^{-/-}$ (DKO). Error bars indicate standard deviation.

line, gDNA of the WT and DKO was extracted and different fragments were amplified by PCR, as shown in figure 2.2 A. The amplified PCR fragments can be seen in figure 2.2 B. Fragment 1 or 2 could only be amplified using the DKO gDNA as template, but not the WT gDNA as template, which indicated that one allele of the *BFR1L* ORF was replaced by *BSD* gene. Otherwise, primers to amplify the ORF of *BFR1L* could only amplify fragment 3 by using the WT genomic DNA (gDNA) as template, but not by using the DKO gDNA as template. Replacement of the second *BFR1L* allele by *PAC* gene was investigated by amplification of fragment 4 and 5. In both cases, the fragment could be amplified by using the DKO gDNA as template, but not by using the WT gDNA. In addition, primers to amplify the *PAC* gene can only amplify fragment 6 by using the DKO gDNA as template, but not by using the WT gDNA as template. Finally, amplification of fragment 7 was investigated. This fragment has different sizes dependent on the sequence length of the different genes. It can be seen that fragment 7 had a size of 2200 bp, when WT gDNA was used as template. When DKO gDNA was used as template two bands with sizes of 1100 bp and 1300 bp appeared indicating replacement of the *BFR1L* gene by the *PAC* and *BSD* gene. After confirmation of successful knock-out of BFR1L, the growth of DKO cells was compared with SKO cells and WT cells over seven days in 3 independent experiments (Figure 2.2 C). The SKO cells showed a negligible growth defect in comparison to the WT cells. I could observe a growth defect of the DKO cells in comparison to the WT and SKO cells, but it was not lethal. This reflects the data of BFR1L knock-down by dsRNAi.

Next, I investigated, whether the growth defect of the DKO bloodstream form cells could be complemented by overexpression of myc-tagged BFR1L. For that reason, double-knockout cells with inducible ectopically expressed N-terminally myc-BFR1L (BF *BFR1L*$^{-/-, MYC}$) were created, which is a conditional DKO (cDKO). The growth of these cells with and without induction of myc-tagged BFR1L expression was compared with the growth of WT BF cells and the *BFR1L*$^{-/-}$ cells (DKO). As described before, I could observe a growth defect of the DKO cells in comparison to the WT cells. The BF *BFR1L*$^{-/-, MYC}$ cells grew as the WT in both cases: with and without tetracycline induction (Figure 2.3 A). However, the overexpression plasmid is leaky as it can be seen by Western Blotting. Even in the uninduced cells (-Tet), a myc-BFR1L band could be observed. Surprisingly, BFR1L appeared as double band, which I could not observe on other WBs (see figure 2.1 C). This experiment showed that the growth defect of the DKO cells can be rescued by overexpression of myc-tagged BFR1L. However, when repeating this experiment, I could not observe a growth defect of the DKO. It is possible that cells have simply been adapted to loss of BFR1L during culturing. In procyclic cells (PC), I was not able to generate a gene knock-down of the protein neither by dsRNAi nor by stem-loop RNAi, despite several attempts (data not shown). In addition, I was also not able to generate a double knockout of *BFR1L* in PCs; however, I could create a SKO of the protein in PCs. These cells were then transfected with the ectopic expression plasmid to create PC *BFR1L*$^{-/+, MYC}$ cells (conditional SKO=cSKO). In these cells, myc-tagged BFR1L was expressed by Tetracycline induction and then they were transfected with a plasmid that replaces the second ORF of *BFR1L* with the *PAC* gene to create a conditional DKO (PC *BFR1L*$^{-/-, MYC}$). Surprisingly, all different cell lines (SKO, cSKO, cDKO) cells grew as the WT in both cases: with and without tetracycline induction (Figure 2.3 B; performed by Pia Hartwig (PH)). As it can be seen by Western Blotting, the overexpression of myc-BFR1L is tightly regulated by tetracycline induction. In the uninduced cells no myc-BFR1L signal can be observed, which suggests that BFR1L is not essential in the procyclic form.

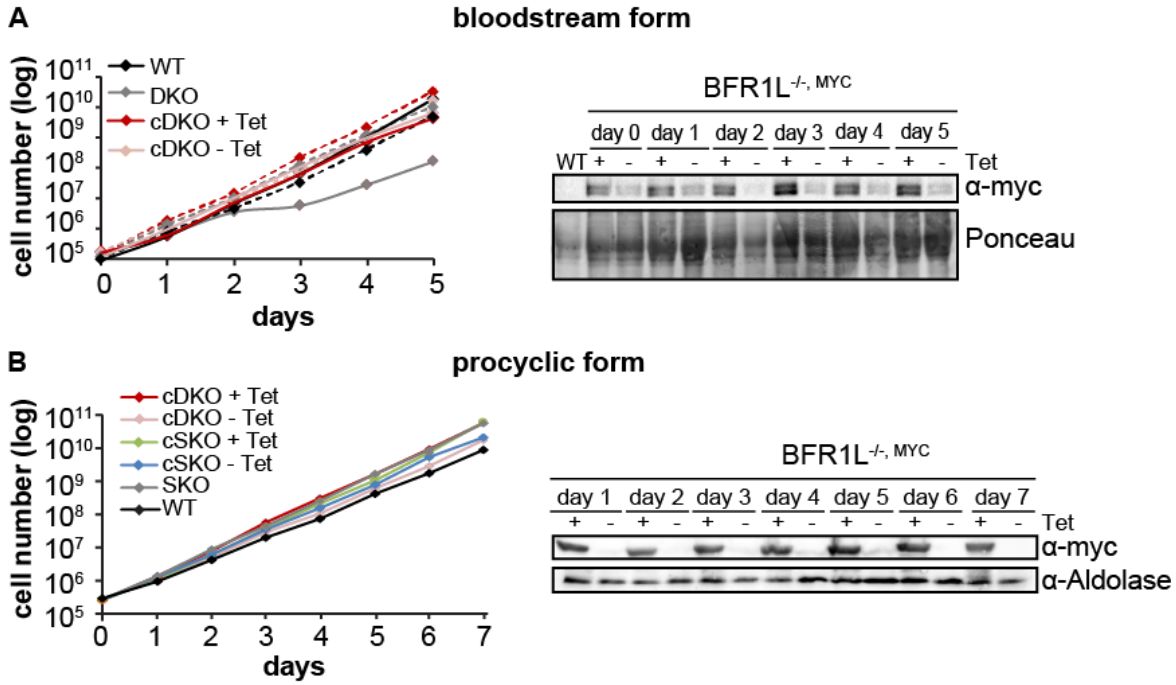


Figure 2.3: Effect of conditional *BFR1L* DKO on growth of bloodstream form and procyclic cells. A. Growth curve for BF (WT), BF *BFR1L*$^{-/-}$ (DKO) and BF *BFR1L*$^{-/-, MYC}$ (DKO plus ectopic expression plasmid for myc-BFR1L) without (-Tet) and with (+Tet) tetracycline induction. 3x10^6 cells were collected each day after counting and a Western Blot for myc-BFR1L was performed. Ponceau staining was used as loading control. Solid and dashed lines are two replicates. **B.** Growth curve for PC (WT), SKO (*BFR1L*$^{+/-}$), cSKO (*BFR1L*$^{+/-, MYC}$) and cDKO (*BFR1L*$^{-/-, MYC}$) without (-Tet) and with (+Tet) tetracycline induction. 3x10^6 cells were collected each day after counting and Western Blot of myc-BFR1L was performed. Aldolase was used as loading control. Representative growth curves out of two replicates is shown (PH).

Taken together, knock-down or knock-out of *BFR1L* in BF trypanosomes leads to a slight growth defect, which can be rescued by ectopic expression of the protein, whereas I was neither able to generate a RNAi-mediated knock-down cell line of *BFR1L* nor to generate a DKO cell line in PC trypanosomes. However, I was able to generate a cDKO in PCs. Surprisingly, growth was not restricted by removing of tetracycline from these cells and cells grew as wildtype even without ectopic expression of myc-tagged BFR1L.

2.1.2. BFR1L localizes to ER and partially to mitochondria

Immunofluorescence microscopy data of the TrypTag project (Dean et al., 2017) suggests a localization of BFR1L at the ER. To investigate the localization of the BFR1L in more detail, immunofluorescence microscopy was performed using bloodstream form cells expressing N-terminally V5-tagged BFR1L from the endogenous locus. Proteins with known localization were used as controls and the kinetoplast and nuclear DNA was stained with DAPI. Wild type BF cells were used as control to validate that the V5-antibody binds specifically to V5-tagged BFR1L. I recorded Z-stacks and deconvoluted the images afterwards. Figure 2.4 A shows the localization of V5-tagged BFR1L and Trypanothione Reductase (TR), a protein located in the cytoplasm. There might be partial co-localization with TR. However, TR is not an exclusively cytosolic marker, because I could observe signal in the nucleus. I could not observe a signal of V5-BFR1L in the nucleus. In figure 2.4 B the localization of V5-tagged BFR1L and Aldolase, a protein located in the glycosomes, can be seen.

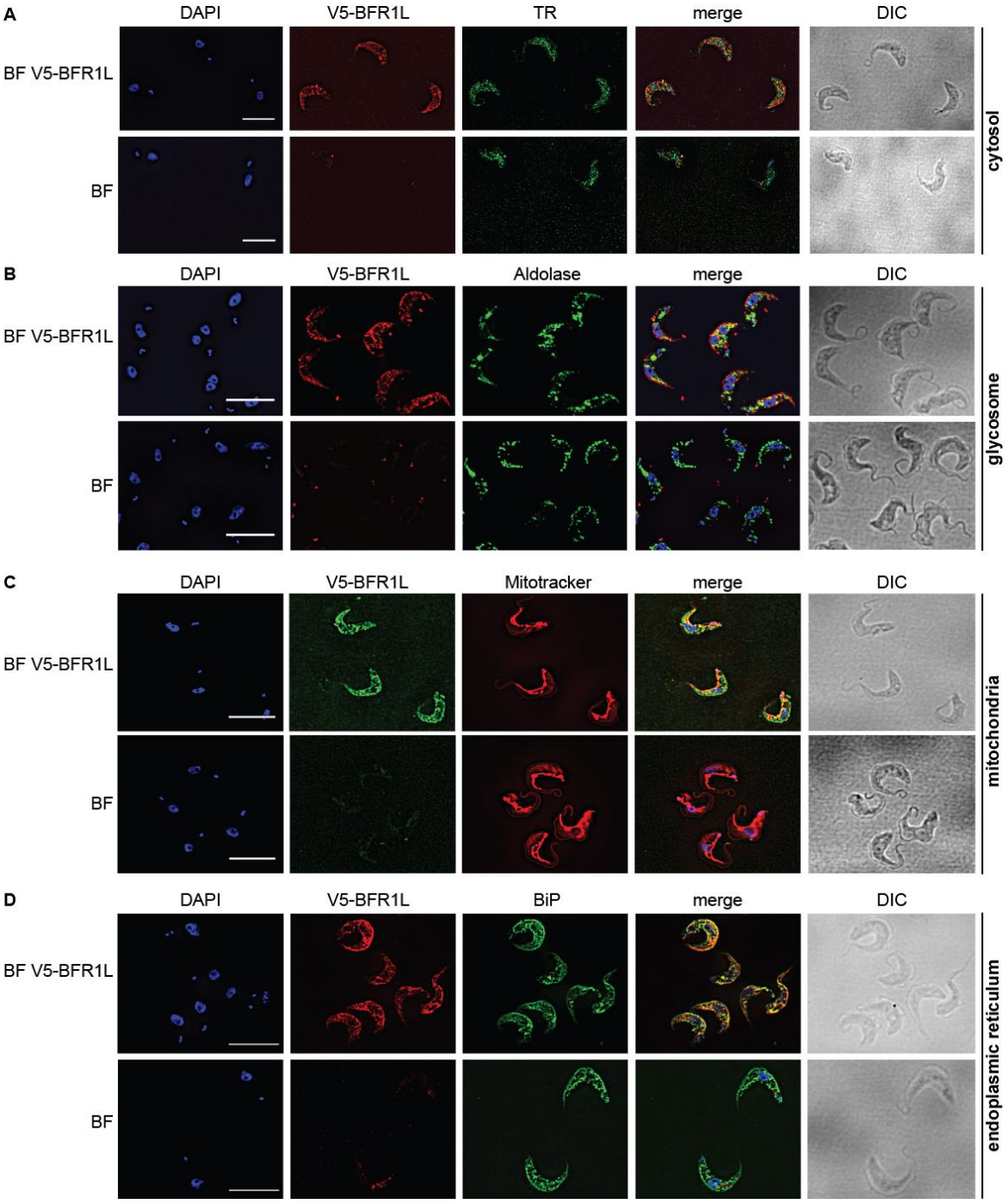


Figure 2.4: Fluorescence microscopy of BF V5-BFR1L cells. BF V5-BFR1L cells were subjected to fluorescence microscopy. BF cells without V5 served as control. Nuclear and kinetoplast DNA was stained with DAPI (blue) in all four immunofluorescence sets. **A.** Parasites were treated with antibodies against V5-tag (red) and Trypanothione Reductase (=TR) (green). **B.** Parasites were treated with antibodies against V5-tag (red) and the glycosomal Aldolase (green). **C.** Parasites were treated with Mitotracker (red), which stains the mitochondria, and with an antibody against V5-tag (green). **D.** Parasites were treated with antibodies against V5-tg (red) and the ER protein BiP (green). All four immunofluorescence sets are representative for three independent experiments. Z-stacks were examined using the Olympus CellR microscope and 100 x magnifications. Images were deconvoluted and one image was chosen for this figure (Scale bar: 10 µm).

I could not observe any co-localization. In addition, I investigated the localization of V5-tagged BFR1L and Mitotracker, which stains the mitochondria (Figure 2.4 C). It looks like there is some co-localization at the membranes of the parasite. As seen in figure 2.4 D V5-tagged BFR1L co-localizes with BiP (binding immunoglobulin protein), a marker of the endoplasmic reticulum. It can be shown that the ER, as expected, formed a continuous network through the cells. Although BFR1L doesn't seem to have a mitochondrial localization signal or a signal peptide according to localization site prediction tools (PSORTII, MitoPro and TargetP), the N-terminal V5-tag could influence the localization of the protein. For that purpose, the immunofluorescence microscopy was also done with a cell line expressing C-terminally myc-tagged BFR1L from the endogenous locus. As described above co-localization of myc-tagged BFR1L with TR, Aldolase, Mitotracker and BiP was investigated (Figure 2.5). I could not observe a co-localization with TR (Figure 2.5 A). In contrast to the immunofluorescence microscopy results of the N-terminally V5-tagged protein, the C-terminally tagged protein showed partial co-localization with Aldolase (Figure 2.5 B) as well as with Mitotracker (Figure 2.5 C) and BiP (Figure 2.5 D).

In addition, the localization of BFR1L was determined by digitonin titration. Digitonin is a non-ionic detergent that has a high affinity for cholesterol and permeabilizes cell membranes. Cells with N-terminally or C-terminally tagged BFR1L were pelleted and treated with increasing digitonin concentrations. The supernatants and pellets were collected and samples were analyzed by Western blotting (Figure 2.6). TR was used as cytoplasmic marker; Aldolase was used as glycosomal marker; LipDH (Lipoamide dehydrogenase) was used as mitochondrial marker and BiP was used as ER marker. V5-BFR1L (Figure 2.6 A), as well as, BFR1L-myc (Figure 2.6 B) could already be found in the early fractions of the supernatant, which suggested that BFR1L was located in the cytosol, and the amount of protein increased going to fractions with higher digitonin concentrations. Nevertheless, TR, which was used as cytoplasmic marker, was almost completely in the supernatant fractions and could not be found in the pellet fractions, whereas BFR1L could also be found in the pellet fractions, especially in the ones from lower digitonin:protein ratios. In contrast, only very low amounts of Aldolase protein could be found in the supernatant fractions with low digitonin:protein ratios and the same was seen for LipDH and BiP: only very low amounts of protein could be found in the supernatant fractions with low digitonin:protein ratio. However, more LipDH and BiP could be found in the pellet fractions than in comparison to Aldolase. The localization of N- and C-terminally tagged BFR1L in the pellet fractions suggests partial co-localization with membranes.

Taken together, no co-localization of BFR1L with DAPI and Aldolase, but partially with TR and Mitotracker could be observed according to immunofluorescence microscopy and digitonin titration. A clear co-localization with BiP could be observed, which suggests that the protein co-localizes fully with the ER and partially with mitochondria. These data agree with the localization of BFR1L to the ER in the procyclic form according to the TrypTag database and with the data about yeast BFR1p, which localizes to the ER under normal growth conditions (Lang et al., 2001). It could also be that BFR1L does not go into the organelles, but attaches to the membranes of the organelles from the cytosolic side and thereby it looks like partial co-localization.

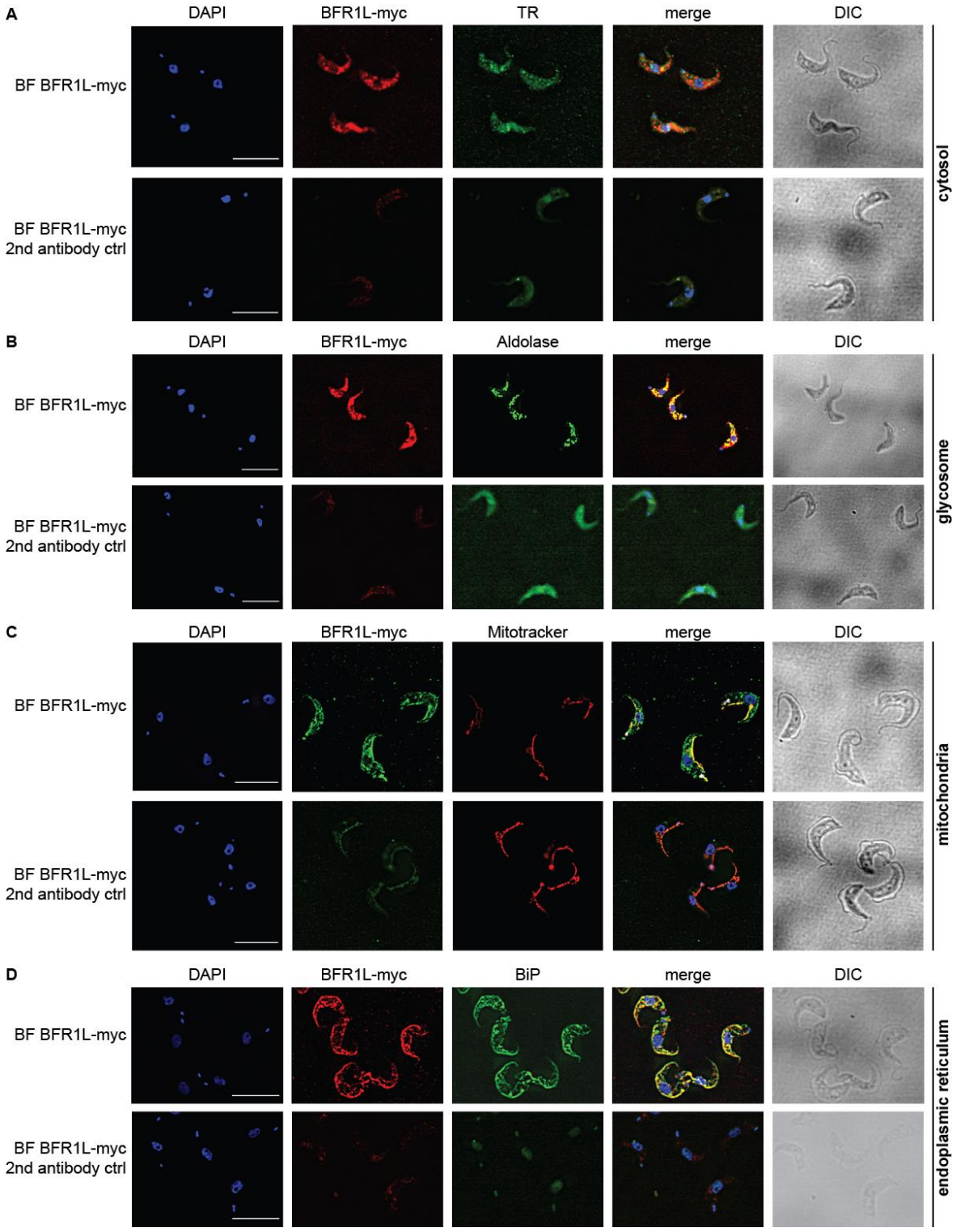


Figure 2.5: Fluorescence microscopy of BF BFR1L-myc cells. BF BFR1L-myc cells were subjected to fluorescence microscopy. Nuclear and kinetoplast DNA were stained with DAPI (blue) in all four immunofluorescence sets. **A.** Parasites were treated with antibodies against V5-tag (red) and the cytoplasmic Trypanothione Reductase (=TR) (green). **B.** Parasites were treated with antibodies against V5-tag (red) and the glycosomal Aldolase (green). **C.** Parasites were treated with Mitotracker (red), which stains the mitochondria, and antibodies against V5-tag (green). **D.** Parasites were treated with antibodies against V5-tag (red) and the ER protein BiP (green). All four immunofluorescence sets are representative for three independent experiments. Z-stacks were examined using the Olympus CellR microscope and 100 x magnifications. Images were deconvoluted and one image was chosen for this figure (Scale bar: 10 µm).

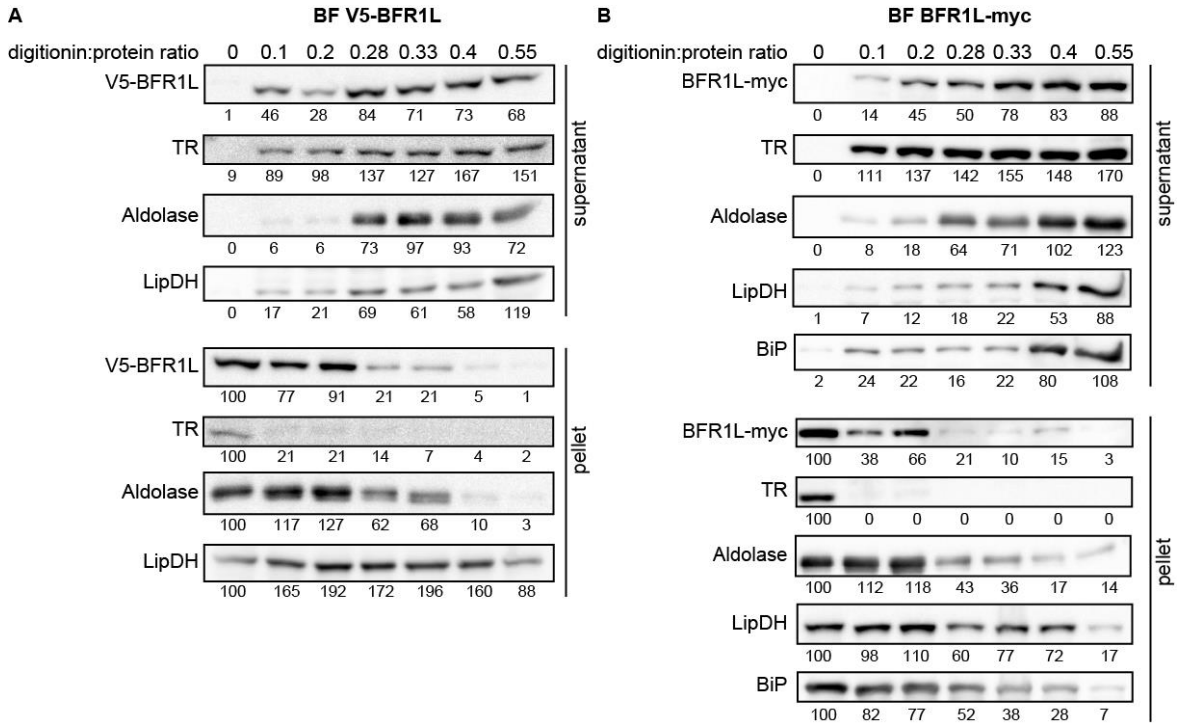


Figure 2.6: Digitonin titration of BF V5-BFR1L and BFR1L-myc. Bloodstream form cells with endogenous N-terminal V5-tagged BFR1L (**A**) or C-terminal myc-tagged BFR1L (**B**) were pelleted and treated with increasing digitonin:protein (mg:mg) ratios as indicated on top. Samples were centrifuged and fractions of supernatant and pellet were collected and Western Blotting was performed. The samples applied for Western Blotting correspond to $3x10^6$ cells. Numbers below the blot indicate quantification of the corresponding signal. TR: Trypanothione Reductase (cytosol); LipDH: Lipoamide Dehydrogenase (mitochondria); BiP: binding immunoglobulin protein (ER).

2.1.3. Protein interactome of BFR1L

In order to identify the protein interaction partners of BFR1L, which could give a hint how BFR1L acts as an RNA-binding protein, tandem affinity purification followed by mass spectrometry analysis was performed using *in situ* expressed TAP-BFR1L, while the other allele was deleted. Samples of all steps of the affinity purification were analyzed by Western Blotting. Figure 2.7 A shows a representative blot. The first purification step shows that cleavage of the TAP-tagged protein by TEV protease released the protein from the beads. However, it could not be detected in the elution, because the tag has been removed. Aldolase was used as negative control, which should not bind to the beads. Since Aldolase is so abundant a small amount was bound to the beads. However, the majority could be found in the unbound fraction. In addition, Aldolase should not be released form the beads after TEV cleavage. It can be seen that it was still bound to the beads after cleavage. I proceeded with a second purification step. As it can be seen by Western Blotting, the elution of CBP-BFR1L from the beads did work well, since the majority of the protein was found in the elution. The eluted protein and its interaction partners were then analyzed by mass spectrometry. Purification of a TAP-tagged ZC3H5 construct served as control. Liquid chromatography-tandem mass spectrometry (LC/MS/MS) of four independent TAP-BFR1L purifications revealed a total list of 4 putative interaction partners (Figure 2.7 B). BFR1L was also significantly enriched, which showed successful pull-down. The candidates were the following: Tb927.6.4300 (GAPDH), Tb927.10.2110 (EEF1A), Tb927.10.5620 (Aldolase) and Tb10.v4.0052 (microtubule-associated protein 2). Treatment with RNase A had no effect on any of these interactions (Figure 2.7 C), indicating that they were not RNA

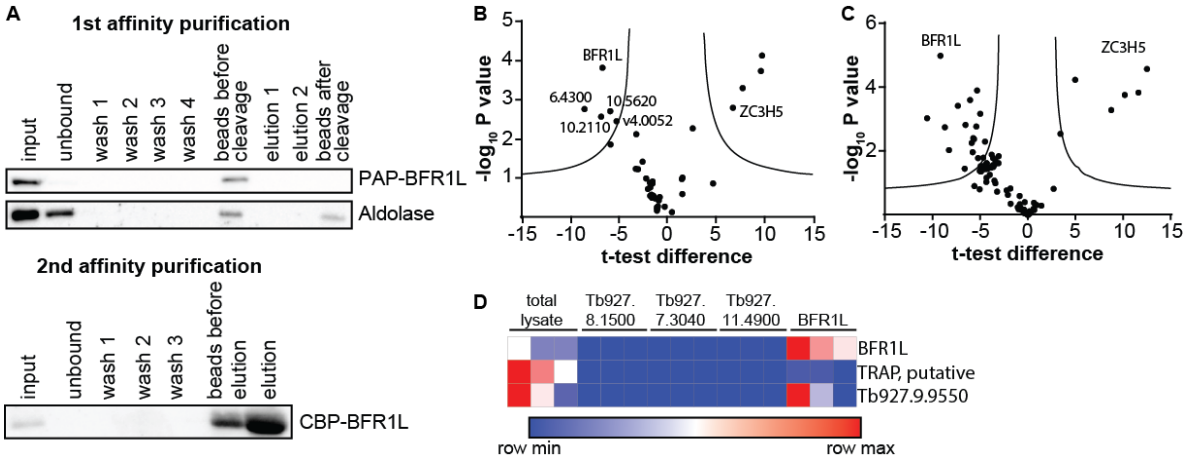


Figure 2.7: Affinity purification of TAP-BFR1L revealed one putative interaction partner. To identify proteins associated with the candidate protein, I performed Tandem Affinity Purification in a BFR1L single knock-out bloodstream form cell line using *in situ* expressed TAP-BFR1L. Associated proteins were analyzed by Mass Spectrometry and compared with TAP-ZC3H5 and TAP-GFP. **A.** 20 µl sample of each step was collected and applied for Western Blotting. **B &C.** In the volcano plot, the ratio of TAP-BFR1L to control (TAP-ZC3H5) in label-free quantification are plotted against the log10 of the false discovery rate (FDR) calculated by a permutation-based FDR adapted t-test. Significant outliers are labeled. B: without RNase A; C: with RNase A. Tb927.6.4300 (GAPDH), Tb927.10.2110 (EEF1A), Tb927.10.5620 (Aldolase) and Tb10.v4.0052 (microtubule-associated protein 2). **D.** Endogenously V5-tagged BFR1L and controls (Tb927.8.1500, Tb927.7.3040 or Tb927.11.4900) were purified three times and analyzed as described above. Heat map shows the spectral counts considering at least two out of 3 samples with one or more peptides detected. The hierarchical clustering was made with one minus Pearson's correlation. I choose these lines out of the whole analysis, because Tb927.9.9550 was the only putative candidate and all other putative interaction partners were abundant proteins, which are most likely unspecific.

dependent. The pull-down with RNase A gave a list of 16 putative interaction partners. However, the putative candidates in both sets of pull-downs were abundant proteins and it did not look like these interactions were specific. By setting a lower threshold, Tb927.9.9550 was detected as putative interaction partner by MS analysis. In addition, a pull-down of V5-BFR1L was performed and the elution was also analyzed by LC/MS/MS. Pull-downs of V5-Tb927.8.1500, V5-Tb927.7.3040 and V5-Tb927.11.4900 served as controls and also total lysate of BF was analyzed. The heat map was done with the spectral counts for all the proteins that are present in two out of three replicates. Tb927.9.9550 was identified in two out of three replicates and especially in the first replicate it had a high number of spectral counts (Figure 2.7 D). The heat map showed that Tb927.9.9550 is a specific putative interaction partner of BFR1L, because it couldn't be identified in any other of the performed pull-downs. The Tb927.9.9550 protein consists of 212 amino acids and has a predicted molecular weight of 24.5 kDa. The polypeptide contains a single transmembrane domain spanning residues 4-26 and two low complexity region (Suppl. Figure 2 A). It is conserved among Kinetoplastida. However, low sequence identity can be found with the Leishmania major orthologue (Suppl. Figure 3). According to the TrypTag website (Dean et al., 2017) a C-terminally tagged version of Tb927.9.9550 is located at the ER and the nuclear envelope (Suppl. Figure 2B) and it is not essential according to the RNA interference target sequencing (Alsford et al., 2011). Tb927.9.9550 protein displays *in vivo* mRNA binding (Lueong et al., 2016) although it lacks any canonical RNA-binding domain and was an up-regulator in the tethering screen (Erben et al., 2014b).

Since the protein interactome of TAP-tagged BFR1L revealed a very short list of putative interaction partners, a yeast-two-hybrid (Y2H) assay was performed to find putative new candidates. The protein was N-terminally fused with the binding-domain (BD) and transformed into yeast (bait). As prey I used yeast containing the mini-ORFeome library fused to the activating domain (AD). The bait and prey strains were then combined by mating. Unfortunately, I did not get any colonies (data not shown).

Since BFR1L was identified as an up-regulator in the high-throughput tethering screen (Erben et al., 2014), I performed CAT assays to validate the screen result. Expression of chloramphenicol acetyltransferase (CAT) was measured in cells expressing different myc-lambda-N-fusion proteins (BFR1L and Tb927.7.2780). Tb927.7.2780, which is a known activator of gene expression, served as positive control (Figure 2.8 A). Expression of the proteins was analyzed by Western Blotting (Figure 2.8 B). Indeed, the different proteins were expressed even though in different concentrations. Expression of the lambda-N proteins was induced with tetracycline for 24h. However, tethering of BFR1L to the CAT reporter did not confirm BFR1L as activator of gene expression.

Taken together, BFR1L might interact with one putative protein (Tb927.9.9550) and it may not be an activator of expression when tethered to an mRNA.

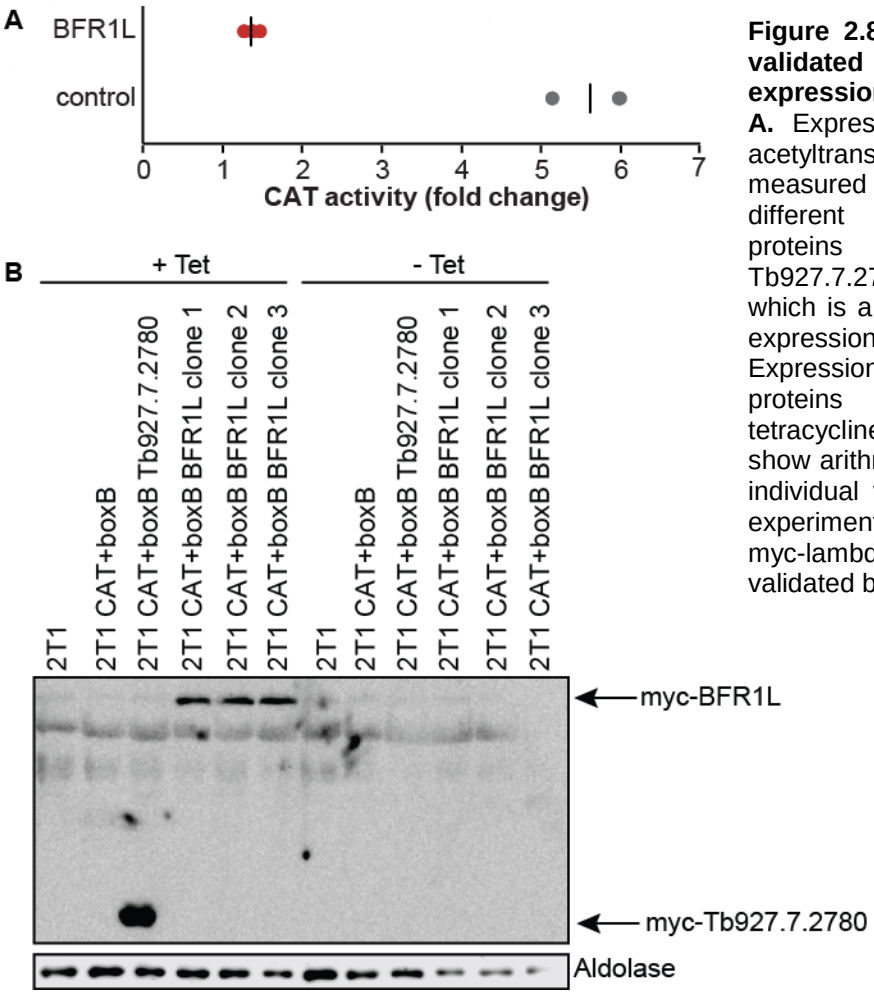


Figure 2.8: BFR1L could not be validated as activator of gene expression by CAT assay. A. Expression of chloramphenicol acetyltransferase (CAT) was measured in cells expressing different myc-lambda-N-fusion proteins (BFR1L and Tb927.7.2780). Tb927.7.2780, which is a known activator of gene expression, served as control. Expression of the lambda-N-myc proteins was induced with tetracycline (+Tet) for 24h. Results show arithmetic mean (black bar) & individual values of 3 independent experiments. **B.** Expression of the myc-lambda-N-fusion proteins was validated by Western Blotting.

2.1.4. Most of the mRNAs associated with BFR1L encode ribosomal proteins

For the identification of the mRNA targets of BFR1L, Affinity Purification was performed from cells expressing TAP-tagged BFR1L. Samples of all steps of the affinity purification were analyzed by Western Blotting as described above (Figure 2.9 A). The Western Blot showed that the protein was released from the beads. I directly analyzed the RNAs of this elution and did not proceed with the second purification step to avoid RNA degradation and because this method has worked for several other proteins in our lab. The sequences of associated RNAs were determined by RNA-sequencing and compared with total RNA. The experiment was performed in duplicates. However, the library preparation of one unbound sample failed, so the eluted RNA was compared with the WT total RNA. Principle component analysis showed that the elution samples clustered as well as the WT samples. Surprisingly, the unbound and WT samples did not cluster well. However, they are still clearly separated from the elution samples (Figure 2.9 B). I could identify 46 bound transcripts that were at least 2-fold enriched compared to total RNA. Strikingly, more than 50% of the bound transcripts encoded ribosomal proteins (Figure 2.9 C). According to the Fisher's exact test, the class 'Ribosome' was strongly enriched (p-value of ~ 2.2e-16). DREME analysis (Bailey, 2011) for the identification of a common binding motif in the 5'UTR, CDS or 3'UTR of the targets was negative. It is known that mRNAs encoding ribosomal proteins have relatively low ribosome densities although they are abundant and very stable (Antwi et al., 2016). The regulation of mRNAs encoding ribosomal proteins after stress is very different from that of most RNAs, since they are neither found in starvation stress granules (Fritz et al., 2015) nor in heat-shock granules (Minia et al., 2016). Thus,

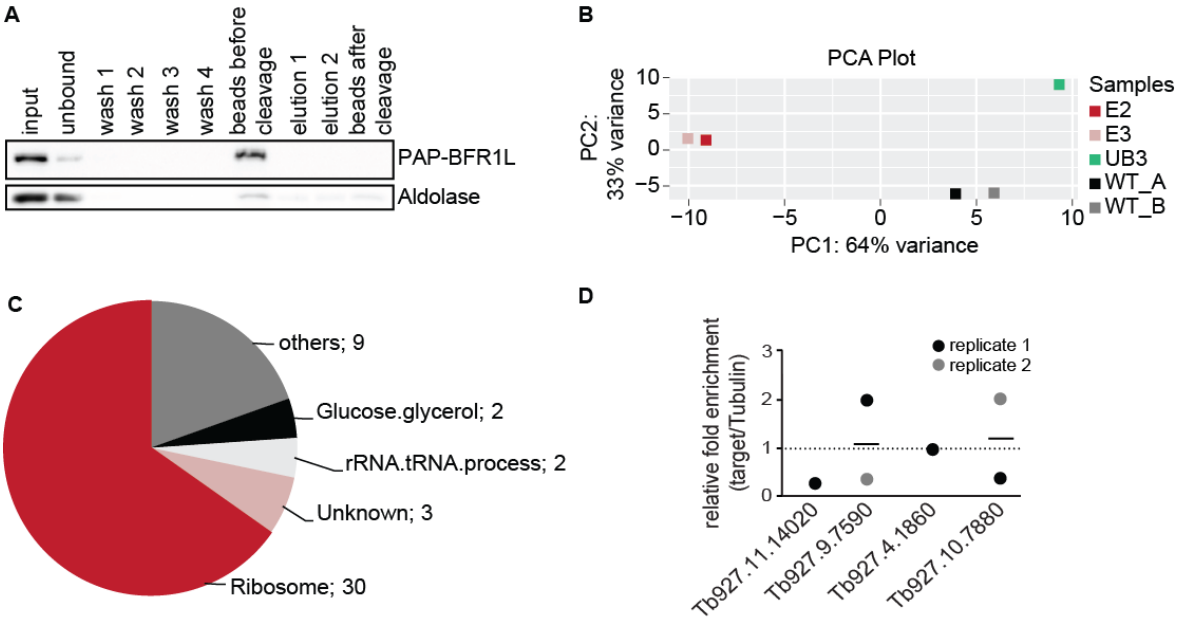


Figure 2.9: Most of the mRNAs associated with TAP-BFR1L encode ribosomal proteins. To identify RNAs bound by BFR1L, Affinity Purification in a BFR1L$^{-/TAP}$ bloodstream form cells was performed. The bound RNA was purified and analyzed by RNA-seq. Bound RNA was compared with total RNA. **A.** 20 µl sample of each step was collected and applied for Western Blotting. **B.** Principle component analysis (PCA) of the eluted (E), unbound (UB) and wild type (WT) RNA in duplicates (KL). **C.** Functional categories enriched in RIP-Seq data. **D.** Unbound and eluted RNA was analyzed by RT-qPCR. Expression fold change was calculated and normalized to tubulin. Results show mean (black bar) and individual values of 1-2 independent experiments.

BFR1L might keep its target mRNAs attached to the ribosomes and prevent sequestration in granules.

I attempted to validate the RIP-Seq results by RT-qPCR. TAP-BFR1L and its bound mRNA targets were purified by Tandem Affinity Purification and RNA was purified from the unbound and eluted samples. The RNA was reverse transcribed into cDNA and qPCR was performed to amplify the cDNA of the target mRNAs (Figure 2.9 D). Three targets that appeared to bind to TAP-BFR1L were used: Tb927.11.14020 (nuclear RNA-binding domain 2), Tb927.9.9750 (60S ribosomal protein L11, putative) and Tb927.4.1860 (putative ribosomal protein S19). Tb927.10.7880 (putative Sperm tail C-terminal domain containing protein), which was not enriched in the RIP-Seq data, was used as negative control. Ct values were normalized to Tubulin and expression fold change was calculated. However, all putative mRNA targets as well as Tb927.10.7880, which served as negative control, were not enriched in the elution.

Taken together, RIP-Seq of BFR1L identified 46 putative mRNA targets and most of the mRNAs were associated with BFR1L encode for ribosomal proteins. Unfortunately, the three tested putative mRNA targets could not be validated by qPCR.

2.1.5. BFR1L does not colocalize with the stress granule marker Scd6 under starvation stress

Next, I investigated whether BFR1L goes to stress granules, since I hypothesized that the protein might prevent the sequestration of its attached mRNAs in granules by keeping them away from stress granules. The experiment was performed in procyclic cells, because the work flow for the stress granule purification was optimized for these cells (Fritz et al., 2015) and the expression of BFR1L protein is the same in procyclic and bloodstream form cells (Dejung et al., 2016). For that reason, procyclic cells expressing *in situ* V5-tagged BFR1L were treated with starvation stress (cells were transferred to 1x PBS for 2h). Stress granules and localization of BFR1L was investigated by immunofluorescence microscopy (Figure 2.10 A; (PH)). The known stress granule marker Scd6 was used as control. After 2 h of PBS treatment Scd6 localized in stress granules in PC trypanosomes, whereas V5-BFR1L was still distributed through the cytosol. In addition, starvation stress granule purification was performed as described by Fritz *et al.* (Fritz et al., 2015) (Figure 2.10 B). The protein content of the different pellet and supernatant fractions was analyzed by Western Blotting (Figure 2.10 C). Known stress granule markers, like DHH1 and SCD6, were used as controls. Both proteins could be found in the final granule enriched fraction (P4), whereas V5-BFR1L was not there. BiP, an ER marker, which should be absent from the final granule enriched fraction was there as a contaminant, but nevertheless it was clear that V5-BFR1L does not localize to stress granules.

The role of BFR1L in the stress response was explored in more detail. The BF BFR1L$^{-/-}$ cells (DKO) were stressed by starvation (2h in 1x PBS) and recovery was compared with BF cells (WT) and the conditional knock-out (cDKO) with (+Tet) and without (-Tet) tetracycline induction. Expression of myc-BFR1L was monitored by Western Blotting. As described above, I could see a growth defect of the bloodstream form DKO cell line under normal growth conditions in one experiment and the cDKO cells grew as the WT in both cases: with and without tetracycline induction (see Figure 2.3 A). However, after 2h of starvation stress, the DKO grew as the WT and as the cDKO (Figure 2.11 A). Since the relative protein abundance of BFR1L is the same in both stages (Dejung et al., 2016), I also

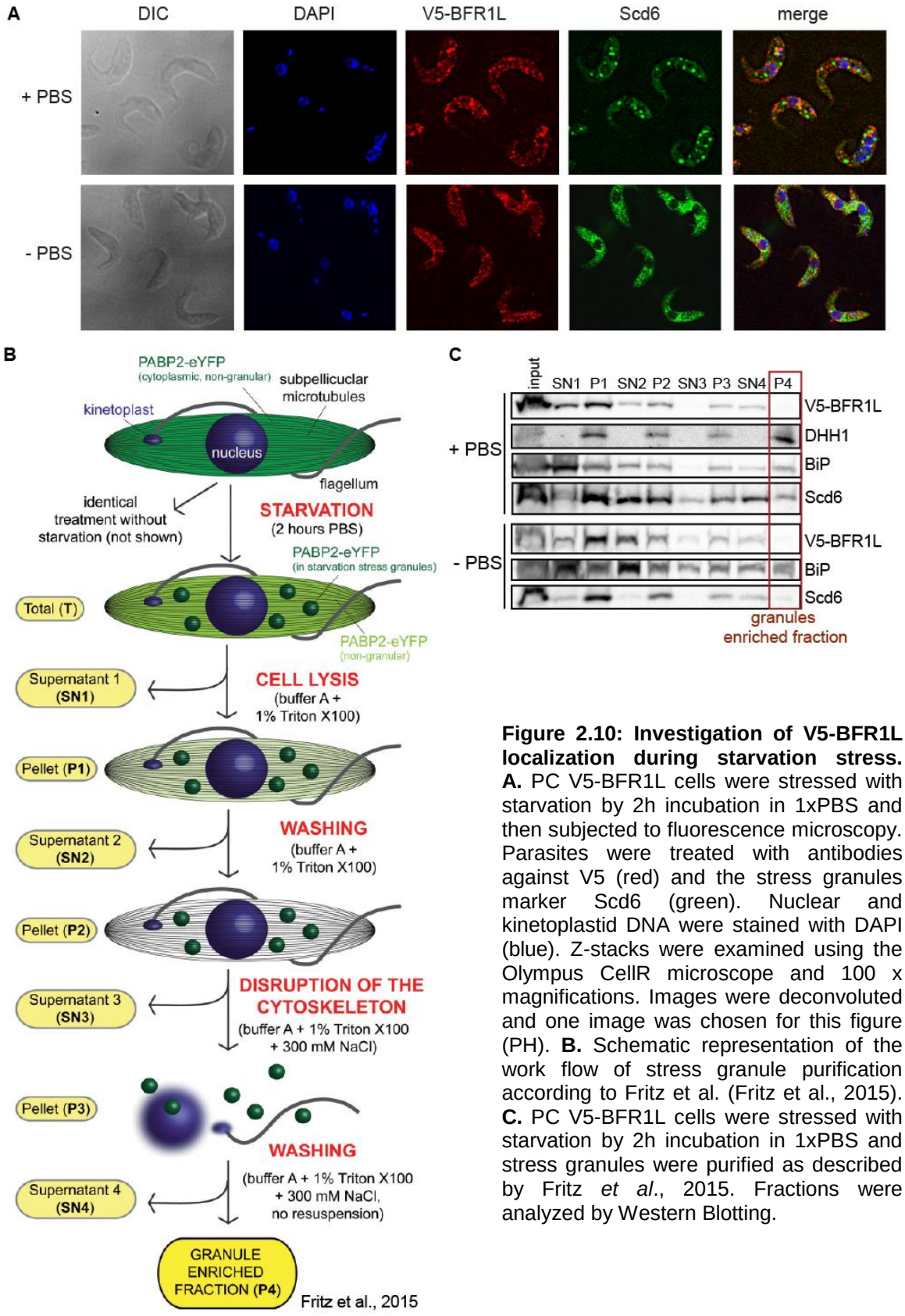


Figure 2.10: Investigation of V5-BFR1L localization during starvation stress. A. PC V5-BFR1L cells were stressed with starvation by 2h incubation in 1xPBS and then subjected to fluorescence microscopy. Parasites were treated with antibodies against V5 (red) and the stress granules marker Scd6 (green). Nuclear and kinetoplastid DNA were stained with DAPI (blue). Z-stacks were examined using the Olympus CellR microscope and 100 x magnifications. Images were deconvoluted and one image was chosen for this figure (PH). **B.** Schematic representation of the work flow of stress granule purification according to Fritz et al. (Fritz et al., 2015). **C.** PC V5-BFR1L cells were stressed with starvation by 2h incubation in 1xPBS and stress granules were purified as described by Fritz *et al*., 2015. Fractions were analyzed by Western Blotting.

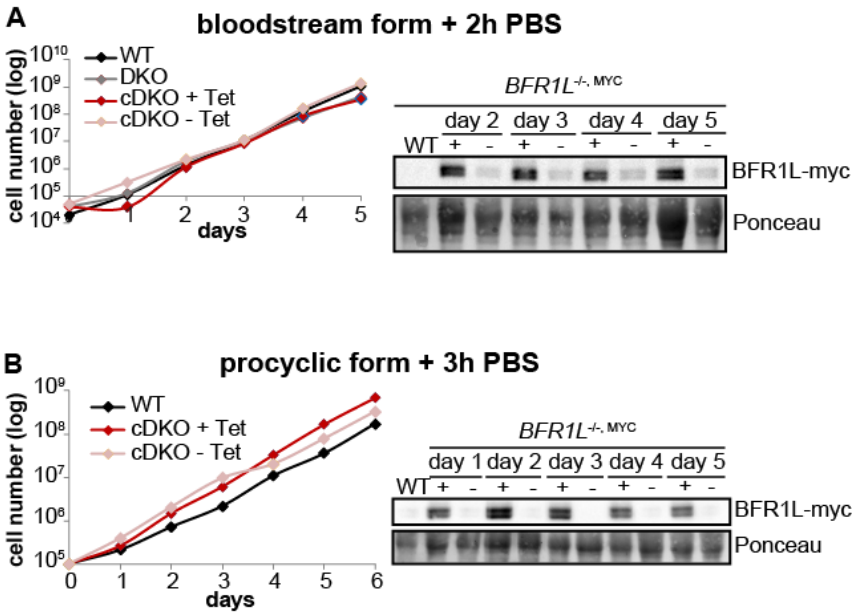


Figure 2.11: Growth of bloodstream form cells and upon double-knockout of *BFR1L* and starvation stress. A. Growth curve for BF (WT), DKO (BF *BFR1L*$^{-/-}$) and cDKO (BF *BFR1L*$^{-/-, MYC}$, +Tet and -Tet) after 2h of starvation stress in 1x PBS. 3x10^6 cells were collected each day after counting and Western Blot of myc-BFR1L was performed. Ponceau staining was used as loading control. **B.** Experiment was performed as described in A using PC (WT) and cDKO (PC *BFR1L*$^{-/-, MYC}$, +Tet and -Tet) cells.

examined the growth kinetics in the procyclic form. As described above, the procyclic cDKO cells grew as the WT, no matter if ectopic expression of myc-BFR1L was induced by tetracycline (+Tet) or not (-Tet) (see figure Figure 2.11 C). In addition, starvation stress (3h in 1x PBS) did not influence the growth kinetics of the cDKO cells (Figure 2.11 D).

Taken together, it seems like BFR1L does not go to stress granules in PCs after 2h starvation stress in 1x PBS and starvation stress does not influence the growth of procyclic cDKO cells. However, BF DKO cells grow as the WT upon starvation stress, whereas they show a growth defect without starvation.

2.2. Discussion

Regulation of gene expression in *T. brucei* mainly depends on post-transcriptional mechanisms. To identify putative trypanosome post-transcriptional regulators a tethering assay was performed (Erben et al., 2014b). The aim of this study was to characterize two candidates in more detail. I choose candidates which were shown to be repressors or activators in the tethering screen (Erben et al., 2014b; Lueong et al., 2016), because this suggests a role in the regulation of gene expression. In addition, the candidate should bind to mRNA by interactome capture (Lueong et al., 2016) as I was interested in the regulation of gene expression by RBPs; and the protein should be located in the cytosol, because this would be compatible with its RNA-binding activity. The candidate should also be essential according to RIT-Seq (Alsford et al., 2011), because that suggests an important role in the cells. Tb927.10.14150 was chosen as candidate because it was an activator in the tethering screen (Erben et al., 2014b; Lueong et al., 2016), and it showed *in vivo* mRNA binding and was one of the top-ranking hits (Lueong et al., 2016), although it lacks a canonical RNA-binding domain. The protein was named as BFR1L, because of its similarities to the yeast protein Bfr1p. Together with its interacting protein Scp160, Bfr1p binds to mRNAs at polysomes under normal growth conditions and protects these mRNAs from P-body formation (Weidner et al., 2014). However, there is no homologue of Scp160 in trypanosomes. Trypanosome BFR1L is not localized in stress granules (Fritz et al., 2015). This led to the hypothesis that BFR1L could prevent sequestration of its target mRNAs in granules, which was investigated in more detail in this study.

I could demonstrate that knock-down or knock-out of BFR1L in BF trypanosomes leads to a slight growth defect, which can be rescued by ectopic expression of the protein, whereas I was neither able to generate an RNAi-mediated knock-down cell line nor to generate a DKO cell line of BFR1L in PC trypanosomes. However, I was able to generate a cDKO in PCs. Surprisingly, growth was not restricted by removing of tetracycline from these cells and cells grew as wildtype even without ectopic expression of myc-tagged BFR1L. These results suggest that BFR1L is not essential neither in the bloodstream form nor in the procyclic form.

The localization of the protein according to TrypTag was not known at the beginning of this study. This is why I did immunofluorescence microscopy and digitonin titration, which showed that the protein does not co-localize with DAPI and Aldolase, but partially with TR and Mitotracker. A clear co-localization with BiP could be observed, which suggests that the protein co-localizes with the ER and partially with mitochondria. In the digitonin titration assay BFR1L was detected in the early fractions of the supernatant, like TR, which is a cytosolic protein. However, BFR1L can also be found in the later fractions of the pellet, like LipDH and BiP. These data agree with the localization of BFR1L to the ER according to the TrypTag database and with the data about yeast BFR1p, which localizes to the ER under normal growth conditions (Lang et al., 2001). Most probably BFR1L is located in the cytosol and does not go into the organelles, but attaches to the membranes of the organelles from the cytosolic side.

To get a clearer idea about how BFR1L could function, I was interested in the protein interaction partners as well as in the target mRNAs. To identify the interaction partners of BFR1L I used an endogenously TAP-tagged BFR1L version as well as a V5-tagged BFR1L version and IPs were performed. Mass spectrometry analysis revealed Tb927.9.9550 as a putative interaction partner. Tb927.9.9550 is a protein of unknown function, which is

conserved among Kinetoplastida. The polypeptide contains a single transmembrane domain and two low complexity regions. According to the TrypTag website (Dean et al., 2017) a C-terminally tagged version of Tb927.9.9550 is located at the ER and the nuclear envelope. Preliminary immunofluorescence microscopy data show a co-localization of BFR1L and Tb927.9.9550 at the ER (data not shown), which suggests that the two proteins could act in concert at the outside of the ER. Tb927.9.9550 is not essential according to RIT-Seq (Alsford et al., 2011). The protein displays *in vivo* mRNA binding (Lueong et al., 2016) although it lacks any canonical RNA-binding domain and was an up-regulator in the tethering screen (Erben et al., 2014b). No homologue of Tb927.9.9550 is known in yeast. One possibility is that the interaction between Tb927.9.9550 and BFR1L depends on their interactions with mRNAs and they could stabilize their targets. However, the function of BFR1L as up-regulator of gene expression, which was suggested by the tethering screen (Erben et al., 2014b; Lueong et al., 2016), could not be validated by the CAT assay. In the tethering screen proteins were fused N-terminally to the lambda-N peptide and inducibly expressed in a trypanosome cell line that contains the Blasticidin reporter construct. The cells were grown under non-inducing (tet -) and inducing (tet +) conditions with increasing Blasticidin conditions. If the tethered proteins increase reporter gene expression, the cells should be resistant to higher Blasticidin concentrations as the control cells. One possible explanation why BFR1L was suggested to be an activator of gene expression in the tethering screen could be that overexpression of the protein makes the cells more resistant to Blasticidin. Blasticidin inhibits the termination of translation and thereby stops the *de novo* synthesis of proteins. Initially, Bfr1p was discovered in the 90s in a genetic screen for high-copy suppressors of Brefeldin A (BFA) in *Saccharomyces cerevisiae*. BFA, which is a fungal toxin, influences the function and structure of the organelles of the secretory pathway (Klausner et al., 1992; Lippincott-Schwartz, 1993). It leads to tubulation of the lysosome, to fusion of the trans-Golgi network with the endosomes and to tubulation and fusion of the Golgi with the ER (Hunziker et al., 1992; Klausner et al., 1992; Pelham, 1991). Bfr1p can partially suppress the phenotypes caused by BFA (Jackson and Kepes, 1994). However, the mechanisms of action of BFA and Blasticidin are not related. BFR1L could partially suppress the phenotype caused by Blasticidin. The resistance of trypanosomes overexpressing BFR1L to Blasticidin should be investigated in the future.

To identify the mRNA targets of BFR1L, I used an endogenously TAP-tagged BFR1L version. RIP-Seq identified 46 putative mRNA targets and most of them encode ribosomal proteins. Unfortunately, the three tested putative mRNA targets could not be validated by qPCR. DREME analysis (Bailey, 2011) of the BFR1L mRNA targets could not identify any motif in the 5'UTR, CDS or 3'UTR. It is known that mRNAs encoding ribosomal proteins have relatively low ribosome densities although they are abundant and very stable (Antwi et al., 2016). In addition, they have an absolutely optimum codon usage (de Freitas Nascimento et al., 2018). The ribosomes may run through very fast giving a low ribosome density. The regulation of mRNAs encoding ribosomal proteins after stress is very different from that of most RNAs, since they are neither found in starvation stress granules (Fritz et al., 2015) nor in heat-shock granules (Minia et al., 2016). According to Fritz and coworkers, almost half of the identified target mRNAs of BFR1L are not enriched in the granules fraction (Fritz et al., 2015). It is known that BFR1L and its protein interaction partner Tb927.9.9550 do not localize to stress granules (Fritz et al., 2015). BFR1L might keep the enriched mRNAs attached to the ribosomes and thus prevent sequestration in granules as it was shown for Bfr1p. Thereby it could have a stress-related function.

To investigate the role of BFR1L under stress conditions, procyclic cells were stressed with 2h starvation and stress granules were purified. The experiment was performed in procyclic cells, since the stress granules purification protocol is optimized for this form (Fritz et al., 2015) and BFR1L has the same relative protein abundance in both stages (Dejung et al., 2016). It seems like BFR1L does not go to stress granules in PCs. This strengthens the hypothesis that BFR1L prevents the sequestration of its target mRNAs to granules. Interestingly, BF DKO cells grow as the WT upon starvation stress, whereas they show a growth defect without starvation.

Taken together, BFR1L as well as its putative interaction partner Tb927.9.9550 are suggested to be located in the cytosol and attach to the membrane of the ER, from the cytosolic side, which suggests that the proteins post-transcriptionally regulate mRNAs at the ER. However, only Tb927.9.9550 has a transmembrane domain, which could be inserted in the ER. The targets of yeast Bfr1p were enriched for mRNAs encoding for proteins involved in cytoplasmic translation, containing ribosomal proteins, and membrane-associated functions (Lapointe et al., 2015). In addition, Bfr1p does not localize to P-bodies under normal growth conditions (Weidner et al., 2014), it is not essential (Jackson and Kepes, 1994), and it localizes to the ER (Lang et al., 2001). Lang and colleagues could show that Bfr1p is part of a polyribosome-associated mRNP complex, which contains Scp160p, Pap1p, additional unidentified proteins and polyadenylated mRNA (Lang and Fridovich-Keil, 2000; Lang et al., 2001). The localization of Bfr1p at the outside of the ER (Lang et al., 2001), the association of Bfr1p with polysomes (Weidner et al., 2014) and the enrichment of its mRNA targets for membrane-related proteins (Lapointe et al., 2015) suggests that the Bfr1p targets are directly translated to the ER. Usually mRNAs encoding proteins of the secretory pathway are translated into the ER in an SRP-dependent manner. However, recently it could be shown in several studies that ribosomes bound to the ER also translate mRNAs encoding cytosolic proteins (Reid and Nicchitta, 2015). In HEK293 cells, 75% of the ribosomes are bound to the ER and approximately 50% of the mRNAs encoding cytosolic proteins are bound to these ER-associated ribosomes (Jagannathan et al., 2014; Reid and Nicchitta, 2012). The recruitment of mRNAs to the ER can influence the translation rate due to a distinct environment of translation factors and regulatory proteins, which suggests a role of these ER-bound ribosomes in the regulation of post-transcriptional gene expression (Reid and Nicchitta, 2015). During oxidative stress, mRNAs localized to the ER can escape from sequestration into stress granules, while free mRNAs in the cytosol are recruited to stress granules (Unsworth et al., 2010). I identified mRNAs encoding ribosomal proteins as putative targets of BFR1L. My hypothesis is that BFR1L regulates the expression of its targets directly at the ER to produce ribosomal proteins. However, ribosome biogenesis in eukaryotes occurs in the nucleolus and the ribosomal proteins have to be imported into the nucleus after translation (Greber, 2016). The remaining question is: Why the mRNAs encoding ribosomal proteins could be translated at the ER? Parts of the rough ER are close to the nucleus, which would explain short ways. The ribosomal proteins could be imported into the nucleus directly for the ribosome assembly, but BFR1L co-localizes with the whole ER and not only around the nucleus. By binding of BFR1L to the target mRNAs, the mRNAs might be kept in ribosomes and thereby could not sequester into stress granules similar to the mechanism described for Bfr1p in yeast (Weidner et al., 2014). The attachment of BFR1L to the ER could be mediated via Tb927.9.9550, which has a transmembrane domain (Figure 2.12). The interaction of BFR1L with mRNAs encoding ribosomal proteins under stress should be investigated in the future by FISH. RIP-Seq of

BFR1L under stress should give a similar result as I obtained without stress, if BFR1L keeps the target mRNAs away from stress granules.

This study shows that BFR1L in Trypanosomes has some similarities to Bfr1p in yeast. Both proteins seem to keep their targets away from different kind of granules. However, the mechanism of action is different for the two proteins. While the targets of Bfr1p are suggested to be directly translated into the ER, the targets of BFR1L are suggested to be translated at the ER into the cytosol.

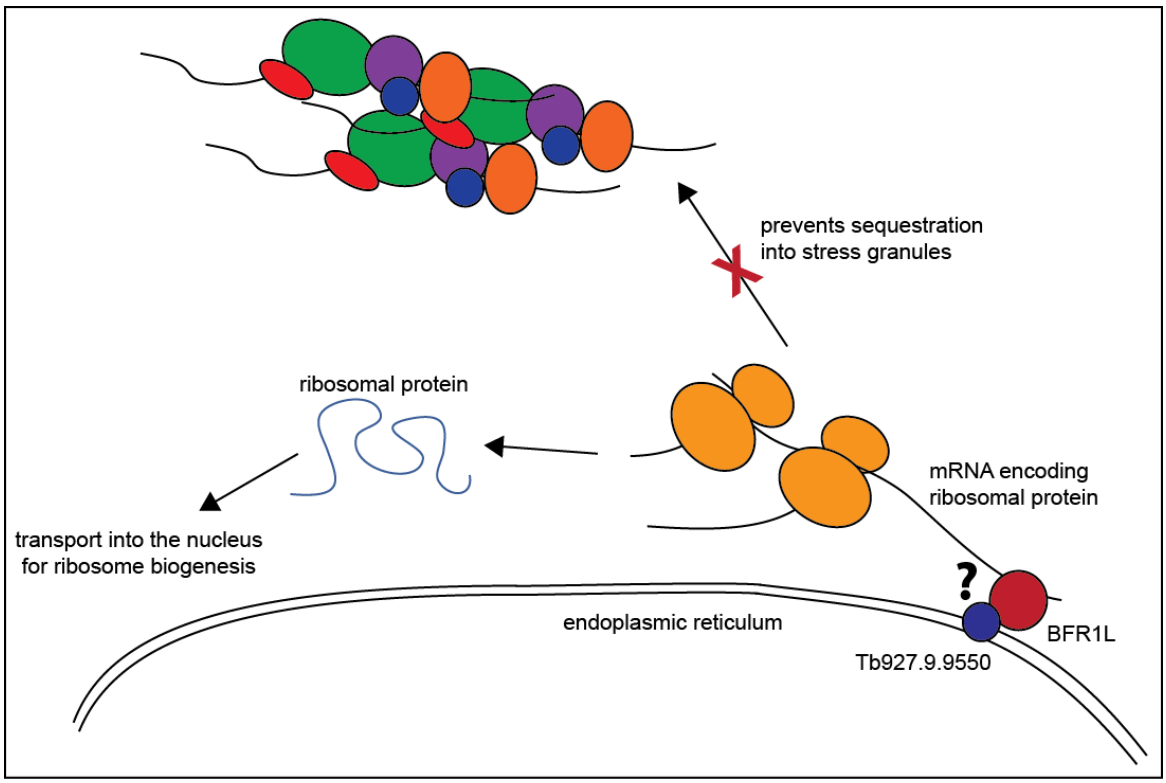


Figure 2.12: BFR1L might prevent sequestration of mRNAs encoding ribosomal proteins into stress granules. By binding of BFR1L to the target mRNAs, the mRNAs might be kept in ribosomes and thereby could not sequester into stress granules (containing proteins like DHH1, SCD6, PABP1, RBPs, etc.). The attachment of BFR1L to the ER could be mediated via Tb927.9.9550, which has a transmembrane domain.

3. ZC3H5 is required for cytokinesis

3.1. Results

3.1.1. Downregulation of ZC3H5 rapidly kills bloodstream form cells

The ZC3H5 (locus Tb927.3.740) protein consists of 246 amino acids with a predicted molecular weight of 25.5 kDa. The polypeptide contains a single C3H1-type zinc finger domain spanning residues 87-112 (Figure 3.1 A) and is conserved among Kinetoplastida (Suppl. Figure 4). Most sequence identity is concentrated around the zinc finger domain

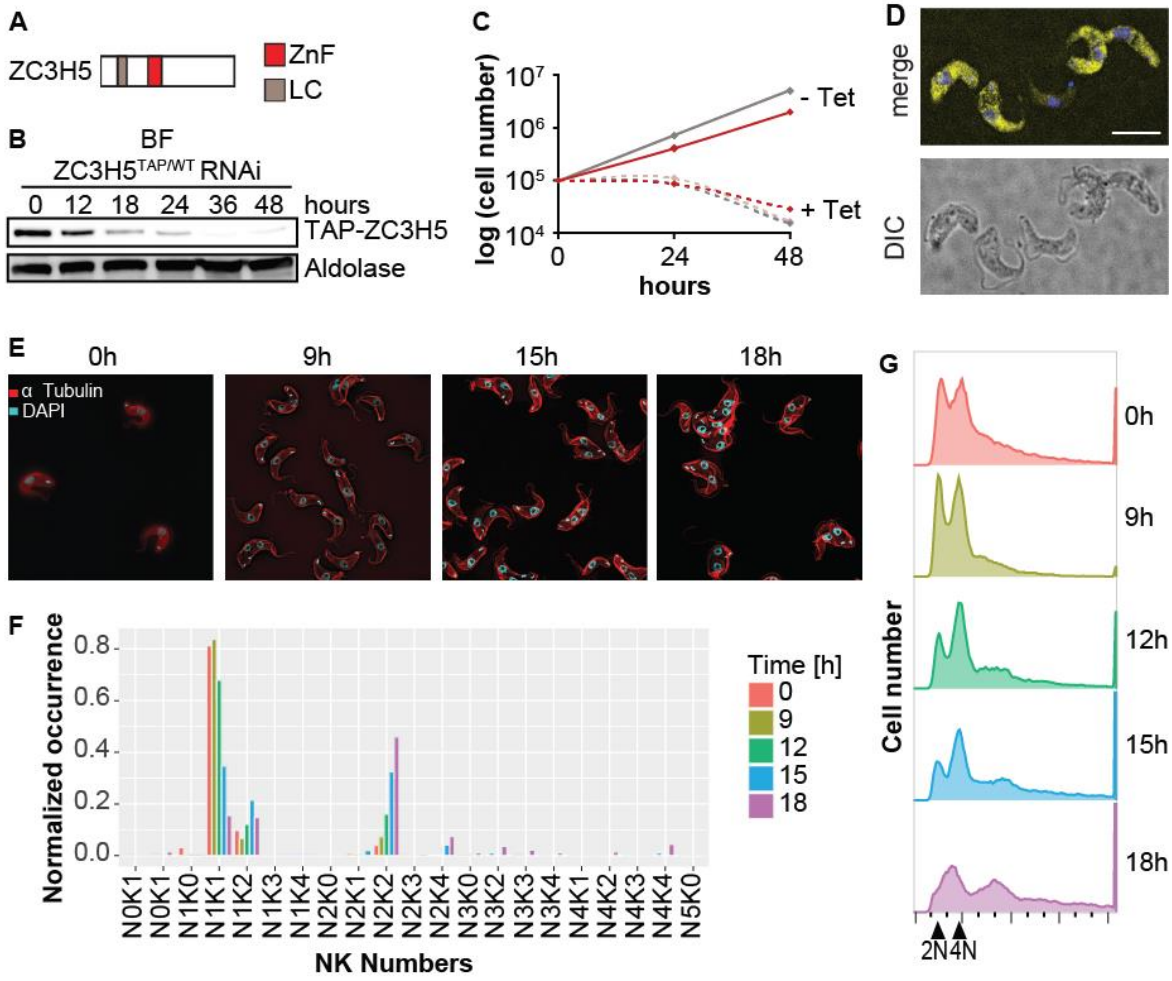


Figure 3.1: In vitro growth and cell cycle analysis for ZC3H5. A. Domains of ZC3H5. ZnF: Zinc finger domain, LC: Low complexity region. **B.** ZC3H5 RNAi cell line clones expressing TAP::ZC3H5 from the endogenous locus were analyzed by Western blotting with an anti-PAP antibody over time following tetracycline (Tet) induction. Anti-aldolase antibody was used as a loading control. **C.** Growth curve showing cumulative cell counts from three independent experiments over time following tetracycline (Tet) induction (+) or not (−) of ZC3H5 RNAi cell line in culture. Tet+ (dashed lines) and Tet- (solid lines) (KL). **D.** BF YFP-ZC3H5 cells were subjected to fluorescence microscopy. Nuclear and kinetoplast DNA were stained with DAPI (blue). Images were examined using the Leica DMi8 spinning disk microscope (Scale bar: 10 µm). **E.** ZC3H5 RNAi bloodstream form cells were analyzed by fluorescence microscopy. Kinetoplast and nuclear DNA was stained with DAPI and parasites were treated with antibodies against tubulin (red) (KL). **F.** The number of nuclei (N) and kinetoplasts (K) per cell of the cells described in E was quantified (n>200) at the time-points indicated (KL). **G.** Knock-down of ZC3H5 was induced for various time points, DNA was stained with Propidium iodide and analyzed by FACS.

and towards a proline-rich C-terminal region. To investigate whether ZC3H5 is essential, cell growth after RNAi-mediated knock-down of ZC3H5 was analyzed in a stable cell line in which one allele of ZC3H5 was tagged *in situ* with an N-terminal TAP-tag. This experiment was done in the bloodstream form (BF) as all the following experiments. Protein samples were collected at each time-point to monitor successful knock-down of ZC3H5 (Figure 3.1 B). Indeed, downregulation of ZC3H5 led to a cell growth arrest after 12-15 h of RNAi induction and killed the cells after 48 h (Figure 3.1 C; performed by Kevin Leiss (KL)), which indicates that ZC3H5 is essential in this form. According to my results endogenously expressed N-terminally YFP-tagged ZC3H5 is located in the cytosol (Figure 3.1 D), which was also observed in procyclic cells by the TrypTag website (Dean et al., 2017). YFP-ZC3H5 was equally distributed throughout the cytoplasm and no overlap with the DAPI signal of the nucleus and kinetoplast DNA was observed. The phenotype observed upon knock-down of ZC3H5 was also analyzed by fluorescence microscopy (Figure 3.1 E; KL). The nuclear and kinetoplast DNA were stained with DAPI and Tubulin staining was used to visualize the cell shape. After staining for DNA, cells were scored for different cell cycle stages: cells with a single nucleus and kinetoplast (1N1K) are in G1 or S phase, cells with two kinetoplasts and one nucleus (1N2K) are in G2 phase, and cells with two kinetoplasts and two nuclei (2N2K) are mitotic or post-mitotic. The number of cells with 2 nuclei and 2 kinetoplasts (2N2K) increased rapidly to approximately 45% of cells after 18 h of ZC3H5 knock-down (Figure 3.1 F; KL). In addition, even larger numbers of xNxK conformations could be observed upon knock-down of ZC3H5. To examine this phenotype in more detail, I did FACS analysis of ZC3H5 knock-down cells (Figure 3.1 G). Knock-down of ZC3H5 was induced for various times, DNA was stained with Propidium iodide and the cells were analyzed by FACS. In the uninduced cells the peaks of the G1-phase (2N) and G2/M-phase (4N) showed the same height. After 12h and 15h an increase of the second peak (G2/M-phase) could be seen. This showed that the number of cells with 2 nuclei and 2 kinetoplasts (2N2K) increased rapidly. 18h after induction of ZC3H5 knock-down a decrease of both peaks could be observed, because the cells were already dying.

I conclude that down-regulation of ZC3H5 rapidly kills bloodstream form trypanosomes. In addition, the proportion of cells in G2/M phase increased rapidly and cells were often arrested in the cytokinesis stage.

3.1.2. Proteins interacting with ZC3H5 form a complex

To identify interaction partners of ZC3H5 the cell line described above with endogenously TAP-tagged ZC3H5 was used for tandem affinity purification. The purified complexes were then analyzed by liquid chromatography-tandem mass spectrometry (Figure 3.2 A; performed by Esteban Erben (EE)). Purification of TAP-tagged GFP and TAP-tagged BFR1L, which are not related to ZC3H5, served as controls. The quantitative analysis revealed a list of 3 putative interaction partners encoded by: Tb927.11.4900, Tb927.7.3040 and Tb927.8.1500. Tb927.11.4900 protein consists of 560 amino acids with a predicted molecular weight of 61.9 kDa. The polypeptide contains four WD40-domains (Figure 3.2 B) and most sequence identity is concentrated around the WD40-domains (Suppl. Figure 5). Tb927.7.3040 protein consists of 635 amino acids with a predicted molecular weight of 69.1 kDa. The polypeptide contains three WD40-domains (Figure 3.2 B). However, there is not much sequence homology to the *Leishmania major* homologue (Suppl. Figure 6). Tb927.8.1500 protein consists of 601 amino acids with a predicted molecular weight of 63.2 kDa. The polypeptide contains several low complexity regions (Figure 3.2 B) and most

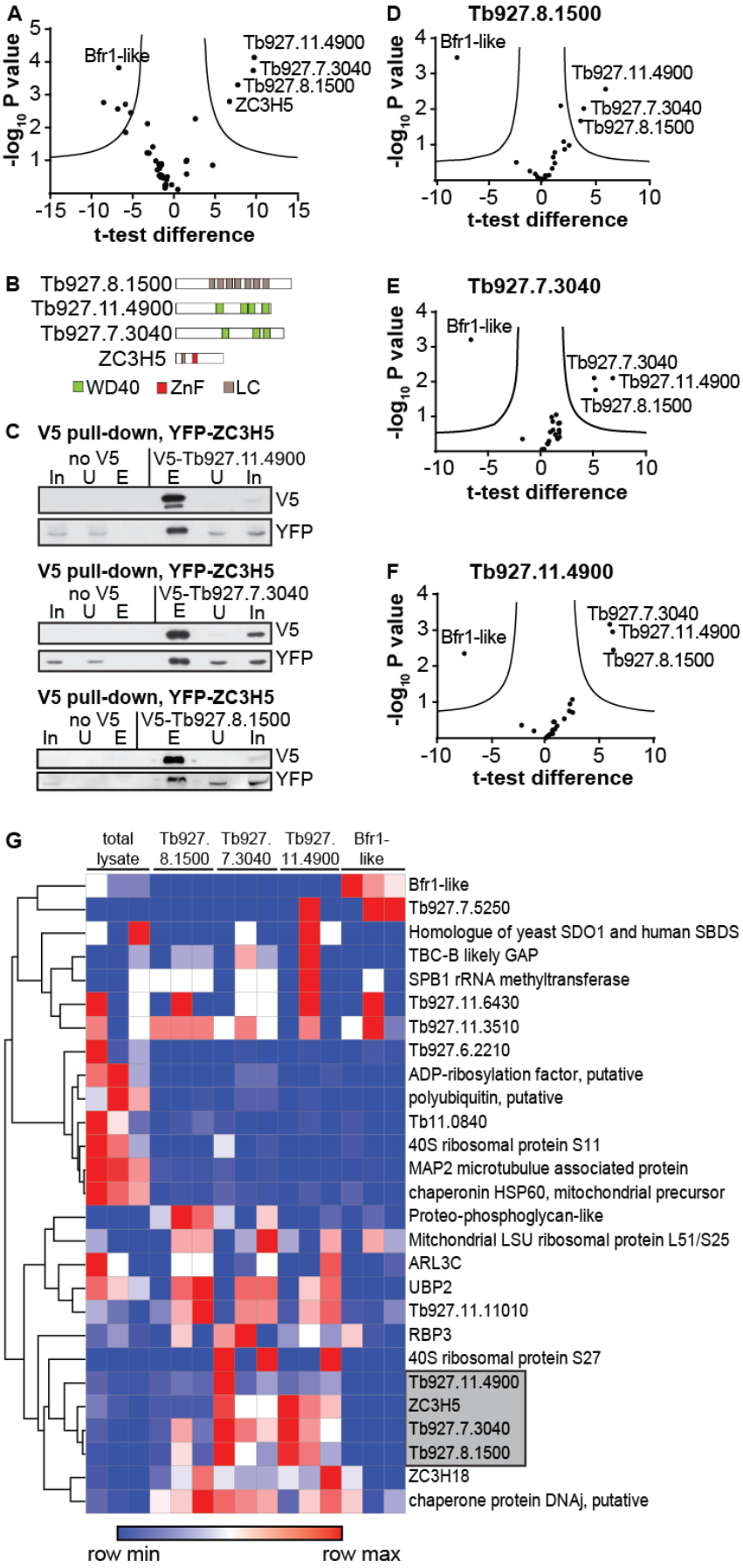


Figure 3.2: TAP-tagged ZC3H5 interacts with three proteins. **A.** Endogenously TAP-tagged ZC3H5 and controls (TAP-BFR1L and TAP-GFP) were purified three times. Raw data were analyzed by MaxQuant, and specific interactors were selected from background using label-free quantification in Perseus. In the volcano plots, the ratio of ZC3H5 to controls are plotted against the log10 of the false discovery rate (FDR) calculated by a permutation-based FDR adapted t-test. Significant outliers are labeled (EE). **B.** Conservation of the four proteins among different Kinetoplastid species was analyzed. **C.** Protein complexes were purified using an anti-V5 affinity matrix. 30 µg of input (In) and unbound (U) proteins and total amount of immuno-precipitated (E) proteins were analyzed by Western blotting (EE: Tb927.7.3040 & Tb927.11.4900 pull-down). **D.-F.** Endogenously V5-tagged proteins (Tb927.8.1500, Tb927.7.3040 or Tb927.11.4900) and control (Tb927.10.14150) were purified three times and analyzed as described above. **G.** Heat map of the samples described in D.-F. It shows the spectral counts considering at least two out of 3 samples with one or more peptides detected. The hierarchical clustering was made with Pearson's correlation. The grey box indicates ZC3H5 and its interaction partners.

sequence identity is concentrated in the N-terminal region (Suppl. Figure 7). According to the TrypTag website (Dean et al., 2017) an N-terminally tagged version of Tb927.11.4900 is located in the cytosol and the localization of the two other proteins is not determined yet. According to RIT-Seq Tb927.11.4900 and Tb927.8.1500 are essential whereas Tb927.7.3040 is not essential (Alsford et al., 2011). ZC3H5 and its interacting proteins are all conserved among Kinetoplastida. The interaction of these proteins with ZC3H5 was validated by co-immunoprecipitation using cells expressing endogenously YFP-tagged ZC3H5 and endogenously V5-tagged protein (Figure 3.2 C; partially EE). After pull-down of V5-Tb927.11.4900, V5-Tb927.7.3040 or V5-Tb927.8.1500, YFP-ZC3H5 was enriched in the eluted fraction, whereas it was not enriched in the elution of control cells without V5-tagged protein. Unfortunately, the cells lost the YFP-tagged ZC3H5 expression quickly. The interaction was RNA-independent, since RNase treatment did not influence the outcome (data not shown). To investigate the protein interactome in more detail, V5-pull-downs of endogenously V5-tagged Tb927.8.1500, Tb927.7.3040 and Tb927.11.4900 were performed and analyzed by mass spectrometry. Purification of V5-tagged BFR1L served again as control. As shown by the volcano plots in figure 3.2 D-F, in each pull-down all three ZC3H5 interaction partners were enriched in the elutions and BFR1L was clearly underrepresented. However, ZC3H5 was not significantly enriched according to this analysis. Comparing the spectral counts for all the proteins that were present in two out of three replicates, ZC3H5 was present in the elution of all three replicates of Tb927.7.3040 and Tb927.11.4900, but not in the elution of Tb927.8.1500 and also not in the elution of the negative control BFR1L (Figure 3.2 G). On one hand, ZC3H5 and its interaction partners had only low spectral counts in the total lysate, which shows the enrichment of the ZC3H5 complex upon pull-down of each of the complex proteins. On the other hand, proteins, which were abundant in the total lysate, were not enriched in the pull-downs, which shows that the purifications were very specific. As already described by the volcano plots, the three proteins interacting with ZC3H5 are enriched in all three pull-downs. These experiments strongly suggest that the three proteins form a complex that interacts with ZC3H5. To identify the stoichiometry of the protein complex subunits the iBAQs of the mass spectrometry data of the tandem affinity purification were analyzed. The iBAQ is the sum of all the protein's total intensity divided by the number of observable peptides of a protein (Schwanhausser et al., 2011). The stoichiometry of the protein complex subunits was determined as 1:2:2 (ZC3H5:Tb927.11.4900:Tb927.7.3040). The stoichiometry calculated for Tb927.8.1500 was 0.76, which is inconclusive. Interestingly, Tb927.11.4900 is suggested to be a guanine nucleotide-binding beta subunit-like protein (G protein). The evidence for Tb927.11.4900 to be a G protein is the conserved G-domains which can be found in the sequence (Suppl. Figure 8). However, there are not in the typical G1-G5 order, but in a circular permutation of G4-G5-G1-G2-G3, which was shown to be the only possible circular permutation that exists in nature (Anand et al., 2006). G-proteins have a GTP-binding domain and hydrolysis of the bound GTP to GDP leads to conformational changes, which can regulate a diversity of functions. G-proteins can be found in all three major kingdoms of life, but mainly in eukaryotes (Wittinghofer and Vetter, 2011). Preliminary results show that the mutation of the GTP-binding pocket of Tb927.11.4900 is lethal (data not shown). The mass spectrometry analysis of the ZC3H5 interaction partners identified four proteins, which could assist in the exchange of GTP/GDP: Homologue of SDO1 (guanine nucleotide exchange factor), TBC-B (likely GTPase activating protein), ADP-ribosylation factor (GTP/GDP exchange protein) and ARL3C (GTP/GDP exchange protein) (Figure 3.2 G).

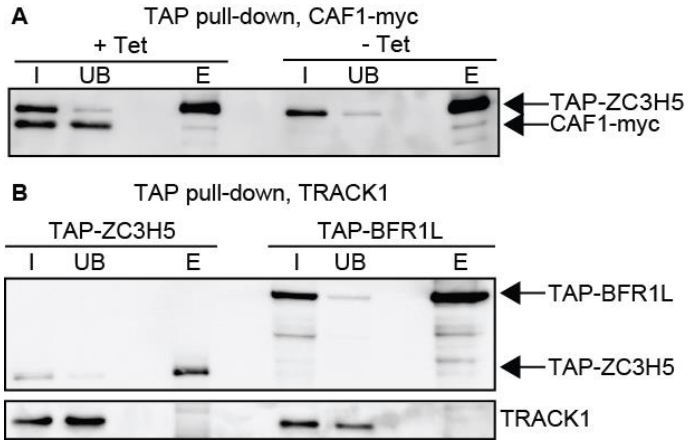


Figure 3.3: ZC3H5 seems not to interact with CAF1 and TRACK1. **A.** CAF1-myc expression was induced (+Tet) or not (-Tet) in a bloodstream form cell line with endogenously TAP-tagged ZC3H5. TAP-ZC3H5 was purified on an anti-PAP affinity matrix. Input (I), Unbound (U) and eluted (E) proteins were analyzed by Western blotting. **B.** Endogenously TAP-tagged ZC3H5 was purified by an anti-PAP affinity matrix. Input (I), Unbound (U) and eluted (E) proteins were analyzed by Western blotting using anti-PAP and anti-TRACK1 antibodies. Endogenously TAP-tagged BFR1L served as control.

Since yeast two-hybrid (Y2H) of the deadenylase CAF1, which is part of the CAF1-NOT complex, gave ZC3H5 as putative interaction partner (Lueong et al., 2016), the interaction of ZC3H5 with CAF1 was investigated. This could provide a first link between the repressive activity of ZC3H5 and a potential mechanism. N-terminally myc-tagged CAF1 was overexpressed in the BF ZC3H5$^{+/TAP}$ cells and affinity purification was performed. Uninduced cells served as control. The proteins in the input, unbound and eluted fractions were then analyzed by Western Blotting (Figure 3.3 A). Pull-down of TAP-ZC3H5 was successful, because it was clearly enriched in the elution. It can be seen that CAF1-myc is expressed upon induction with tetracycline. However, only very little CAF1-myc could be detected in the elution and this band could also be detected in the control, suggesting that ZC3H5 does not interact with CAF1. Finally, the interaction between ZC3H5 and TRACK1 was explored by Co-IP (Figure 3.3 B), since TRACK1 was detected as putative interaction partner by MS analysis by setting a lower threshold. Affinity purification of TAP-tagged ZC3H5 was performed and proteins were analyzed by Western Blotting. TAP-BFR1L served as control. However, TRACK1 could not be identified neither in the elution of TAP-ZC3H5 nor in the elution of TAP-BFR1L.

The previous results suggest that ZC3H5 and the three interaction partners form a complex, which was investigated in more detail. A Y2H assay was performed doing pairwise interactions of ZC3H5 with the validated interaction partners to map intra-complex interactions (Figure 3.4 A). In addition, the deadenylase CAF1 was included in the Y2H assay, for the reasons explained above. Successful expression of the constructs in the different yeast clones was monitored by western blotting (Figure 3.4 B). All proteins were expressed, although at different levels. The Y2H results showed that ZC3H5 as well as CAF1 were auto activators as baits. However, ZC3H5 as prey interacted with pBD-Tb927.11.4900 and pAD-Tb927.11.4900 interacted with pBD-Tb927.7.3040, which suggests that Tb927.11.4900 is the linker between ZC3H5 and Tb927.7.3040. Figure 3.4 C summarizes all protein interactions between ZC3H5 and the three proteins that were identified so far by MS analysis, Co-IPs and Y2H.

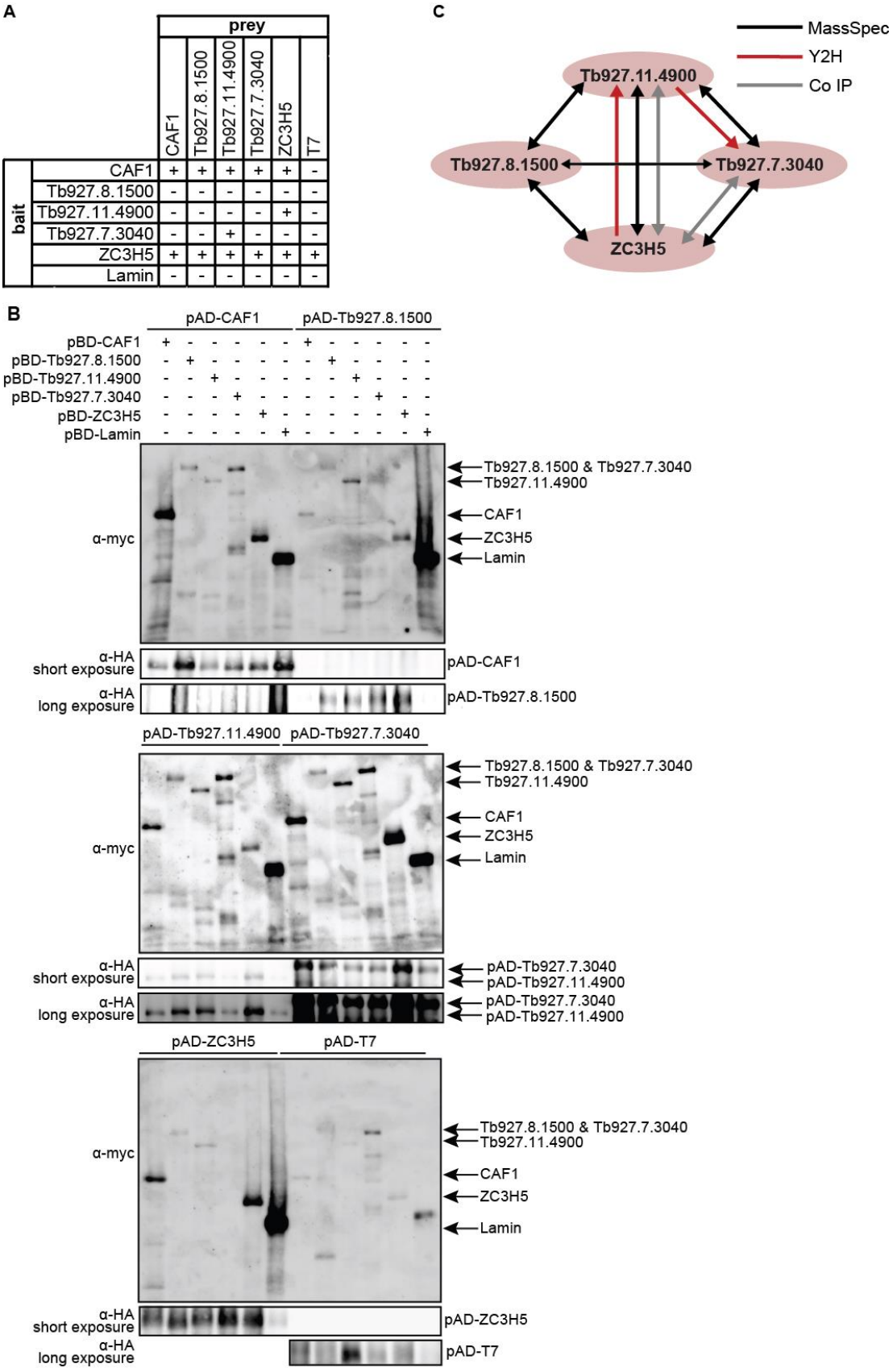


Figure 3.4: ZC3H5 and its interacting proteins form a complex. A. Two-hybrid interactions of ZC3H5 with its putative interaction partners. Interactions marked with + were positive both in quadruple dropout plates and via alpha-galactosidase assay. **B.** Expression of proteins was analyzed by Western Blotting. C. Summary of interactions between ZC3H5, Tb927.7.3040, Tb927.8.2500 and Tb927.11.4900 by Y2H, Co-IPs and mass spectrometry analysis.

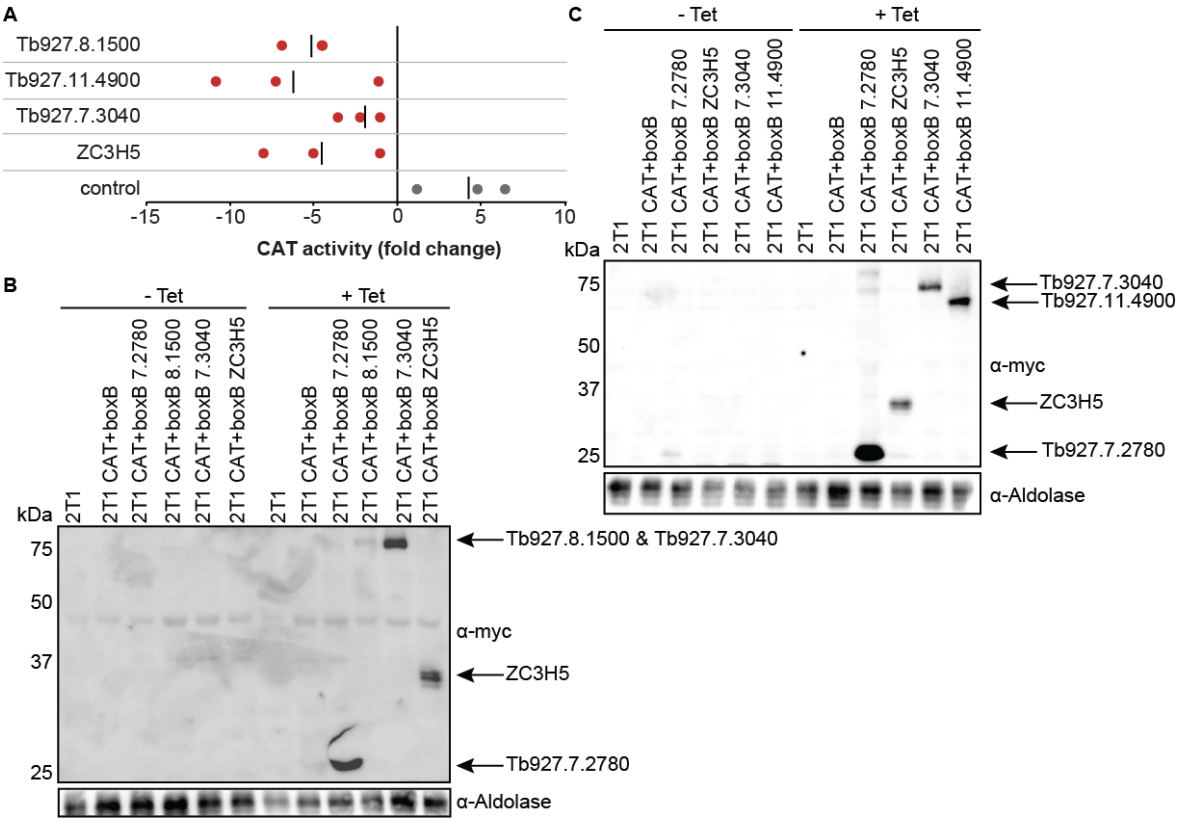


Figure 3.5: ZC3H5 and its interacting proteins decrease reporter gene expression. A. Expression of chloramphenicol acetyltransferase (CAT) was measured in bloodstream form cells expressing different myc-lambda-N-fusion proteins (Tb927.7.2780, ZC3H5, Tb927.7.3040, Tb927.8.1500 or Tb927.11.4900). Tb927.7.2780, which is a known activator of gene expression, served as control. Expression of the lambda-N proteins was induced with tetracycline (+Tet) for 24h. Results show arithmetic mean (black bar) and individual values of 3 independent experiments. **B&C.** Expression of the myc-lambda-N-fusion proteins was validated by Western Blotting.

Since ZC3H5 was identified as a down-regulator in the tethering screen (Erben et al., 2014), I performed CAT assays to validate the screen result and to investigate whether the interacting proteins also act as repressors of gene expression. Tethering of these interaction partners to a CAT reporter showed that ZC3H5 as well as Tb927.8.1500, Tb927.7.3040 and Tb927.11.4900 did act as repressors, whereas the positive control (Tb927.7.2780) acted as an activator (Figure 3.5 A). Expression of the proteins was confirmed by Western Blotting (Figure 3.5 B&C).

I conclude from these results that ZC3H5 interacts with Tb927.8.1500, Tb927.7.3040 and Tb927.11.4900 and the three proteins also interact with each other. iBAQ analysis suggests a stoichiometry of 1:2:2 (ZC3H5:Tb927.11.4900:Tb927.7.3040) and according to the Y2H data Tb927.11.4900 is the linker between ZC3H5 and Tb927.7.3040. In addition, all three proteins as well as ZC3H5 are repressors of gene expression according to CAT assays. These results suggest that ZC3H5 and its interacting proteins form a repressive complex.

3.1.3. Knock-down of proteins interacting with ZC3H5 results in reduced growth

Since knock-down of ZC3H5 led to an increase of cells in G2/M phase, I was wondering whether knock-down of the three interacting proteins would result in similar cell cycle phenotypes. Cell growth after RNAi-mediated knock-down was analyzed and expression of proteins was monitored by Western blotting. Down-regulation of all three proteins led to a

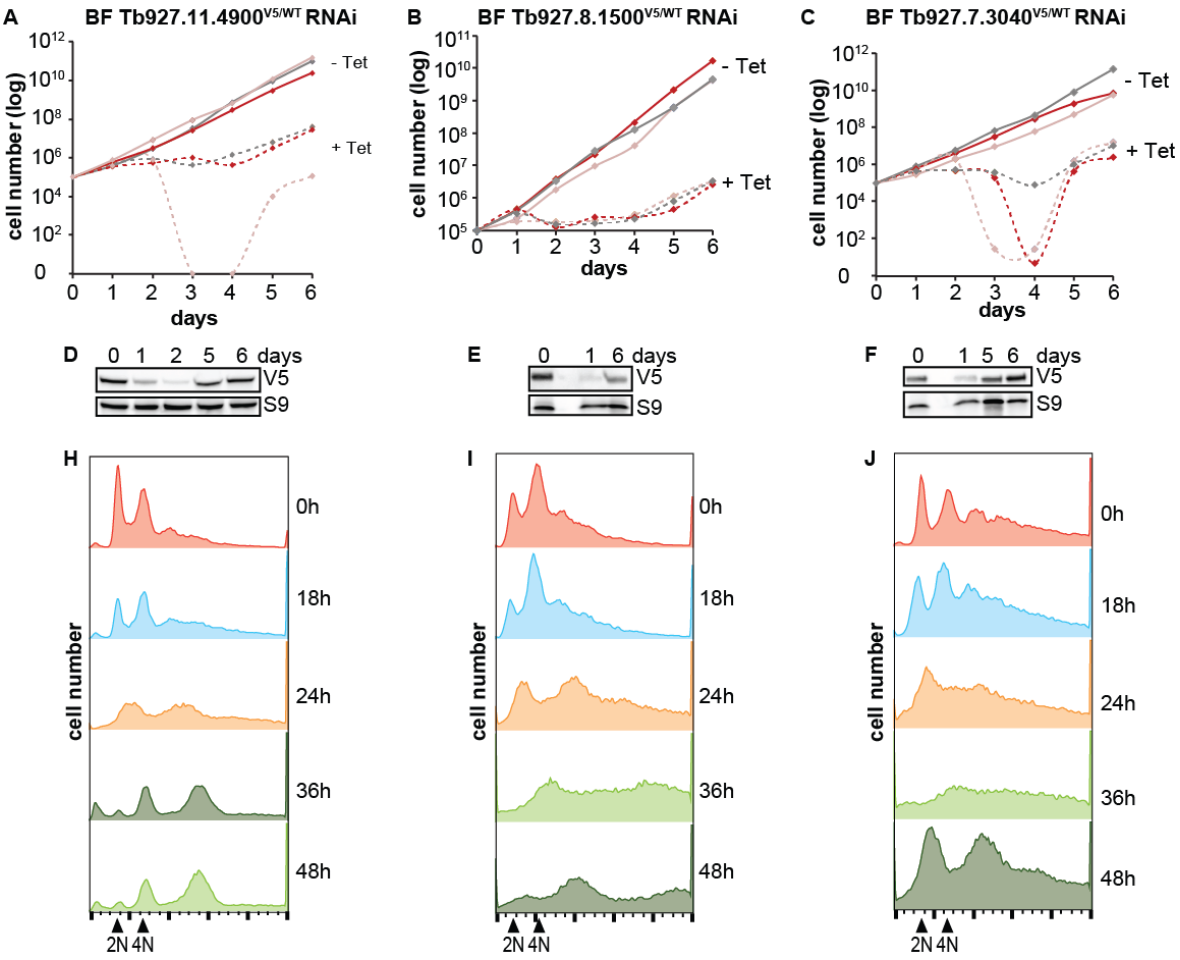


Figure 6: Knockdown of the interaction partners of ZC3H5 in bloodstream form cells. A-C. Growth curve of three independent experiments for RNAi targeting Tb927.11.4900, Tb927.8.1500 or Tb927.7.3040 in BF cells with endogenously V5-tagged protein. Error bars indicate standard deviation. Tet + (dark red) and Tet - (light red). **D-F.** $3x10^6$ cells were collected and Western Blot of V5-tagged protein was performed. S9 was used as loading control. **H-J.** Knock-down of Tb927.11.4900, Tb927.8.1500 or Tb927.7.3040 was induced for various times, DNA was stained with Propidium iodide and analyzed by FACS.

cell growth arrest after 2 days of RNAi induction (Figure 3.6 A-C). In contrast to ZC3H5 knock-down, the cells did not die and recovered form reduced growth after 4 days. However, the Western Blots showed that the cells escaped from the knock-down after some days, because also protein levels increased again (Figure 3.6 D-F). To examine cell cycle behavior of the three proteins in more detail, I did FACS analysis upon knock-down of the proteins. Knock-down was induced for various time points, DNA was stained with Propidium iodide and analyzed by FACS (Figure 3.6 H-J). Upon knock-down of the proteins for 18h an increase of the peak of the G2/M-phase (4N) could be seen. This showed that the number of cells with 2 nuclei and 2 kinetoplasts (2N2K) increased. 24h after induction of knock-down a decrease of both peaks could be observed, because the cells were already dying. Upon knock-down of Tb927.8.1500 (Figure 3.6 I), the peak of the G2/M-phase (4N) was already higher than the peak of the G1-phase (2N) in the uninduced cells. This showed that the number of cells with 2 nuclei and 2 kinetoplasts (2N2K) was already increased without induction of RNAi. However, this was not surprising, because the cells already grew slower without induction of RNAi. The calculated division time was 8.9±0.6 h, whereas it was 7.3±0.5 h (BF Tb927.11.4900$^{V5/WT}$ RNAi) and 7.6±0.6 h (BF Tb927.7.3040$^{V5/WT}$ RNAi) for the

other two cell lines without induction of RNAi. This suggests that the Tb927.8.1500 RNAi plasmid is leaky and the protein is already knocked-down without induction.

Taken together, RNAi of the three interacting proteins leads to a growth defect from which the cells recover after 4 days, when the RNAi is not efficient anymore and the protein level increases. In addition, the increase of cells with 2N2K, as it was observed for ZC3H5 knock-down, could be observed for all three interacting proteins with knock-down of Tb927.8.1500 showing the strongest effect.

3.1.4. RIP-Seq identifies mRNAs encoding cytoskeleton proteins as ZC3H5 targets

To gain further information about the mRNA targets of ZC3H5, RNA affinity purification followed by next-generation sequencing (RIP-Seq) was performed using the described endogenously TAP-tagged ZC3H5 cell line. The RIP-Seq identified 918 putative mRNA targets with a log2 fold change between bound/unbound >1 (Figure 3.7 A, EE). Further analysis revealed that longer transcripts were slightly more enriched than shorter (p=0.38) (Figure 3.7 B, EE) with a negative correlation between cellular mRNA abundance and enrichment after RIP-Seq (Pearson correlation -0.31) (Figure 3.7 C, EE). Principle component analysis showed that the elution and unbound samples clustered and one component that separated the two sample types accounted for nearly all variance (Figure 3.7 D, KL). Procyclin-associated gene mRNA (*PAG1*), was the most highly enriched transcript (29-fold) and the mostly enriched class of biological processes was ‘cytoskeleton’ (79 transcripts) (Figure 3.7 E, EE). On the other hand, transcripts of the category ‘ribosomal proteins’ were strikingly underrepresented. The targets were also compared with the cell cycle dependent transcriptome (Archer et al., 2011) and 261 putative targets peak in the S-phase of the cell cycle. The RIP-Seq data were validated by RT-qPCR. TAP-ZC3H5 and its bound mRNA targets were affinity purified and the RNA from the unbound and eluted samples was purified. RNA was reverse transcribed into cDNA and qPCR was performed to amplify the cDNA of the target mRNAs. Five targets that appeared to bind to TAP-ZC3H5 were used: Tb927.10.10350 (putative Protein Kinase), Tb927.10.7880 (putative Sperm tail C-terminal domain containing protein), Tb927.11.10900 (component of motile flagella 9), Tb927.9.15050 (nexin-dynein regulatory complex 4) and Tb927.11.5810 (putative surfeit locus protein 6). Tb927.4.1860 (putative ribosomal protein S19), which was underrepresented in the elution of the RIP-Seq data, was used as negative control. Ct values were normalized to Tb927.4.1860 and expression fold change was calculated. All putative mRNA targets were enriched in the elution in 3 independent experiments (Figure 3.7 F). I also normalized the Ct values to Tubulin, which showed that Tb927.10.10350 was enriched and the other targets were slightly enriched, whereas the negative control Tb927.4.1860 was not enriched (Figure 3.7 G). DREME analysis (Bailey, 2011) for the identification of a common binding motif in the 5’UTR, CDS or 3’UTR of the targets identified a set of putative binding motifs. Set of mRNAs, which were not enriched in the RIP-Seq data, were used as control set of mRNAs. The motifs with the lowest E-value were showing a TAG motif surround by a small variety of nucleotides. The motif identified in the 5’UTR of the targets had the highest E-value of 1.1E-8 (Figure 3.7 H), the motif identified in the CDS of the targets had the lowest E-value of 2.8E-78 (Figure 3.7 I) and the motif identified in the 3’UTR of the targets had an E-value of 3.3E-19 (Figure 3.7 J). Taken together, RIP-Seq identified cytoskeleton enriched ZC3H5 targets and some targets could be validated by RT-qPCR.

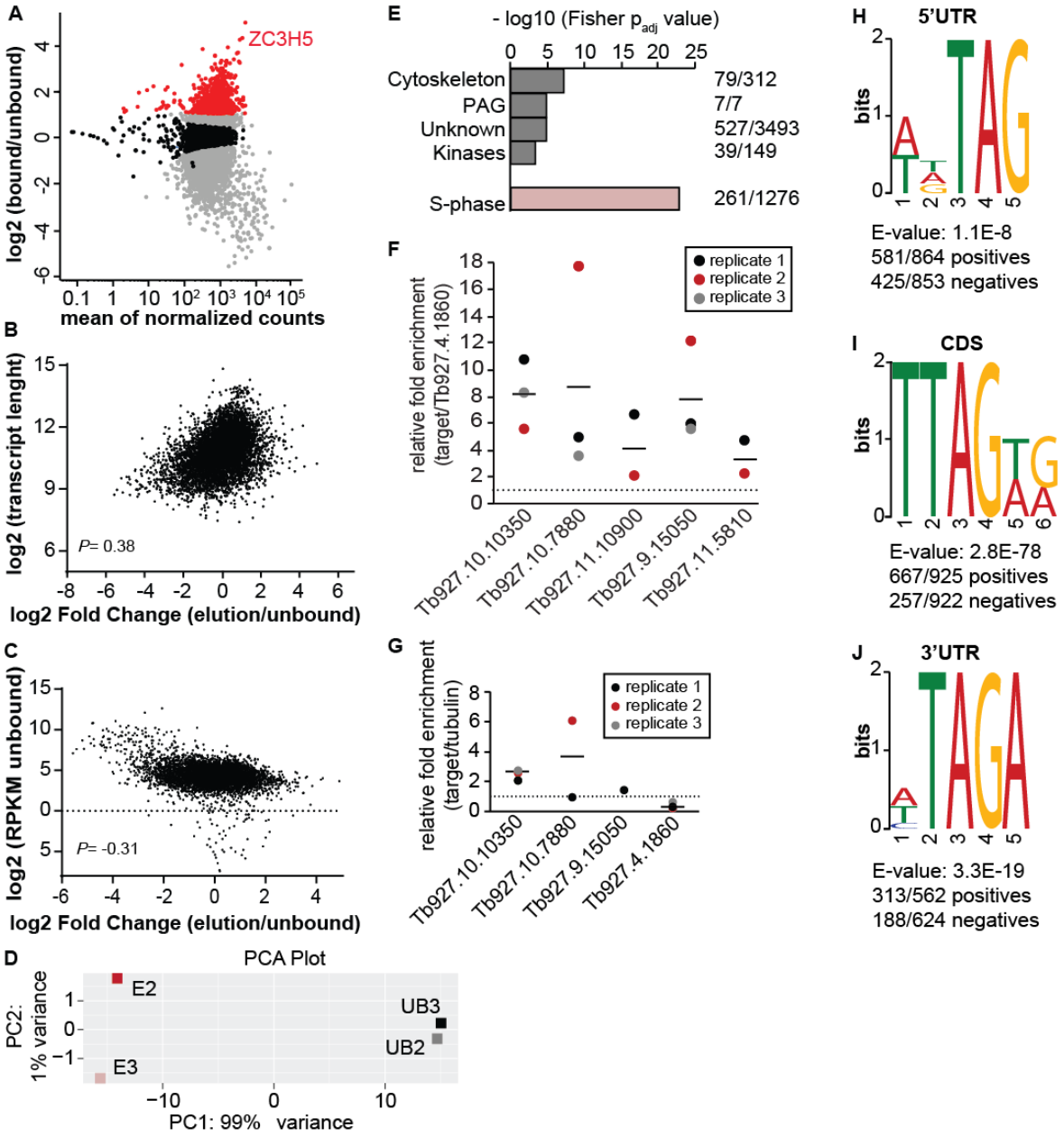


Figure 3.7: RIP-Seq identified 918 putative mRNA targets. A: To identify RNAs bound by ZC3H5, Tandem Affinity Purification in a cell line expressing *in situ* tagged TAP-ZC3H5 was performed. The bound RNA was purified and sequenced, and the results were analyzed by DEseq2. Bound RNA was compared with total RNA. We could identify 981 bound transcripts, which were at least 2-fold enriched compared to total RNA (Padj >0.05) (EE). **B.** Correlation analysis of enriched mRNAs and transcript length. **C.** Comparison of enriched mRNAs and RPKM (reads assigned per kilobase of target per million mapped reads). **D.** Principle component analysis (PCA) of the eluted (E) and unbound (UB) in duplicates (KL). **E.** Functional categories (grey) enriched in RIP-Seq data and cell cycle phase, in which targets peak (light red) (EE). **F & G.** Validation of the RIP-Seq data by qPCR. Unbound and eluted RNA was analyzed by RT-qPCR. Expression fold change was calculated and normalized to Tb927.4.1860 or tubulin. Results show arithmetic mean (black bar) and individual values of 3 independent experiments. **H-J.** Identification of possible binding motifs in the ZC3H5 mRNA targets. Depicted are the sequence motif logos of possible ZC3H5 binding sites in the 5'UTR, CDS or 3'UTR. The 918 significantly overrepresented and underrepresented mRNAs in the elution fraction of the RIP-Seq experiments were used and all sequences were removed which were shorter than 8 nucleotides. Motif identification was done using the DREME algorithm (Bailey, 2011). Positives / negatives give information about how many sequences tested contained the motif. As control mRNA set, mRNAs that were not enriched in the RIP-Seq analysis and having on average the same size, were used. Analysis was restricted to the sense strand.

3.1.5. Short-term down-regulation of ZC3H5 results in a minor effect on the transcriptome

To identify the mRNAs that are affected by ZC3H5 knock-down, transcriptome analysis after 9h and 12h of ZC3H5 knock-down was performed. ZC3H5 knock-down was induced in two independent clones (Figure 3.8 A, EE) and the reduction of ZC3H5 protein level upon knock-down was monitored by Western Blotting (Figure 3.8 B, EE). After 9h and 12h of ZC3H5 knock-down, RNA was extracted and RNA sequencing was performed. Principle component analysis showed that the uninduced (WT) and induced samples (9h and 12h) clustered within their group very close. PC1 (83%) and PC2 (10%) accounted for nearly all variance within the samples (Figure 3.8 C, KL). However, this showed only minor effects. In fact, after 12 h induction only 44 transcripts were up-regulated by >2-fold and only 31 transcripts were down-regulated by >2-fold (Figure 3.8 D, EE). Reassuringly, the ZC3H5 mRNA was decreased, which indicated a successful knock-down. Further analysis revealed that longer transcripts were slightly more enriched than shorter transcripts in the mRNAs identified upon knock-down of ZC3H5 for 9h and 12h (Pearson correlation 0.22) (Figure 3.8 E, EE) with a positive correlation between cellular mRNA abundance and increasing after RNAi (Pearson correlation 0.13) (Figure 3.8 F, EE). To analyze the group of genes that were up-regulated upon knock-down of ZC3H5 a threshold of at least 1.5-fold enrichment was set, since 2-fold enrichment only gave 44 candidates. The functional group that was up-regulated most was ‘ribosomal proteins’ (Figure 3.8 G, EE). Since the tethering screen suggested ZC3H5 to be a post-transcriptional repressor, it was expected that target mRNAs of ZC3H5 would be increased upon ZC3H5 knock-down. However, comparison of the transcriptomic data with the RIP-Seq data did not show any correlation or overlap (Figure 3.8 I). Transcripts showing the greatest increase in abundance after knock-down of ZC3H5 were neither selectively enriched nor underrepresented in the RIP-Seq. To validate the transcriptomic data, I induced knock-down of ZC3H5 for 12h and 18h, harvested the cells and isolated the RNA from the different time points. The RNA was then analyzed by Northern blotting using probes specific for the putative up-regulated mRNAs (Figure 3.8 H). Upon knock-down of ZC3H5, I observed decrease of the ZC3H5 mRNA and an increase on the mRNA levels encoding the proteins Tb927.7.3040 (F-box and WD40 domain containing protein), Tb927.11.10340 (putative Serine/threonine-protein phosphatase PGAM5) and Tb927.2.4550 (2’-O-ribose RNA methyltransferase SPB1 homologue), which were up-regulated in the transcriptomic data. mRNA levels of Tb927.3.1660 (SUMO-interacting motif-containing protein), which was not up-regulated in the transcriptomic data, remained unchanged. Taken together, the abundance of only a few mRNAs changed upon knock-down of ZC3H5.

3.1.6. Effect of ZC3H5 down-regulation on ribosomal biogenesis, transcription and mRNA processing

The mRNAs that go up upon ZC3H5 knock-down, are characterized by a long half-life (e.g. mRNAs encoding for ribosomal proteins) (Figure 3.9 A), which led to the question of whether the transcription and mRNA processing machinery (splicing, transport and turnover) are affected. To analyze that, ZC3H5 was knocked-down in bloodstream form trypanosomes for 18 h, RNA was isolated and the spliced-leader RNA was analyzed by Northern blotting. However, the amount of spliced leader-containing RNA did not change after 18 h induction of ZC3H5 knock-down (Figure 3.9 B). This showed that even after knock-down of ZC3H5

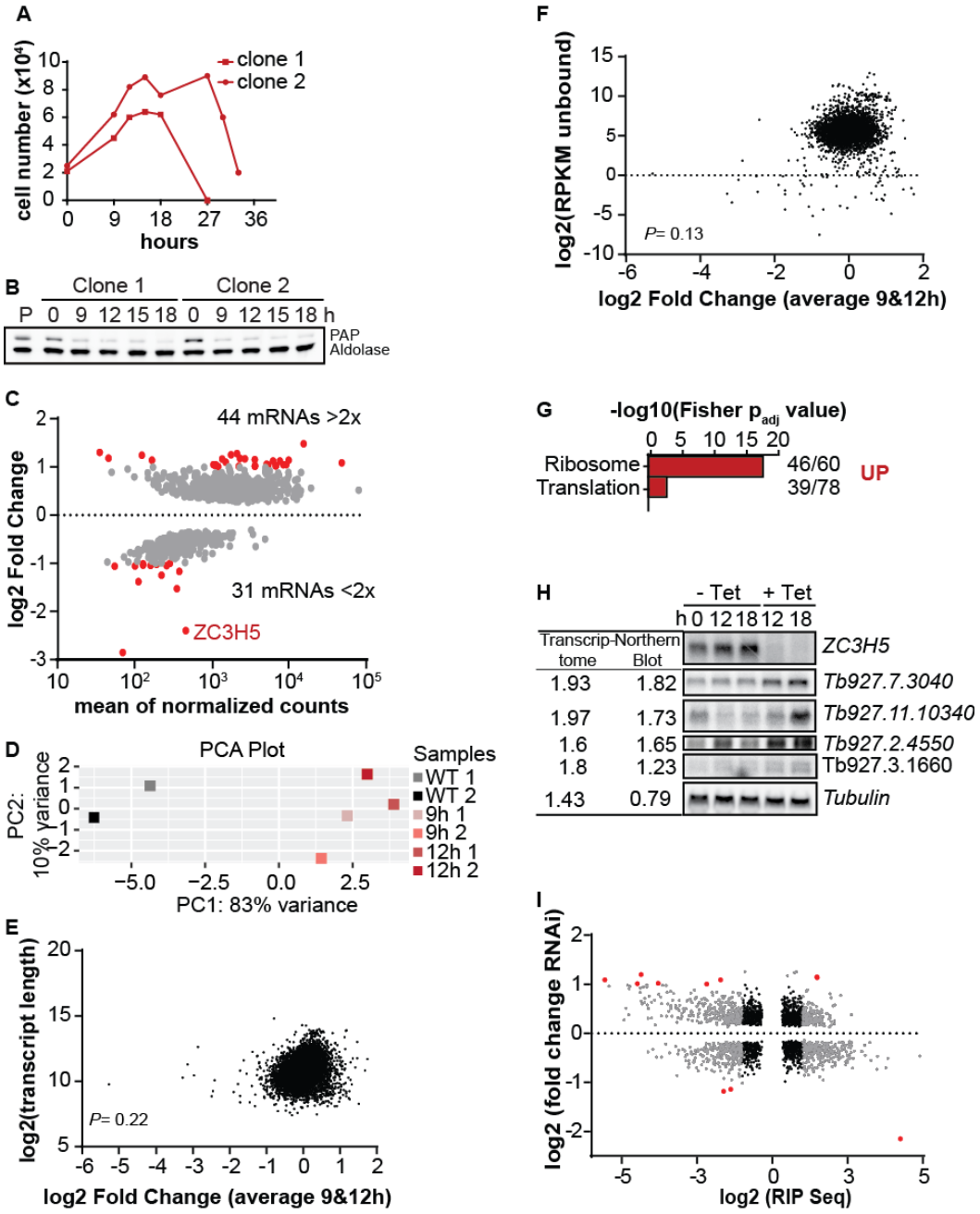


Figure 3.8: Short term down-regulation of ZC3H5 only slightly affects the transcriptome. **A.** Growth curves showing cell counts over time following knock-down of ZC3H5 RNAi by tetracycline induction for two independent clones (EE). **B.** ZC3H5 RNAi cell line was analyzed by Western blotting with an anti-PAP antibody over time following tetracycline induction. Anti-aldolase antibody was used as a loading control (EE). **C.** Downregulation of ZC3H5 was induced for 9h and 12h and total RNA was analyzed by DEseq2. The figure shows an MA plot. As down-regulated proteins only these were considered, which go down both at 9h and 12h, and being down-regulated more after 12h than 9h (12h<9h<WT), and vice versa for the up-regulated proteins (EE). **D.** Principle component analysis (PCA) of the WT and ZC3H5 knock-downs in duplicates (EE). **E.** Correlation analysis of fold change average of all mRNAs identified at 9h&12h and transcript length. Pearson correlation: P=0.22 (EE). **F.** Correlation analysis of fold change average up-regulated at 9h&12h and RPKM (reads assigned per kilobase of target per million mapped reads). Pearson correlation: P=0.13 (EE). **G.** Functional categories up-regulated upon knock-down of ZC3H5 (EE). **H.** Downregulation of ZC3H5 was induced for 12h and 18h and total RNA was analyzed by Northern blotting. Blots were probed with ZC3H5, Tubulin and four mRNAs that were increased in the transcriptomic data upon ZC3H5 RNAi. **I.** Comparison of the fold change of the RIP seq data with the fold change of the RNAs of the transcriptomic data upon average knock-down of ZC3H5 for 9h & 12h (EE). Red dots are the candidates that are affected by 2-fold in both experiments. All transcripts quantified in both experiments with an adjusted P-value of less than 0.05 are shown.

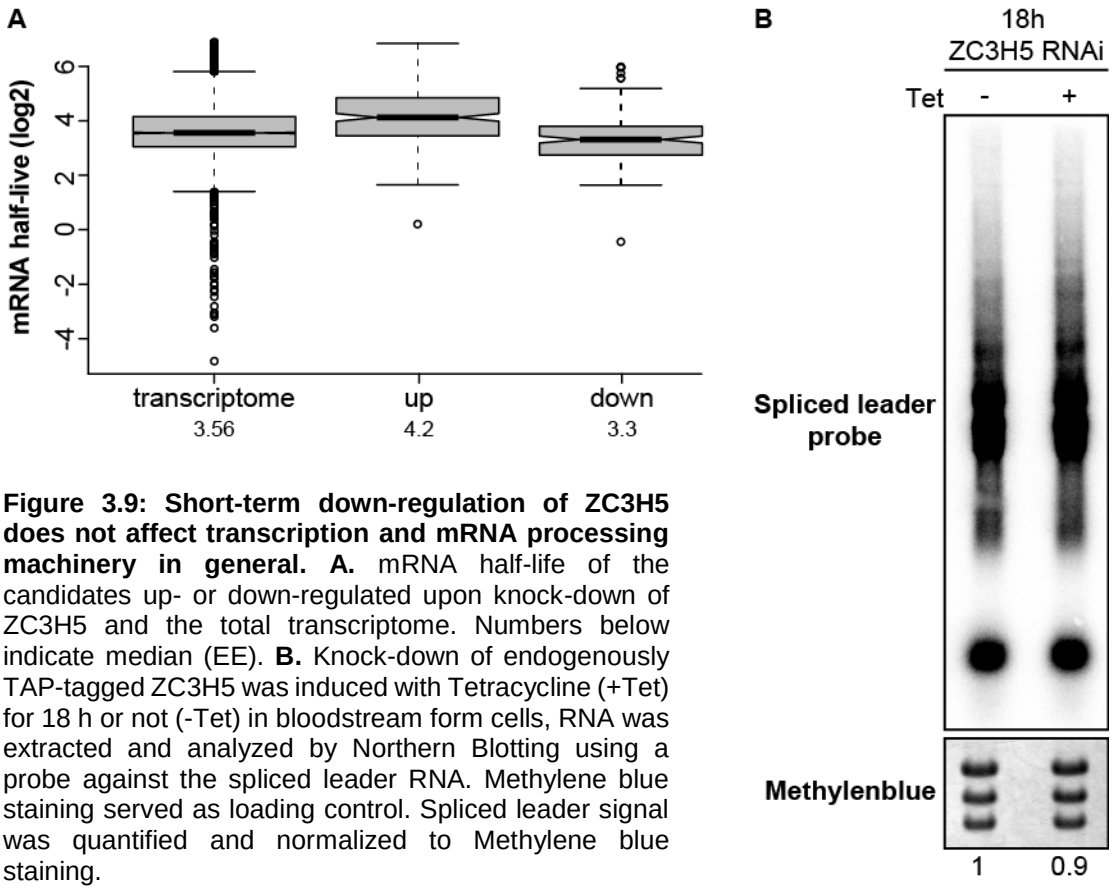


Figure 3.9: Short-term down-regulation of ZC3H5 does not affect transcription and mRNA processing machinery in general. A. mRNA half-life of the candidates up- or down-regulated upon knock-down of ZC3H5 and the total transcriptome. Numbers below indicate median (EE). **B.** Knock-down of endogenously TAP-tagged ZC3H5 was induced with Tetracycline (+Tet) for 18 h or not (-Tet) in bloodstream form cells, RNA was extracted and analyzed by Northern Blotting using a probe against the spliced leader RNA. Methylene blue staining served as loading control. Spliced leader signal was quantified and normalized to Methylene blue staining.

for 18h the transcription was still functional and thereby also the mRNA processing machinery, because a deregulation of this machinery would affect the mRNA levels, which should be seen in the SL Northern Blot.

The transcriptomic data showed an up-regulation of the functional groups 'ribosomal proteins' and 'translation', which suggests that ribosome biogenesis or translation could be affected by ZC3H5 knock-down. To analyze if ribosome biogenesis was affected, ZC3H5 RNAi was induced for 24 h, RNA was isolated and the levels of pre-5.8S rRNA and pre-18S rRNA were analyzed by Northern blotting using probes to detect the rRNA precursors as described by Sakyiama *et al.* (Sakyiama et al., 2013) (Figure 3.10 A). Successful knock-down of ZC3H5 was monitored by Western Blotting (Figure 3.10 B). A decrease of the 0.17 kb pre-5.8S rRNA precursor could be seen in one out of three replicates upon knock-down of ZC3H5 (Figure 3.10 C&D). The amount of the 5.9 kb pre-5.8S rRNA precursor and 2.7 kb and 3.7 kb pre-18S rRNA precursors did not change upon knock-down of ZC3H5 (Figure 3.10 E&F), whereas I could see an increase of the 9.6 kb pre-18S rRNA precursor in one out of three replicates. To analyze, whether translation was affected, a ^{35}S-Methionine incorporation assay upon ZC3H5 knock-down was performed. BF cells served as control and cycloheximide (CHX) treated cells, which inhibits protein synthesis, served as negative control. It can be seen that CHX inhibited protein synthesis efficiently and almost no ^{35}S-Methionine was incorporated in BF as well as uninduced BF ZC3H5$^{TAP/WT}$ RNAi cells. Only a slight reduction of ^{35}S-Methionine incorporation after 9h and 18h of *ZC3H5* RNAi induction could be observed (Figure 3.11).

Taken together, neither transcription and the mRNA processing machinery nor translation are heavily affected upon knock-down of ZC3H5. There was also no reproducible effect on rRNA processing.

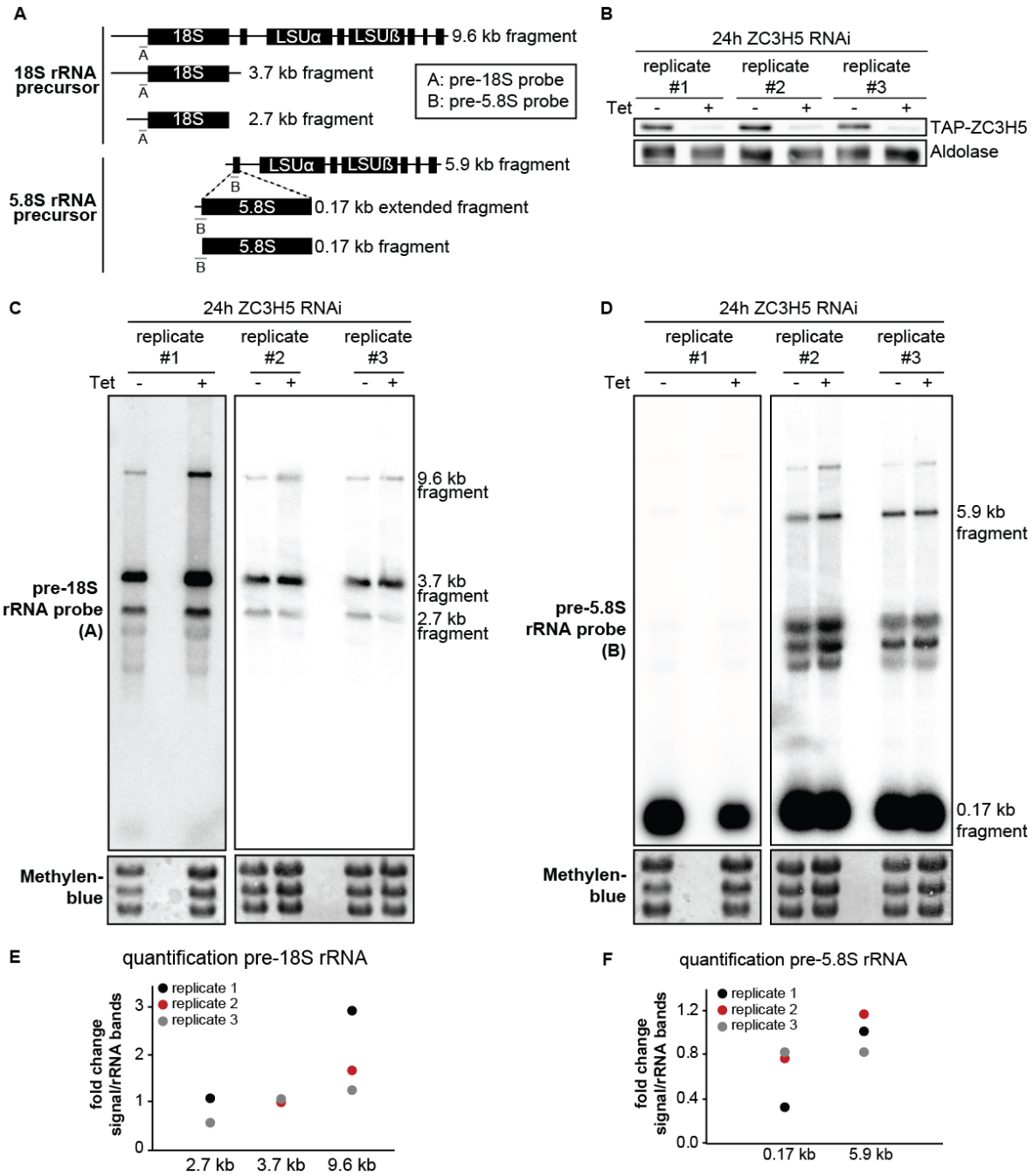


Figure 3.10: Knock-down of ZC3H5 does not affect pre-rRNA processing. A. Schematic representation of the *T. brucei* rRNA transcription unit, probes used for Northern blotting (gray bars) labeled with A (18S rRNA precursor) or B (5.8S rRNA precursor). Processing products observed by Northern blotting with corresponding length. Figure modified from: (Sakyiama et al., 2013). **B.** ZC3H5 RNAi cell line was analyzed by Western blotting with an anti-PAP antibody following tetracycline induction for 24h. Anti-aldolase antibody was used as a loading control. **C & D.** Knock-down of ZC3H5 was induced with Tetracycline (+Tet) for 24 h or not (-Tet) in bloodstream form cells, RNA was extracted and analyzed by Northern Blotting using a probe against the pre-18S-rRNA or pre-5.8S-rRNA in triplicates. Methylene blue staining served as loading control. **E & F.** Indicated bands were quantified and normalized to Methylene blue staining in all three replicates.

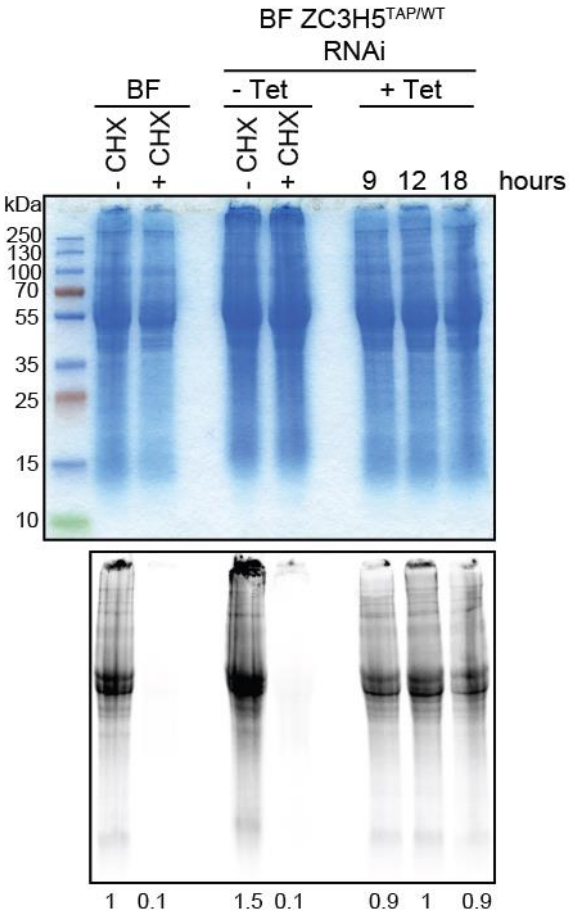


Figure 3.11: Knock-down of ZC3H5 leads to a slight decrease in ^{35}S-Methionine incorporation. Knock-down of endogenously TAP-tagged ZC3H5 was induced with Tetracycline (+Tet) for 9h, 12h and 18h or not (-Tet) in bloodstream form cells. BF cells served as control. Treatment with Cycloheximide (CHX) served as negative control. Cells were washed in 1xPBS + 0.5% glucose and incubated with radioactively labelled ^{35}S-Methionine. Proteins were separated by SDS-PAGE and incorporation of ^{35}S-Methionine incorporation was analyzed using phosphor imager plates. Coomassie staining served as loading control. ^{35}S-Methionine incorporation was quantified and normalized to Coomassie signal (see numbers below the blot).

3.1.7. Short-term down-regulation of ZC3H5 leads to an increase in monosomes

To investigate the effect of ZC3H5 knock-down on its target mRNAs in more detail, polysome fractionation upon ZC3H5 knock-down was performed. Knock-down of ZC3H5 resulted in a slight increase of the monosomal peak after 12h induction (Figure 3.12). This peak increased heavily after 18h induction. However, the calculated monosome/polysome ratio after 18h induction is the minimum value, because the monosome peak after 18h of knock-down was so high that it could not be detected completely due to technical limitations. In addition, a slight decrease of the heavy polysomal fractions could be observed and an increase of the 60S peak. I could also observe a peak between the 2mer (2 ribosomes attached to the mRNA) and 3mer (3 ribosomes attached to the mRNA) polysome peaks upon knock-down of ZC3H5, which suggested a 2 ½mer, that consists of a mRNA with 2 bound ribosomes and one 40S subunit.

Since an increase of the monosomal peak could be seen upon knock-down of ZC3H5, it was investigated in which fractions of the polysome fractionation ZC3H5 and its protein interaction partners can be found. Migration of endogenously TAP-tagged ZC3H5 and endogenously V5-tagged Tb927.8.1500, Tb927.7.3040 and Tb927.11.4900 in the different fractions of the polysome fractionation was analyzed by Western Blotting (Figure 3.13). TR, which is known to not go to the polysomal fractions, and S9, a ribosomal protein known to go to the polysomal fraction, were used as controls. It can be seen that TR was mainly in the free and monosomal fractions, whereas S9 accumulated in the polysomal fractions. All four proteins that were investigated, can be found mainly in the free fractions, but a small amount of protein could even be detected in the polysomal fractions. However, the cells

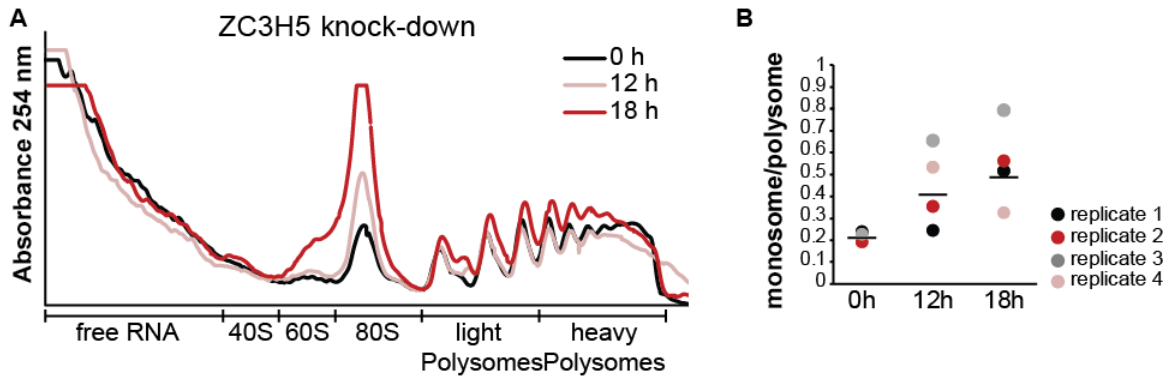


Figure 3.12: Loss of ZC3H5 results in an increase in monosomes and a loss of heavy polysomes. A. Extracts from bloodstream form cells grown with (12h or 18h) or without (0h) Tetracycline to induce ZC3H5 RNAi, were separated on sucrose gradients. The absorbance profiles of three typical gradients are shown. Arrow indicates half-mers. **B.** Average of monosome/polysome ratio calculated from the polysome profiles. Error bars indicate the standard deviations around the means of three biological replicates.

expressing V5-tagged Tb927.8.1500 grew slowly and therefore less cells were used for the polysome profiling, which explains the lower absorbance peaks and less detected protein.

The increase of the monosomal peak upon ZC3H5 knock-down suggests that translation is reduced in these cells. Since knock-down of ZC3H5 did not alter mRNA abundance and translation is not globally affected, polysome profiling was performed to analyze whether the effect is target specific. RNA sequencing of free&40S, monosome&60S, light polysomes and heavy polysomes of BF cells and ZC3H5 knock-down after 0h and 12h induction, as described in figure 3.14, was performed. To decide which fractions to pool, RNA of each fraction was analyzed by Northern Blotting using probes against the spliced leader sequence (to detect mRNAs). The strong spot below 200nt is the spliced leader precursor RNA (*SLRNA*) and the smear above is *trans* spliced mRNA. There was no obvious difference of the spliced-leader-containing mRNA between the control cells and the ZC3H5 knock-down. The Northern blot using a probe against spiked in globin served as loading control and the methylene blue staining visualized the rRNA. The fractions were pooled individually for each polysome fractionation dependent on the absorbance profile and the northern blot as indicated. The fractions of the pooled mRNA were analyzed again by Northern Blotting, which could be used for normalization of the RNA-seq results (Figure 3.15 A&B). A shift of the ZC3H5 mRNA to the free fractions could be observed upon knock-down of ZC3H5. However, the only fraction in which I could see a difference between the RNAs in the uninduced and induced samples was the free fraction (Figure 3.16 A-D). An increase of mRNAs ‘encoding ribosomal proteins’ and ‘PAG’ and a decrease of the functional groups ‘GRESAG’, ‘Cytoskeleton’, ‘Mitochondrial DNA’, ‘RNA binding’ and ‘citric acid cycle’ was observed (Figure 3.16 E). However, due to the pooling of several fractions a shift of mRNAs from, for example, 8 to 5 ribosomes would not have been detected.

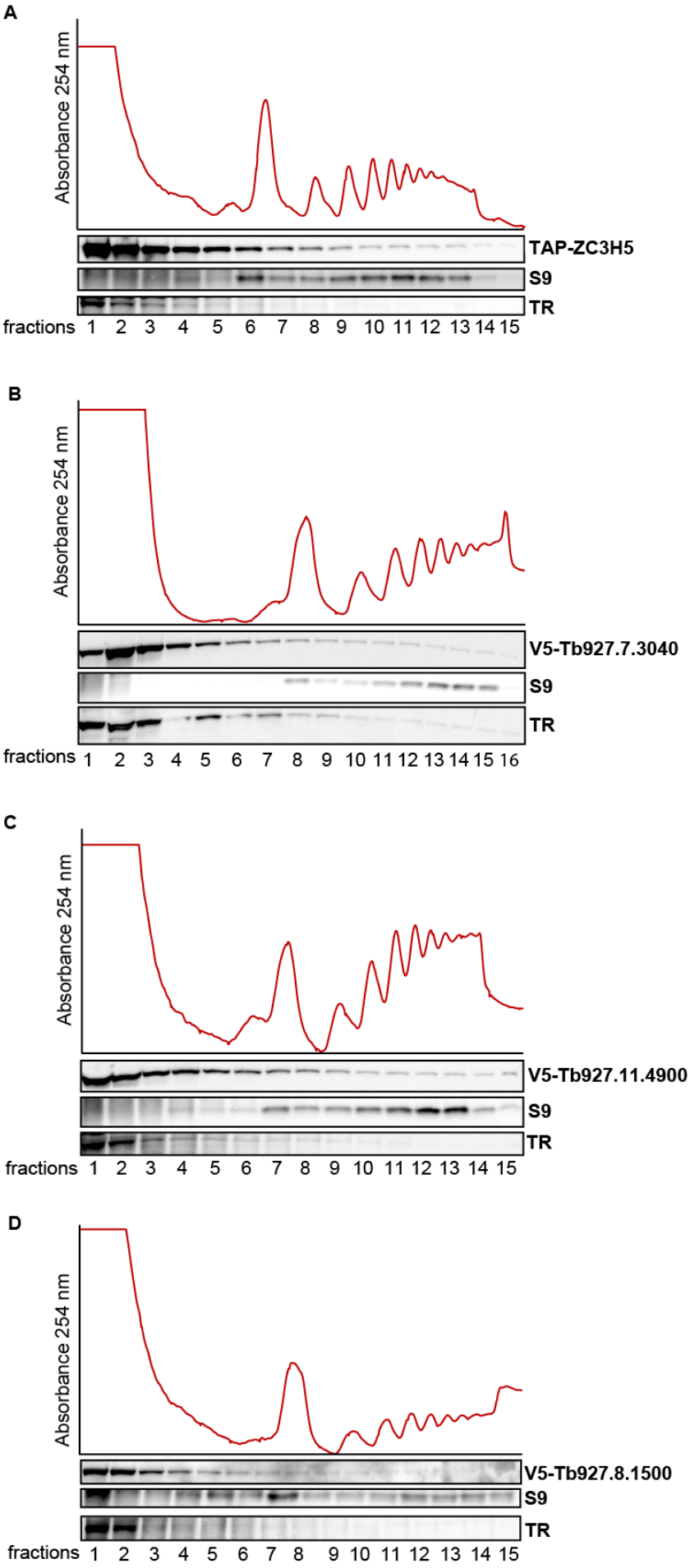


Figure 3.13: ZC3H5 and its interaction partners do mainly localize in early fractions of polysome profiling. Extracts from bloodstream form cells endogenously expressing TAP-tagged ZC3H5 (**A**), endogenously V5-tagged Tb927.7.3040 (**B**), Tb927.11.4900 (**C**) or Tb927.8.1500 (**D**) were separated on sucrose gradients. Proteins of each fraction were analyzed by Western blotting using antibodies against the protein-tag (TAP or V5), TR and S9.

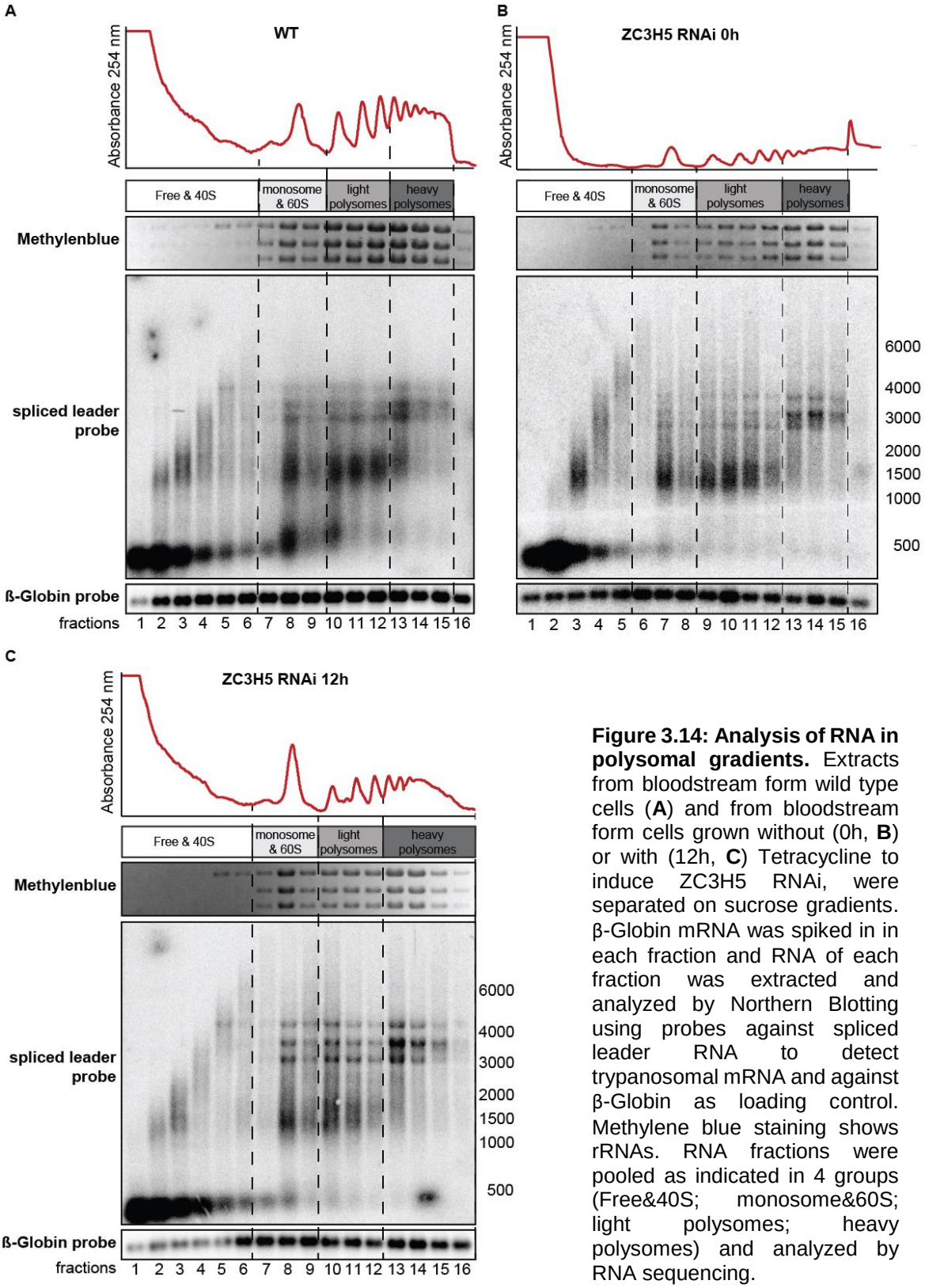


Figure 3.14: Analysis of RNA in polysomal gradients. Extracts from bloodstream form wild type cells (**A**) and from bloodstream form cells grown without (0h, **B**) or with (12h, **C**) Tetracycline to induce ZC3H5 RNAi, were separated on sucrose gradients. β-Globin mRNA was spiked in in each fraction and RNA of each fraction was extracted and analyzed by Northern Blotting using probes against spliced leader RNA to detect trypanosomal mRNA and against β-Globin as loading control. Methylene blue staining shows rRNAs. RNA fractions were pooled as indicated in 4 groups (Free&40S; monosome&60S; light polysomes; heavy polysomes) and analyzed by RNA sequencing.

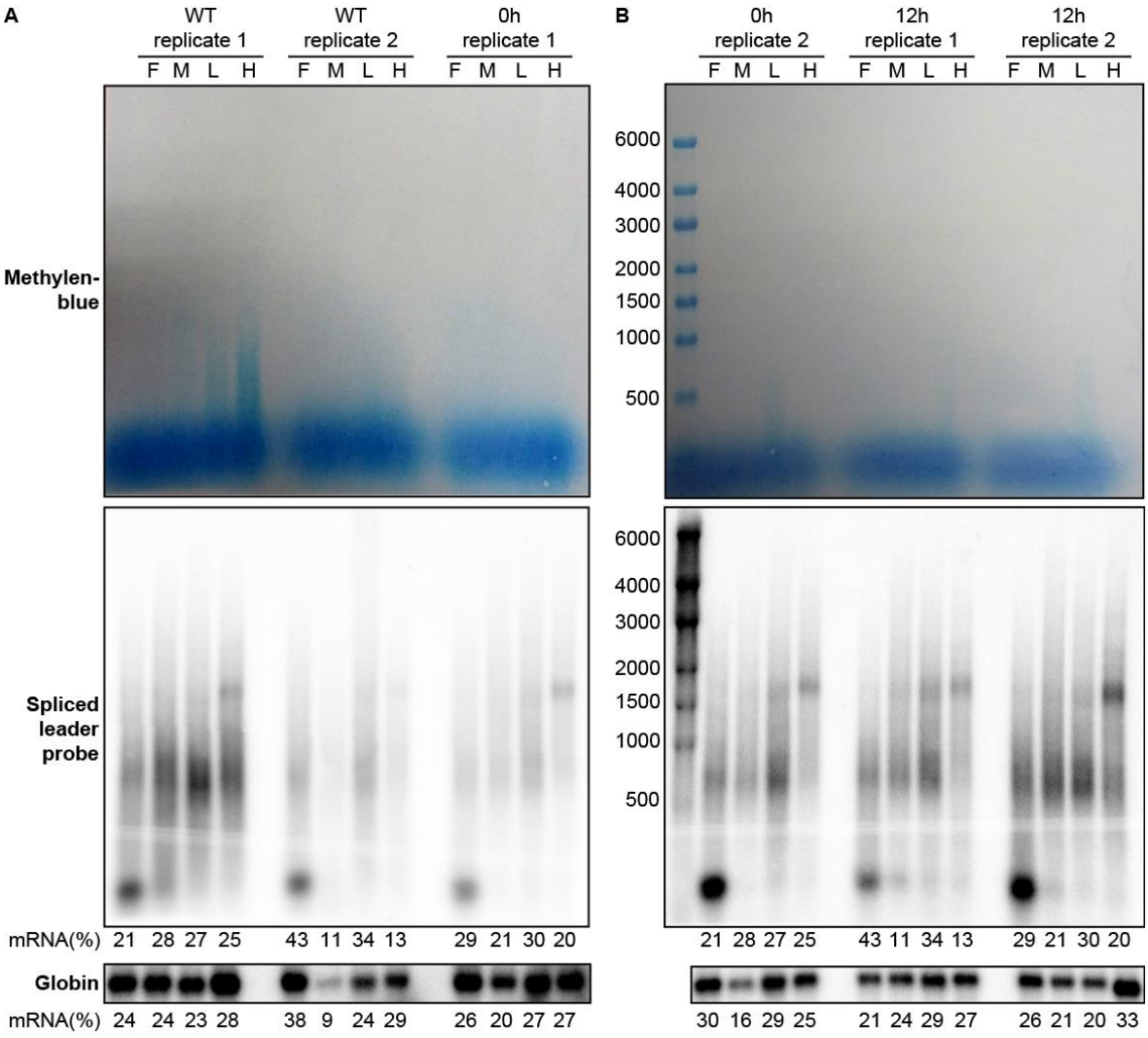


Figure 3.15: Analysis of pooled fractions for RNA Seq. A&B. Northern blots of the combined fractions described in figure 13 using probes against spliced leader RNA to detect trypanosomal mRNA and against β-Globin as loading control. Methylene blue staining shows rRNA depletion. F: Free&40S; M: monosome&60S; L: light polysomes; H: heavy polysomes. The Northern blots were quantified and the percentage of total signal in each sucrose gradient fraction is shown.

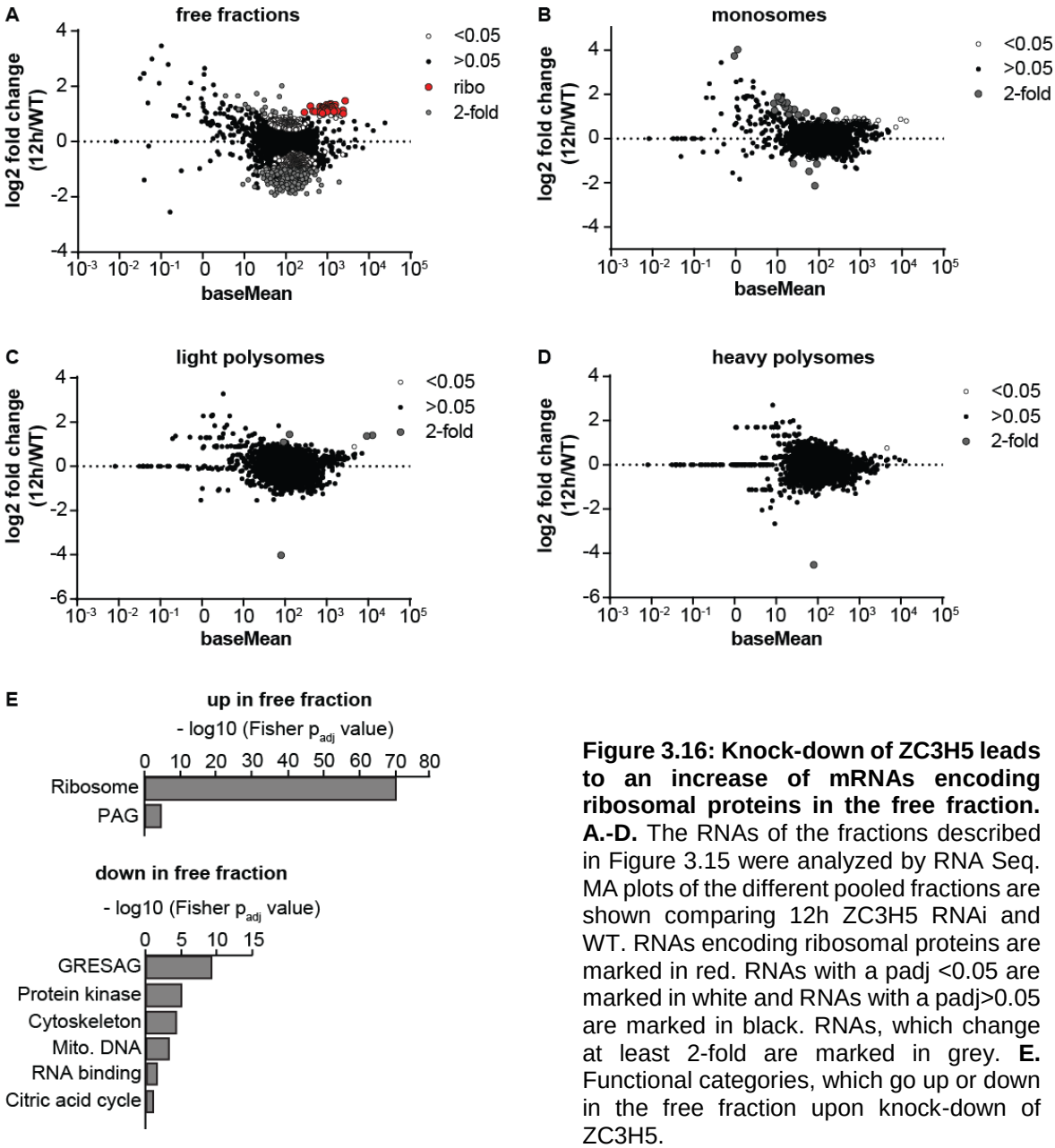


Figure 3.16: Knock-down of ZC3H5 leads to an increase of mRNAs encoding ribosomal proteins in the free fraction. A.-D. The RNAs of the fractions described in Figure 3.15 were analyzed by RNA Seq. MA plots of the different pooled fractions are shown comparing 12h ZC3H5 RNAi and WT. RNAs encoding ribosomal proteins are marked in red. RNAs with a padj <0.05 are marked in white and RNAs with a padj>0.05 are marked in black. RNAs, which change at least 2-fold are marked in grey. **E.** Functional categories, which go up or down in the free fraction upon knock-down of ZC3H5.

Taken together, short-term down-regulation of ZC3H5 leads to an increase of the monosomal peak and the occurrence of half-mers. RNA-Seq of the monosomal peak could not identify a change in the mRNA levels in this fraction. However, an increase of mRNAs encoding ribosomal proteins in the free fractions could be observed. This suggests that the increase of the monosomal peak is due to empty ribosomes and not due to a decrease of translation, because then an increase of mRNAs in the monosomal fraction should be observed.

3.2. Discussion

ZC3H5 is an RNA-binding protein containing a single C3H1-type zinc finger domain that is conserved in Kinetoplastids and it was shown to be essential by RIT-Seq in bloodstream form cells (Alsford et al., 2011). I could show that ZC3H5 is located in the cytosol by immunofluorescence microscopy of a N-terminally tagged version. This coincides with the TrypTag data, which show that a N-terminally tagged version is located in the cytosol and a C-terminally tagged version is located in the cytosol as patchy structures (Dean et al., 2017).

Down-regulation of ZC3H5 rapidly killed bloodstream form trypanosomes. In addition, the proportion of cells in G2/M phase increased rapidly and cells were often arrested in cytokinesis. In Trypanosoma, cytokinesis can be divided in three steps: 1. Initiation of cytokinesis by signaling events; 2. Ingression of cleavage furrow to bisect the cell; 3. Abscission to separate the two daughter cells (Hammarton et al., 2007). The arrest of cells in cytokinesis that we observed upon knock-down of ZC3H5 is most likely a defect of the abscission, because the cells can build a cleavage furrow, but cannot divide. However, in most cases the appearance of cells arrested in cytokinesis is a secondary effect and does not mean that the protein of interest is directly involved in cytokinesis (Hammarton et al., 2007). This is why the phenotype had to be investigated in more detail. To get a clearer idea about how ZC3H5 could function, we were interested in the protein interaction partners as well as in the target mRNAs.

To gain further information about the mRNA targets of ZC3H5, which might help to explain the cytokinesis arrest, RIP-Seq was performed using the described endogenously TAP-tagged ZC3H5 cell line. RIP-Seq identified 79 cytoskeleton enriched ZC3H5 targets and selected targets could be validated by RT-qPCR. Cytoskeleton proteins are known to play a crucial role in cytokinesis in humans, but also other organisms (Fremont and Echard, 2018). The separation of the two daughter cells by abscission might require remodeling of the cytoskeleton and membrane leading to accumulation of proteins at the site of abscission (Hammarton et al., 2007). In Trypanosoma, it is also suggested that the rotational force of the flagellar beat supports abscission. Knock-down of subunits of the dynein regulatory complex, which transmits signals to the axonemal dynein motor, lead to a disturbance of the flagellar beat and to cells arrested in cytokinesis (Ralston et al., 2006). The list of putative targets of ZC3H5 contains several mRNAs encoding dynein proteins and deregulation of this proteins upon knock-down of ZC3H5 could lead to a defect in the rotational force of the flagellar beat. Two recent studies identified two cytokinesis initiation factors, CIF1 and CIF3, which promote cytokinesis (Kurasawa et al., 2018; Zhou et al., 2016). If CIF1 is knocked-down the KAT60/KAT80 complex, which is required for the cytokinesis furrow ingression, is not localized to the new FAZ tip and the cytokinesis is inhibited (Zhou et al., 2016). Secondly, CIF3 interacts with CIF1 and this complex localizes to the new FAZ. On one hand CIF1 stabilizes CIF3 and on the other hand CIF3 sustains the localization of CIF1 at the new FAZ. These data suggest that both proteins, CIF1 and CIF3, are needed for the initiation of cytokinesis (Kurasawa et al., 2018). The mRNAs encoding CIF1 and CIF3 as well as an mRNA encoding for katanin (Tb927.11.3870) and several mRNAs encoding for flagellar proteins (flagellar attachment zone protein 18, paraflagellar rod protein 8, Flagellar member 4, paraflagellar rod protein 10, etc.) were found by RIP-Seq as putative targets of ZC3H5. Taken together, an enrichment of mRNA targets of the group ‘cytoskeleton’ could explain the cell arrest in cytokinesis upon knock-down of

ZC3H5. If the translation of mRNAs encoding cytoskeleton proteins were decreased, the cells would lack proteins needed for proper cytokinesis.

To identify the interaction partners of ZC3H5 an endogenously TAP-tagged ZC3H5 version was used. However, we don't know if it is functional or not, because we never investigated that. But we were not able to knock-out the second copy and since the cells lose the YFP-tag quickly, the tagging might affect the function. Mass spectrometry analysis revealed three putative interaction partners, which were later validated by Co-IPs: Tb927.7.3040, Tb927.8.1500 and Tb927.11.4900. Tb927.11.4900 and Tb927.7.3040 contain WD40 domains while Tb927.8.1500 contains several low complexity regions. Tb927.11.4900 is located in the cytosol according to TrypTag and the localization of Tb927.7.3040 and Tb927.8.1500 is not determined yet (Dean et al., 2017). Further IPs of the three candidates followed by MS analysis showed that the three proteins associate with each other, which indicates that they form a complex. iBAQ analysis of the TAP purification suggests a stoichiometry of 1:2:2 (ZC3H5:Tb927.11.4900:Tb927.7.3040). However, this could not be validated by the pull-downs of the three proteins interacting with ZC3H5. According to the Y2H data Tb927.11.4900 is the linker between ZC3H5 and Tb927.7.3040. All three proteins as well as ZC3H5 are repressors of gene expression according to CAT assays. These results suggest that ZC3H5 and its interacting proteins form a repressive complex. In addition, the complex is conserved in other Kinetoplastid species, which suggests that it could play a role in these species as well.

Interestingly, Tb927.11.4900 is suggested to be a guanine nucleotide-binding beta subunit-like protein (G-protein). These proteins have a GTP-binding domain and hydrolysis of the bound GTP to GDP leads to conformational changes, which can regulate a diversity of functions. G-proteins can be found in all three major kingdoms of life, but mainly in eukaryotes (Wittinghofer and Vetter, 2011). The mass spectrometry analysis of the ZC3H5 interaction partners identified four proteins, which could assist in the exchange of GTP/GDP: Homologue of SDO1 (guanine nucleotide exchange factor), TBC-B (likely GTPase activating protein), ADP-ribosylation factor (GTP/GDP exchange protein) and ARL3C (GTP/GDP exchange protein). All of them are located in the cytoplasm according to TrypTag (Dean et al., 2017), which would enable an interaction with Tb927.11.4900. Since Tb927.11.4900 is supposed to be the linker between ZC3H5 and TB927.7.3040, binding of Tb927.11.4900-GTP to ZC3H5 could recruit Tb927.7.3040, which makes the complex active. Hydrolysis of GTP to GDP might lead to the dissociation of the complex.

RIT-Seq suggested that Tb927.11.4900 and Tb927.8.1500 are essential, while Tb927.7.3040 was not (Alsford et al., 2011). However, I could show that RNAi of all three proteins leads to a strong growth defect. After 4 days, the cells recovered, when the RNAi was not efficient anymore and the protein level increased. In addition, a slight increase of cells with 2N2K, as it was observed for ZC3H5 knock-down, could be observed after depletion of all three interacting proteins with Tb927.8.1500 down-regulation showing the strongest effect. This suggests that knock-down of any protein of the complex leads to an arrest in cytokinesis. However, fluorescence microscopy data (DAPI staining) are missing at the moment. This should be investigated in the future.

To identify the mRNAs that are affected by ZC3H5 knock-down, transcriptomic analysis upon short-term down-regulation of ZC3H5 was performed. Short-term down-regulation was used, because we were interested in catching the primary direct effect. Interestingly,

the abundance of only a few mRNAs changed upon knock-down of ZC3H5, which suggests that knock-down of ZC3H5 does not affect the transcriptome in general. Perhaps as a consequence, the RIP-Seq data show negligible overlap with the transcripts changing at least 1.5-fold. The mRNAs that go up upon ZC3H5 knock-down, are characterized by a long half-life (e.g. mRNAs encoding for ribosomal proteins), which led to the question, if the transcription and the mRNA processing machinery (splicing, transport) were affected. In addition, the transcriptomic data showed an up-regulation of the functional groups 'ribosomal proteins' and 'translation', which suggests that translation could be affected by ZC3H5 knock-down. However, neither general transcription or the mRNA processing machinery nor translation were heavily affected upon knock-down of ZC3H5.

To investigate the effect of ZC3H5 depletion on its target mRNAs in more detail, polysome fractionation upon ZC3H5 knock-down was performed. Knock-down of ZC3H5 resulted in a slight increase of the monosomal peak after 12h induction and a heavy increase after 18h induction. The increase of the monosomal peak upon ZC3H5 knock-down suggests that translation is reduced in these cells. However, only a slight decrease of the heavy polysomes can be observed, which agree with the *de novo* protein synthesis results (methionine labelling). If the translation is inhibited, the amount of polysomes would heavily decrease. The increase of the monosomal peak could be explained by empty ribosomes, which were identified in bacteria as a small and large subunit ribosomal subunit without bound mRNA (Noll et al., 1973), or by mRNAs, which are bound by only one ribosome. In yeast, it was shown by ribosome profiling that monosomes can actively elongate. They translate mRNAs encoding low abundance, regulatory proteins, short ORFs, NMD targets and uORFs in yeast (Heyer and Moore, 2016). In addition, I could observe in *T. brucei* a peak between the 2mer (2 ribosomes attached to the mRNA) and 3mer (3 ribosomes attached to the mRNA) polysome peaks upon knock-down of ZC3H5, which suggested a 2 ½mer, that consists of a mRNA with 2 bound ribosomes and one 40S subunit. It was shown in yeast that this half-mers represent stalled translation initiation complexes (Helser et al., 1981). Half-mers are thought to reflect inefficiencies in translation initiation and to comprise monosomes or polysomes with an additional 43S complex (40S ribosomal subunit with attached initiation factors) at the initiation codon, before the addition of a 60S ribosomal subunit. Half-mers could be observed in *T. brucei* and yeast under conditions that lead to large subunit defects (Jensen et al., 2005; Rotenberg et al., 1988). I suggest that the increase in the levels of transcripts characterized by long half-lives upon ZC3H5 down-regulation might be a consequence of cell cycle arrest. While transcription and mRNA processing continue virtually unaltered, mRNAs begin to accumulate. Thus, an apparent enrichment of mRNAs characterized by long half-lives become apparent relatively to control (uninduced) cells.

To investigate the increase of the monosomal peak in more detail, RNA-seq of free&40S, monosome&60S, light polysomes and heavy polysomes of BF cells and ZC3H5 knock-down after 0h and 12h induction was performed. The only fraction in which I could see a difference between the RNAs in the WT and induced samples was the free fraction. However, due to the pooling of several fractions a shift of mRNAs from, for example, 8 to 5 ribosomes would not have been detected due to insufficient resolution of the experiment. Ribosome profiling would be needed to investigate it in general or northern blots across the polysome gradient for a few of the target mRNAs of ZC3H5. An increase of mRNAs encoding 'ribosomal proteins' (as seen in the transcriptome analysis) and 'PAG' (putative targets of RIP-Seq) and a decrease of the functional groups 'GRESAG', 'Cytoskeleton',

'Mitochondrial DNA', 'RNA binding' and 'citric acid cycle' was observed in the free fraction. This suggests that the increase of the monosomal peak is due to empty ribosomes and not due to a decrease of translation, because then an increase of mRNAs in the monosomal fraction should be observed. An increase of mRNAs in the free fraction could be due to the arrest in cytokinesis: the cells are not dividing, which leads to an accumulation of mRNAs. The phenotype observed is similar to the arrest of stumpy forms, which are not dividing, and stationary procyclic form cells in G0/G1. In the stumpy form, the translation is reduced, which leads to an increase of the monosomal peak and a decrease of the shoulder of heavy polysomes and a shift of the rRNA from the heavy polysome fractions to the free and monosome fractions (Brecht and Parsons, 1998).

Taken together, the increase of mRNAs encoding ribosomal proteins upon knock-down of ZC3H5 in the transcriptome analysis, and the increase of the monosomal peak without an increase of mRNAs in this fraction, the occurrence of half-mers as well as the increase of mRNAs encoding ribosomal proteins in the free fraction upon knock-down of ZC3H5 in the polysome profiles, suggest that the ribosome assembly is disturbed upon knock-down of ZC3H5. However, the defect in ribosome assembly seems to be a secondary effect. Knock-down of ZC3H5 leads to an arrest of cells in cytokinesis, which then leads to an adaptation of the cell to the new situation and to the deregulation of ribosome assembly. The primary function of the ZC3H5 complex is the regulation of mRNAs encoding cytoskeleton proteins and thereby the regulation of cytokinesis (Figure 3.17). The exact mechanism of action is not known at the moment, but it is tempting to speculate that the function of Tb927.11.4900 as G-protein is responsible for the association and dissociation of the complex. This could be cell cycle dependent, because the target mRNAs peak in S-phase, which is before

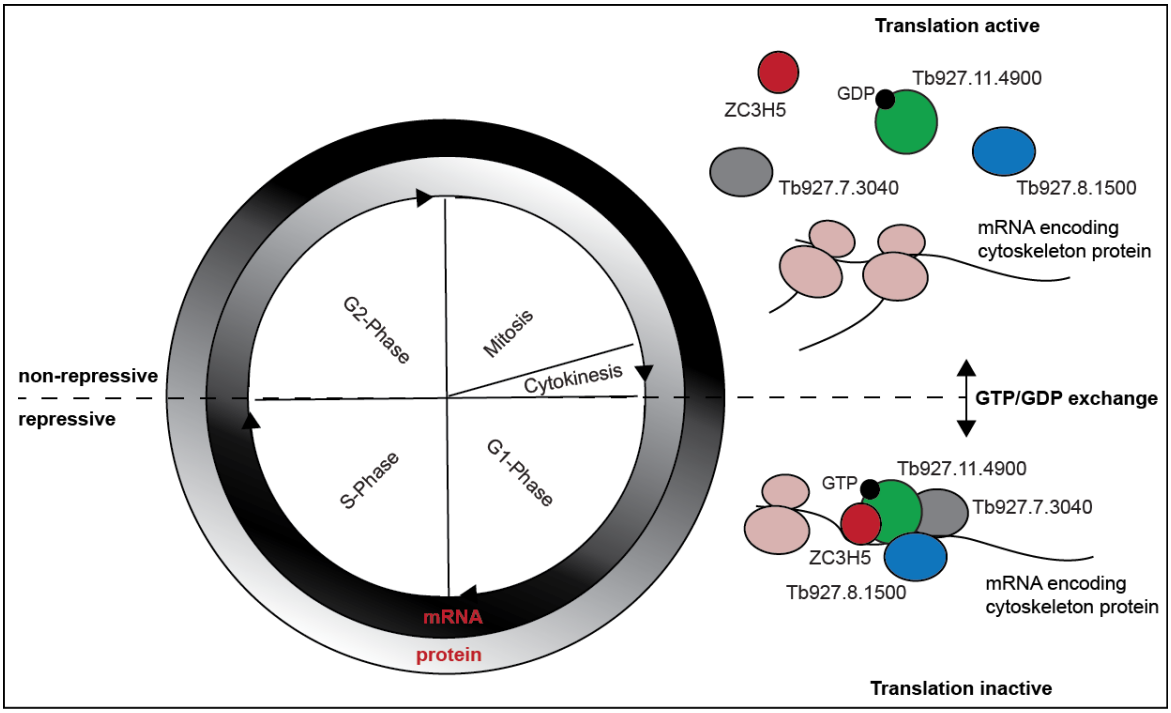


Figure 3.17: ZC3H5 is required for cytokinesis. ZC3H5 binds to mRNAs encoding cytoskeleton proteins. The mRNAs peak in S-phase and the cytoskeleton proteins are needed for cytokinesis. ZC3H5 interacts with Tb927.7.3040, Tb927.8.1500 and Tb927.11.4900 as a repressive complex. Tb927.11.4900 is suggested to be a G-protein, which could interact with GTP. The exchange of GTP to GDP could lead to the disassembly of the complex.

cytokinesis happens and the cytoskeleton proteins are needed during cytokinesis. Maybe the ZC3H5 complex represses its targets during the rest of the cell cycle and targets are de-repressed in S-phase to produce the proteins needed for cytokinesis. However, the remaining question is how a repressive complex can act on the targets in a non-constitutive way? Maybe the assembly/disassembly of the complex is regulated by post-translational modifications of ZC3H5 or one of the other complex proteins or by regulation of the GTP/GDP ratio. ZC3H5 depletion neither affects target turnover (transcriptome does not change) nor the translation heavily (no difference in methionine labelling and still polysomes upon knock-down). Translation might be affected, but not in the extent that we expected. One surprising finding of our study is the translational regulation of the molecular machinery responsible for executing cytokinesis.

4. Material and Methods

4.1. Trypanosoma cell culture

4.1.1. Bloodstream form cells

Monomorphic Lister 427 bloodstream-form trypanosomes were cultured in supplemented HMI-9 medium (components listed below) at 37°C in an incubator with 5% CO_2. The cell concentration was determined by counting using an improved Neubauer counting chamber. All work was done under sterile conditions in a laminar flow hood.

Table 1: Supplemented HMI-9

quantity	ingredients
17.66g/l	Iscove's modified Dulbecco's medium
36 mM	$NaHCO_3$
1 mM	Hypoxanthine
1 mM	Sodium pyruvate
160 mM	Thymidine
50 mM	bathocuprono disulfonic acid disodium salt, pH 6.3
10% (v/v)	Heat-inactivated (56°C, 30 min) FBS
50U/l	Penicillin/Streptomycin
1.5 mM	L-Cysteine
0.14%	β-Mercaptoethanol

4.1.2. Procyclic cells

Monomorphic Lister 427 procyclic trypanosomes were cultured in supplemented MEM-Pros medium (components listed below) at 27°C. The caps of the cell culture flasks were closed tightly and the cells were grown at densities between 0.2-4x10^6 cells/ml. As described above for bloodstream-form cells, an improved Neubauer counting chamber was used for counting and work was done in a laminar flow hood.

Table 2: Supplemented MEM-Pros medium

quantity	ingredients
1 pkg.	MEM-Pros mixture
1 pkg.	MEM vitamins
1 pkg.	MEM non-essential amino acid solution
100 mg	Phenol red
10% (v/v)	Heat-inactivated (56°C, 30 min) FBS
50U/l	Penicillin/Streptomycin
7.5 mg/l	Hemin
0.14%	β-Mercaptoethanol

4.1.3. Antibiotics

For the selection of transgenic trypanosomes, the appropriate antibiotics were added in the following concentrations (see table below). All growth experiments were performed in the absence of antibiotics. For the inducible expression of genes tetracycline was added to a final concentration of 500 ng/ml.

Table 3: Antibiotic concentrations

Antibiotic	Bloodstream form	Procyclic form
Phleomycin	1 µg/ml	1 µg/ml
G418	5 µg/ml	15 µg/ml
Hygromycin	15 µg/ml	50 µg/ml
Puromycin	0.2 µg/ml	1 µg/ml
Blasticidin	5 µg/ml	10 µg/ml

4.1.4. Transfection of bloodstream-form trypanosomes

$2.5x10^7$ cells were pelleted by centrifugation at 2300 rpm for 7 min (centrifuge 5804, Eppendorf). Cell pellet was resuspended in 100 µl transfection buffer (components listed below) and 5-10 µg linearized plasmid was added. The solution was transferred to an electroporation cuvette (2 mm electrodes gap, Peqlab) and electroporation was performed using an electroporation machine and program X-001 (Amaxa Biosystems, Nucleofactor II). The cells were transferred to 25 ml supplemented HMI-9 medium and grown for 6-8 h at 37°C and 5% CO_2. Afterwards, the appropriate antibiotics were added and 1ml of culture was transferred into the first four wells of a 24-well plate each. Culture was diluted by serial dilution throughout the plate. Cells were grown at 37°C and 5% CO_2 and after 5 days wells were analyzed by microscopy to identify single clones.

Table 4: Transfection buffer

quantity	ingredients
90 mM	Sodium phosphate buffer, pH 7.3
5 mM	Potassium chloride
0.15 mM	Calcium chloride
50 mM	HEPES, pH 7.3

4.1.5. Transfection of procyclic trypanosomes

$1.5x10^7$ cells were pelleted by centrifugation at 2300 rpm for 7 min (centrifuge 5804, Eppendorf). Cell pellet was resuspended in 1 ml ZPFM buffer (Zimmerman's Post Fusion Medium, components listed below) and pelleted again by centrifugation as described above. Cell pellet was resuspended in 500 µl transfection buffer and 5-10 µg linearized plasmid was added. The solution was transferred to an electroporation cuvette (2 mm electrodes gap, Peqlab) and electroporation was performed using 1.5 kV and resistance R2 (Electro Cell Manipulator 600, BTX electroporation systems). The cells were transferred to 25 ml supplemented MEM-Pros and grown for 6-8 h at 27°C. Afterwards, the appropriate antibiotics were added and 1ml of culture was transferred into each well of a 24-well plate. Cells were grown at 27°C and after 7 days wells were analyzed by microscopy to identify single clones.

Table 5: ZPFM buffer

quantity	ingredients
132 mM	Sodium chloride
8 mM	Potassium chloride
8 mM	Na_2HPO_4
1.5 mM	KH_2PO_4
1.5 mM	MgAc x 4 H_2O
90 µM	Calcium chloride
Adjust pH to 7.0 with NaOH	

4.2. Cloning

Genes were amplified and cloned into plasmids using standard molecular biology cloning techniques. PCRs were performed using Q5® High-Fidelity DNA Polymerase (NEB) and GoTaq® Hot Start Polymerase (Promega) following the manufacturer's protocol. All restriction enzymes were purchased from NEB as well as the T4 DNA ligase, which was used for ligations. One step Gateway ligations (Fu et al., 2008) were performed using the Gateway LR Clonase II enzyme mix or BP Clonase II enzyme mix (Thermo Fisher Scientific). Gibson cloning (2x Gibson Master Mix, NEB, E2611S) was used to assemble multiple DNA fragments into a plasmid.

4.3. SDS-PAGE and Western Blotting

3-5x10^6 cells were collected per sample, resuspended in 6x Laemmli Buffer and heated at 95°C for 10 min. The samples were subjected to SDS-PAGE gel electrophoresis using 10% polyacrylamide gels. The gels were then stained with SERVA blue G or blotted on a 0.45 µm nitrocellulose blotting membrane (Neolabs). To verify the protein transfer, the membrane was stained with Ponceau S (SERVA). The membrane was blocked with 5% milk in TBS-Tween and incubated with appropriate concentrations of first and secondary antibodies (see table below). Western Lightning Ultra® (Pekin Elmer) was used as chemiluminescence system and signals were detected with the LAS-4000 imager (GE Healthcare) and CCD camera (Fujifilm™).

Table 6: 6x Laemmli buffer

quantity	ingredients
375 mM	Tris-HCl pH 6.8
12%	SDS
45 mM	EDTA
30%	ß-Mercaptoethanol
60%	Glycerol
0.01%	Bromophenol Blue

Table 7: Antibodies for Western Blotting

antibody	company	Product number	host	dilution
Aldolase			rabbit	1:50000
c-myc (9E10)	Santa Cruz Biotechnology	B0614	mouse	1:2000
PAP (Peroxidase-anti Peroxidase)	Sigma	P-2026	rabbit	1:20000
S9			rat	1:1000
TR			rabbit	1:2000
V5	Biorad	MCA1360	mouse	1: 2000
BiP	J. Bangs, Buffalo (from Krauth-Siegel lab)		rabbit	1:1000
CBP	Millipore	07-482	rabbit	1:4000
LipDH			rabbit	1:4000
Scd6	used in A. Singh paper		rabbit	1:10000
Dhh1	from Susanne Kramer		rabbit	1:15000
TRACK	from Esteban Erben		rabbit	1:2000
HA	Roche	11867423001	rat	1:1000
GFP	Santa Cruz		mouse	1:2000
ECL Anti mouse IgG	GE Healthcare	NA931V		1:2000
ECL Anti rabbit IgG	GE Healthcare	NA934V		1:2000
ECL Anti rat IgG	GE Healthcare	NA935V	goat	1:2000
ECL Anti mouse IgG true blot	Rockland	18-8817-33		1:2000

4.4. Digitonin Titration

For each sample $3x10^7$ cells were collected by centrifugation at 2000g for 10 min at 4°C. The pellet was resuspended in 100 µl 1x PBS and centrifuged at 10000g for 5 min at 4°C. Pellet was resuspended in 50 µl STE buffer (components listed below) and centrifuged at 10000g for 5 min at 4°C. A 10 µg/µl digitonin stock solution was heated at 98°C for 5 min and cooled down before use. Seven different digitonin containing solutions, ranging from 0-1.65 µg/µl Digitonin, were prepared and each pellet was resuspended properly in 60 µl of one solution. The samples were incubated at 25°C for 5 min and then centrifuged immediately at 10000g and 4°C for 5min. The supernatant was transferred to another tube containing 20 µl 4x SDS-PAGE sample buffer. The pellet was washed twice with 1x PBS by centrifugation (4°C, 10000g, 5 min) and finally resuspend it in 80 µl 1x Laemmli buffer. Samples were analyzed by Western Blotting.

Table 8: STE buffer

quantity	ingredients
10 mM	Tris-HCl, pH 8.0
0.15 M	NaCl
1 mM	EDTA, pH 8.0

4.5. Tandem affinity purification

4.5.1. 1st step of TAP

$1x10^9$ bloodstream-form trypanosomes with a concentration of $1x10^6$ cells/ml were pelleted by centrifugation at 3000 rpm for 13 min. The pellet was resuspended in 50 ml 1x PBS and the cells were UV-crosslinked (2x2400 µJoules, Stratagene UV crosslinker) in two P15 Petri dishes on ice. The cells were transferred to a Falcon and pelleted by centrifugation at 2300 rpm for 7 min. Cells were snap-frozen in liquid nitrogen. Cells were lysed in 0.5 ml lysis buffer (components listed below) by passing them 20 times through a 21G x 1 ½" needle using a 1 ml syringe and 20 times through a 27G x ¾ needle. Cell debris was pelleted by centrifugation at 10,000 g for 15 min at 4°C. Afterwards the supernatant was transferred to a new tube and salt concentration was adjusted to 150 mM KCl. 250 µl IgG Sepharose 6 Fast Flow beads (GE Healthcare, 17-0969-01) were washed 3 times with 2 ml IPP-150 buffer (components listed below) by centrifugation at 3000 rpm, 3 min, 4°C and let beads settle down for 10 min afterwards. Cell lysate was added to the washed beads and incubated for 2h at 4°C while rotating. Afterwards unbound fraction was collected. Beads were transferred to a 10 ml Poly-Prep Chromatography column (Biorad, 731-1550) and beads were washed 4 times with 10 ml IPP-150 buffer. At the end, beads were transferred in 2 ml IPP-150 buffer to a 2 ml tube and centrifuged as described before to remove the remaining buffer. 0.5 ml TEV cleavage buffer (0.5 ml IPP-150 buffer + 5 µl TEV Protease) was added to the beads and beads were rotated at 16°C for 2h. Beads were centrifuged as described above and elution was collected. Beads were washed again in 500 µl IPP-150 buffer to elute remaining protein.

Table 9: lysis buffer for TAP

quantity	ingredients
20 mM	Tris pH 7.5
5 mM	$MgCl_2$
0.1%	IGEPAL
1 mM	DTT
100U	RNasin
200 µl of 1 pill complete Mini, EDTA-free in 1 ml H2O per $1x10^9$ cells	

Table 10: IPP-150 buffer

quantity	ingredients
20 mM	Tris pH 7.5
5 mM	$MgCl_2$
0.1%	IGEPAL
1 mM	DTT
100U	RNasin
150 mM	KCl

4.5.2. 2nd step of TAP

200 µl calmodulin affinity bead suspension (Agilent Technologies, #214303-52) was transferred to a 10 ml Biorad column and washed 3 times with 10 ml IPP-150 Calmodulin-binding buffer (components listed below). 3 ml Calmodulin-binding buffer and 3 µl $CaCl_2$ was added to 1ml TEV eluate and solution was transferred to the washed beads. Columns

were rotated at 4°C for 1h. The unbound sample was collected and the beads were washed 3 times with 10 ml Calmodulin-binding buffer. Afterwards, the proteins were eluted in 1 ml Calmodulin-elution buffer (components listed below). For the elution the beads were transferred in an Eppendorf tube and rotated for 1h at 4°C. The beads were pelleted by centrifugation at 3000 rpm at 4°C for 3 min and elution was removed from the beads.

Table 11: Calmodulin-binding buffer

quantity	ingredients
add the following ingredients to IPP-150 buffer:	
10 mM	ß-mercaptoethanol
1 mM	magnesium acetate
1 mM	imidazole
2 mM	$CaCl_2$

Table 12: Calmodulin-elution buffer

quantity	ingredients
add the following ingredients to IPP-150 buffer:	
10 mM	ß-mercaptoethanol
1 mM	magnesium acetate
1 mM	imidazole
10 mM	EGTA

4.6. Trichloroacetic acid (TCA)-Acetone precipitation

The proteins of the unbound and eluate fractions were concentrated by TCA precipitation. 6 volumes of 13% TCA in acetone was added to the sample and the proteins were precipitated at 4°C overnight. At the next day, samples were centrifuged at 7000 rpm for 2h at 4°C. The supernatant was removed with a glass pipet keeping 1ml left. The pellet was resuspended in remaining liquid and was transferred to an Eppendorf tube and centrifuged at 13200 rpm for 20 min at 4°C. Supernatant was removed and pellet was washed with 1ml acetone. Finally, the pellet was resuspended in 2x Laemmli buffer and boiled at 95°C for 10 min. Samples were loaded on a 4-12% Bis-Tris 1.5 Nu PAGE gel (#15071080-2289, Novex). The gel was run at 30 mA until the running front reached 1.5 cm in NuPAGE Mops SDS Running Buffer (NP0001; Invitrogen). Gel was stained with colloidal Coomassie blue G-250 and handed to the Mass Spec facility for protein analysis.

4.7. decrosslinking

RNA of unbound and eluate fractions was de-crosslinked from the protein with Proteinase K. 10 µl 10% SDS, 10 µl 0.4 M EDTA and 20 µl Proteinase K (NEB, P8107S) were added to 500 µl sample and incubated for 15 min at 42°C. Afterwards total mRNA was extracted using peqGold Trifast (peqLab) according to manufacturer's protocol. If RNA was used for cDNA synthesis it was, in addition, purified with the Nucleospin RNA purification kit (Macherey-Nagel).

4.8. rRNA depletion

rRNA of the unbound sample was depleted by an RNaseH. RNA was mixed with hybridization buffer (components listed below) and anti-rRNA oligo mix (131 oligos ~50b, in 3,275µl of the mix – 0.5µM conc. of each oligo) and hybridized for 2 min at 95°C. Afterwards the temperature was cooled down to 37°C in steps of 0.1°C per second. RNaseH (Thermo Fisher) and 10x RNaseH buffer (Thermo Fisher) was added and sample was incubated at 37°C for 20 min. DNA was removed by adding Turbo DNase (Ambion) and another incubation at 37°C for 20 min. RNA was cleaned up with RNA clean-up & concentrator columns (Zymo Research, R1015) according to manufacturer's protocol.

Table 13: 5x hybridization buffer

quantity	ingredients
500 mM	Tris pH 7.5
1 M	NaCl

4.9. RNA Sequencing

RNA-seq was done by David Ibberson of the CellNetworks Deep Sequencing Core Facility at the University of Heidelberg. NEBNext Ultra RNA Library Prep Kit for Illumina (New England BioLabs Inc.) was used for library preparation. The libraries were multiplexed and sequenced with a HiSeq 2000 system or NEXTseq system, generating 50 bp single-end sequencing reads. The quality of the raw sequencing data was checked using FastQC (http://www. bioinformatics.babraham.ac.uk/projects/fastqc), and then the sequencing primers were removed using Cutadapt (Martin, 2011). The data was aligned to the T. brucei TREU 927 reference genome using Bowtie2 (Langmead and Salzberg, 2012), then sorted and indexed using SAMtools (Li et al., 2009). Reads aligning to open reading frames of the TREU 927 genome were counted using custom python scripts and a custom pipeline was used (Leiss et al., 2016). Analysis for differentially expressed genes was done in R using the DESeq2 package (Love et al., 2014), using a custom tool for trypanosome transcriptomes (Leiss and Clayton, 2016) which also yields principal component analysis plots. Comparative analysis was limited to a list of unique genes modified from (Siegel et al., 2010). Gene annotations are manually updated versions of those in TriTrypDB and categories were assigned manually. Other statistical analysis was done in R. The 5'UTR, CDS and 3'UTR motif enrichment search was done using DREME (Bailey, 2011); annotated 5'UTR, CDS and 3'UTR sequences were downloaded from TriTrypDB and we considered only the mRNAs with 3'UTRs >8 nt.

4.10. RNA isolation and Northern Blotting

$5x10^7$ cells were used for the extraction of total mRNA using peqGold Trifast (peqLab) according to manufacturer's protocol. The RNA was separated on Agarose-Formaldehyde gels and blotted on a nylon membrane (Amersham Hybond-N+, GE Healthcare, RPN203B). RNA was cross-linked on membrane by UV light (2x240 mJoules) and stained afterwards with methylene blue (SERVA). In case of radioactively labelled DNA probes, the blot was pre-hybridized for 1 h at 65°C with hybridization solution (components listed below). The northern blots were then hybridized with the appropriate probes. [α-^{32}P]dCTP radioactively labelled DNA probes (Prime-IT RmT Random Primer Labelling Kit, Stratagene) were used to detect the mRNAs. Membrane was incubated with the probe overnight at 65°C. At the next day, blot was washed twice with 2xSSC/0.1% SDS at RT for 10 min and once with

1xSSC/0.1% SDS at 65°C for 10 min. To detect the SL RNA, the appropriate oligonucleotide (CZ4490) was labelled with [γ-^{32}P]ATP using T4 polynucleotide kinase (NEB). In this case, that membrane was pre-hybridized for 1 h at 42°C with hybridization solution (components listed below) before the radioactively-labelled oligonucleotide was added and incubated with the membrane overnight at 42°C. At the next day, blot was washed 3x with 6xSSC/0.05% Na-Pyrophosphate at RT for 15 min and once with 6xSSC/0.05% Na-Pyrophosphate at 42°C for 10 min. Afterwards, the blots were exposed to autoradiography films and signal was detected with the phosphoimager. The images were processed using ImageJ.

Table 14: Hybridization solution for DNA probes

quantity	ingredients
2.5 ml	20x SSC
5.9 ml	H_2O
0.5 ml	10% SDS
1 ml	50x Denhardt's Solution
0.1 ml	10 mg/ml Salmon Sperm (denatured at 95°C for 5 min)

Table 15: Hybridization solution for oligonucleotide probes

quantity	ingredients
2.5 ml	20x SSC
5.9 ml	H_2O
0.5 ml	10% SDS
1 ml	50x Denhardt's Solution
0.1 ml	5% Na-pyrophosphate
0.1 ml	10 mg/ml Salmon Sperm (denatured at 95°C for 5 min)

4.11. Immunofluorescence microscopy

Tissue culture glass slides with 8 chambers (Falcon, 354108) were treated with 0.1% Poly-Lysine (Sigma, P-8920). If mitochondria staining was applied, Mitotracker Red CMXRos (Thermo Fisher Scientific) was added to the cells to a final concentration of 50 nM five minutes before collection. For each chamber of the slide, 2.5x10^6 *T. brucei* cells were collected and pelleted by centrifugation (2 min, 2300 rpm). The cells were washed once in 1x PBS. Next, the cell pellet was resuspended in 20 µl 1xPBS and 0.5 ml 4% Paraformaldehyde in 1x PBS was added and incubated for exactly 18 min. Cells were pelleted by centrifugation (2 min, 2300 rpm) and washed 3 times with 1x PBS. Finally, the pellet was resuspended in 200 µl 1x PBS and transferred to the chamber of the chamber glass slide. The chamber glass slide was left at 4°C overnight that a sufficient number of cells can settle down. At the next day, liquid and unbound cells were removed from the chamber and 0.2% (w/v) Triton X-100 in 1x PBS was added. Slide was incubated at RT on a shaker for 20 min. The chamber glass slide was washed 3 times with an excess of 1x PBS to remove the residual Triton X-100. 0.5% (w/v) gelatin in 1x PBS was added to the chambers and incubated at RT on a shaker for 20 min. The blocking solution was removed and the first antibody was added to each chamber diluted in 0.5% gelatin in 1x PBS (dilutions: see table below) and incubated at RT on a shaker for 60 min. Afterwards, slide was washed twice with an excess of 1x PBS and twice with an excess of 0.5% gelatin in 1x

PBS. The washing solution was removed and the secondary antibody was added to each chamber diluted in 0.5% gelatin in 1x PBS (dilutions: see table below) and incubated at RT on a shaker for 60 min in the dark. Slide was washed twice with an excess of 1x PBS followed by 15 min incubation with 1x PBS containing 100 ng/ml DAPI (D9542, Sigma-Aldrich) to stain the nuclear and kinetoplast DNA. Slide was washed twice with an excess of 1x PBS. The chamber scaffold was removed and the glass slide was air-dried. One drop of mounting medium (H-1000, VECTASHIELD) was added to each part of the slide and a cover slide was placed on top. Cover slide was fixed on the glass slide with nail polish and slides were stored at 4°C in the dark until they were analyzed by fluorescent microscopy. All images were examined with the Olympus IX81 microscope. A 100x Oil objective with a numerical aperture of 1.45 was used. Digital images were taken with ORCA-R2 digital CCD camera C10600 (Hamamatsu) and using the xcellence rt software. The bright field images were taken using differential interference contrast (DIC). Fluorescent images were taken as Z-Stacks with a high of roughly 4 µm and a step width of 0.2 µm. The images were deconvoluted (Wiener Filter, Sub-Volume overlap: 20) and then processed using ImageJ. At first the background was subtracted and brightness and contrast were adjusted automatically. The most in focus image of the deconvoluted stack was used.

Table 16: Antibodies for immunofluorescence microscopy

antibody	company	product number	host	dilution
Aldolase			rabbit	1:500
TR			rabbit	1:500
V5	Biorad	MCA1360	mouse	1:200
BiP	J. Bangs, Buffalo (from Krauth-Siegel lab)		rabbit	1:1000
Scd6	used in A. Singh paper		rabbit	1:1000
c-myc	Sigma		mouse	1:1000
Alexa 488 anti-mouse				1:500
Alexa 488 anti-rabbit				1:500
Cy3 anti-mouse				1:500

4.12. Expression and Purification of TEV Protease

Rosetta (DE3)pLysS cells with pHT24 TEV were grown in LB medium containing 10 mg/ml Ampicillin and 40 mg/ml Chloramphenicol at 37°C to an OD_{600} of 0.6 before they were inducted with 1 mM IPTG. Cells were then shifted to 20°C and grown over night. Cells were harvested by centrifugation (5000 g, 4°C, 20 min) and lysed in buffer A (components listed below). For lysis, cells were sonicated 6x for 30 sec. Triton X-100 was added 1/100 to the lysate and cell debris was pelleted by centrifugation (7000 rpm, 4°C, 60 min). The soluble fraction was filtered through a 0.45 µm filter. The cell extract was loaded on Ni-NTA super flow beads (1018611, Qiagen), which were equilibrated with buffer A before, and incubated for 1 h at 4°C while rotating. Beads were washed with wash buffer (buffer A + 20 mM Imidazole). Afterwards TEV protease was eluted with elution buffer (buffer A + 500 mM Imidazole). EDTA (to a final concentration of 2 mM) and DTT (to a final concentration of 10 mM) was added to the eluate and protein concentration was measured with Bradford assay. Enrichment of the TEV protease was investigated by analyzing the different fractions of the purification steps on a 10% SDS gel followed by Coomassie staining. The TEV protease was used for TAP.

Table 17: Buffer A

quantity	ingredients
50 mM	Na_2HPO_4
300 mM	NaCl
5 µg/ml	Leupeptin
0.5 mg/ml	Heparin

4.13. Stress granules purification

Stress granules were purified from $5x10^8$ control or starvation-stressed cells (2h in 1x PBS) procyclic cells as described by (Fritz et al., 2015).

4.14. ^{35}S-Methionine labeling

$5x10^6$ cells were collected by centrifugation (2300 rpm, 7 min) and washed twice with 1x PBS + 0.5% glucose (3000 rpm, 5 min). The pellet was resuspended in 500 µl labeling medium (components listed below) and incubated for 15 min at 37°C. 10 µCi ^{35}S-Methionine was added and incubated for 30 min at 37°C. Cells were washed twice with 1x PBS + 0.5% glucose (3000 rpm, 5 min). Finally, supernatant was removed and pellet was resuspended in 1x Laemmli buffer and boiled at 94°C for 10 min. Samples were run on a 10% SDS gel and stained with Coomassie afterwards. Gel was dried, exposed exposed to autoradiography films and signal was detected with the phosphoimager. The images were processed using ImageJ.

Table 18: Labelling medium

quantity	ingredients
Dulbecco's modified eagle medium (GIBCO) supplemented with:	
25 mM	HEPES
2 mM	Glutamine
0.1 mM	Hypoxanthine
1.5 mM	L-Cysteine
0.0028%	β-Mercaptoethanol
0.05 mM	Bathocuproine sulfate
10%	heat-inactivated FCS previously dialyzed against 30 mM HEPES pH7.3/150 mM NaCl

4.15. CAT assay

Expression of protein of interest tagged lambda-myc was induced with 500 ng/ml Tetracycline overnight. $3x10^6$ cells were collected for a Western Blot. $1x10^7$ cells were pelleted by centrifugation (2300 rpm, 7 min) and washed twice in 1 ml cold 1x PBS (spin down at 3000 rpm for 5 min at 4°C). Finally, the pellet was resuspended in 200 µl 100 mM Tris-HCl, pH 7.8. For lysis, the cells went through two freeze-thaw cycles on dry ice. Cell lysate was centrifuged at 13000 rpm for 3 min at 4°C and supernatant was transferred to a new tube. Protein concentration was determined by Bradford assay and 1 µg protein was used for the CAT assay. 1 µg protein in solution was filled up to 50 µl with 100 mM Tris-HCl, pH 7.8 and mixed with 200 µl 100 mM Tris-HCl, pH 7.8, 2 µl 40 mg/ml Chloramphenicol (Serva, 16785.03) and 10 µl ^{14}C-butyryl-CoA in a scintillation tube (NEB). 4 ml Ultima Gold

F Scintillation Cocktail (PerkinElmer, 6013171) was added and measurement was started using a scintillation counter (LS6000IC, Beckman) measuring ^{14}C.

4.16. Bradford assay

The Bradford assay was used to determine protein concentrations. Bradford reagent (Protein Assay Dye Reagent Concentrate, Bio-Rad, #5000006) was diluted 1:5 in water and 200 µl of this dilution was added in each well of a 96-well plate. 0 µg, 0.5 µg, 1 µg, 2 µg and 2.5 µg BSA was prepared in H_2O for a BSA standard curve, respectively. 10 µl of these solutions were added to separate wells of a 96-well plate. In addition, 5-10 µl of the protein solutions of which the protein amount should be determined, was added to separate wells. The plate was incubated at RT for 5 min and then measured with the plate reader (Tecan). The wavelength at 562 nm was determined using FLUOR4.excel. A BSA standard curve was calculated in Excel and the linear regression line was determined. With the help of the linear regression line the protein concentration was calculated.

4.17. IP using magnetic beads

$1x10^9$ bloodstream-form trypanosomes with a concentration of $1x10^6$ cells/ml were pelleted by centrifugation at 3000 rpm for 13 min. The pellet was resuspended in 50 ml 1x PBS and the cells were UV-crosslinked (2x2400 µJoules, Stratagene UV crosslinker) in two P15 Petri dishes on ice, when the RNA was used for further experiments. The cells were transferred to a conical tube and pelleted by centrifugation at 2300 rpm for 8 min. Cells were snap-frozen in liquid nitrogen. Cells were lysed in 0.5 ml lysis buffer (components listed below) by passing them 20 times through a 21G x 1 ½" needle using a 1 ml syringe and 20 times through a 27G x ¾ needle and salt concentration was adjusted to 300 mM KCl. 1x DNase buffer (components listed below) and 10 µl DNaseI (NEB, M0303S) was added and sample was incubated at 37°C for 10 min. 5 mM stop solution (components listed below) was added, sample was mixed and incubated on ice for some minutes. Cell debris was pelleted by centrifugation at 10,000 g for 10 min at 4°C. Afterwards the supernatant was transferred to a new tube. The protein concentration was determined by Bradford assay and 700 µg protein in 700 µl wash buffer (components listed below) was used for the Co-IP. 30 µl magnetic beads (Dynabeads™ M-280 Tosyl-activated, Invitrogen, 14203) coupled to the appropriate antibody according to manufacturer's protocol were washed 3 times with wash buffer for 3 min each. Magnetic beads were pelleted with the help of a magnetic rack (DynaMag-2 magnet, Invitrogen, 12321D) and the washes were removed by flipping tubes attached to the magnet. Beads were incubated with the protein for 1 h at 4°C on a rotator. Unbound sample was removed and 30µl sample was collected for a Western Blot. Beads were washed 6 times by rotating for 5 min in-between. Proteins were eluted by adding 50 µl 2x Laemmli and boiling at 95°C for 10 min. Tubes were attached to the magnetic rack to remove the eluate. 30 µg total protein, 30 µg unbound sample and half of the eluate were analyzed by SDS-PAGE and Western Blotting.

Table 19: Lysis buffer for IP with magnetic beads

quantity	ingredients
25 mM	Tris pH 7.5
0.1%	IGEPAL
0.5 mM	DTT
100U	RNasin
100 µl	of 1 pill complete Mini, EDTA-free in 1 ml H_2O per $1x10^9$ cells

Table 20: 200x DNase buffer

quantity	ingredients
500 mM	$MgCl_2$
100 mM	$CaCl_2$

Table 21: Stop solution

quantity	ingredients
250 mM	EDTA
250 mM	EGTA

Table 22: Wash buffer

quantity	ingredients
25 mM	Tris pH 7.5
0.1%	IGEPAL
0.5 mM	DTT
100U	RNasin
200 mM	KCl

4.18. Mass Spectrometry analysis

Proteins that co-purified with BFR1L, Tb927.8.1500, Tb927.7.3040 and Tb927.11.4900 were analyzed in three independent experiments by LC/MS by the ZMBH Mass Spectrometry facility. Cell lines expressing TAP-GFP served as control for TAP-BFR1L pull-down. V5-BFR1L served as control for V5-Tb927.8.1500, V5-Tb927.7.3040 and V5-Tb927.11.4900. Raw data were analyzed using MaxQuant 1.5.8.3, with label-free quantification (LFQ), match between runs (between triplicates), and the iBAQ algorithm enabled. The identified proteins were filtered for known contaminants and reverse hits, as well as hits without unique peptides. Statistical analysis was performed in Perseus (Tyanova et al., 2016). Data were filtered for at least two valid values in at least one condition and remaining missing values were imputed with a normal distribution based on the whole data set (width = 0.3; shift = 1.8). We determined significant outlier with t-test statistics (permutation-based false discovery rate of 1% and S0 of 1).

4.19. Yeast two-hybrid assay

The ORFs of the genes of interest (ZC3H5, CAF1, Tb927.8.1500, Tb927.11.4900, Tb927.7.3040) were amplified by PCR and cloned into pBD-gate2 and pAD-GW (Maier et al., 2008). As negative controls the pGBKT7 plasmid containing Lamin and pAD-T7 were used. The pBD-gate2 plasmid was used as bait and contains an N-terminal-GAL4 DNA binding domain and a myc-tag. The pAD-GW plasmid was used as prey and contains an N-terminal fused GAL4 activation domain and a HA-tag. The Matchmaker Yeast Two-Hybrid System (Clontech) was used according to the manufacturer's protocol for pairwise co-transformation of the bait and prey plasmids into the AH109 yeast strain. Clones expressing the bait and prey plasmid were selected on double dropout medium (minimal SD medium lacking tryptophan and leucine). Positive interactions were indicated by the change to blue color on quadruple dropout medium (minimal SD media lacking tryptophan, leucine,

histidine and adenine) containing X-α-gal. Western Blotting confirmed the expression of the proteins containing the myc- and HA-tags. In addition, the plasmids were extracted (see DNA isolation of yeast) from the different clones and PCRs using primer pairs, which are specific for the different ORFs, were performed. The plasmids were also retransformed in DH5α for amplification and analyzed by Sanger sequencing.

4.20. DNA isolation of yeast

A yeast colony was inoculated in 500 µl SD medium overnight at 30°C by shaking at 250 rpm. Cells were pelleted by centrifugation at 14000 rpm for 5 min. The pellet was resuspended in 40 µl SD medium and 10 µl Lyticase (5 units/µl, Sigma-Aldrich, L4025) was added and incubated at 37°C for 60 min by shaking at 250 rpm. 20 µl 10% SDS was added and vortexed for 1 min. Samples went through three freeze-thaw cycles at -20°C. Plasmids were then purified using the NucleoSpin Plasmid Purification Kit (Macherey-Nagel, 740588.50) according to the manufacturer's protocol.

4.21. qPCR

RNA was extracted using peqGold Trifast (peqLab) according to manufacturer's protocol and cDNA was synthesized using Maxima First Strand cDNA synthesis kit for RT-qPCR (Thermo Scientific, K1671) according to manufacturer's protocol. For qPCR the Luna Universal qPCR Master Mix (NEB, #M3003S) was used according to manufacturer's protocol and samples were measured in triplicates using different cDNA concentrations. Tubulin was used as loading control and Ct values were normalized according to tubulin.

4.22. Polysome fractionation

On the day before the experiment, sucrose gradients were prepared out of 5 different sucrose solutions in polysome buffer (50%, 42.5%, 35%, 22.5% and 15%; components listed below) in polyallomer 14 x 89mm tubes (Beckman). To pour the gradients, 790µL of each sucrose solution was added sequentially in a polyallomer tube, starting with 50% sucrose. Tubes were covered with aluminum foil and frozen at -80°C for 20min, before adding the next sucrose solution. In the evening before the polysome fractionation experiment, gradients were transferred to the cold room to thaw them. $5x10^8$ cells per gradient were collected by centrifugation at 3000 rpm for 15 min. The cell pellet was resuspended in 50 ml serum-free medium and transferred to a 50 ml conical tube. The cells were incubated with 100µg/ml Cycloheximide for 7 min at RT. Cells were pelleted by centrifugation at 2300 rpm for 7 min at 4°C and washed with 1 ml ice-cold 1x PBS and transferred to an Eppendorf tube. Cells were lysed in 350 µl lysis buffer (components listed below) and passed 15 times through the 21-gauge needle using a 1 ml syringe and then 15 times through the 27-gauge needle. The lysate was cleared by centrifugation at 15000 g for 10 min at 4°C in a microfuge. Salt concentration was adjusted to 120 mM KCl and the lysate was loaded on the 4 ml continuous linear 15-50% sucrose gradient. The gradients were centrifuged at 40000 rpm in the Beckmann SW60 centrifuge for 2h at 4°C using the swinging bucket rotor. Afterwards, 16 fractions with a volume of 300 µl were collected by fractionation with the UV/VIS detector (Teledyne Isco). For RNA purification, 900µL TriFast was added to each tube and RNA was purified as described above.

Table 23: Polysome buffer

quantity	ingredients
20 mM	Tris pH 7.5
2 mM	$MgCl_2$
10 µg/ml	Leupeptin
1 mM	DTT
100 µg/ml	Cycloheximide
120 mM	KCl

Table 24: Lysis buffer polysome fractionation

quantity	ingredients
20 mM	Tris pH 7.5
20 mM	KCl
2 mM	$MgCl_2$
2 mM	DTT
1000U	RNasin
10 µg/ml	Leupeptin
0.2%	IGEPAL
200 mM	sucrose
100 µg/ml	Cycloheximide
1 pill complete Mini, EDTA-free per $1x10^9$ cells	

4.23. FACS analysis

$2x10^7$ cells were collected by centrifugation (1500 rpm, 10 min, 4°C). Cells were washed twice with 5 ml cold TDB buffer (components listed below) + 2 mM EDTA and finally resuspended in 200 µl TDB buffer + 2 mM EDTA in a 15 ml conical tube. Cells were fixed by adding 2ml of ice-cold 70% ethanol dropwise while vortexing on high speed. Cells can be stored at 4°C. Nuclear and kinetoplast DNA was stained with Propidium iodide for FACS analysis. 1 ml of cells in 70% Ethanol were pelleted by centrifugation (1500 rpm, 10 min, 4°C) and resuspended in 500 µl staining solution (components listed below). Cells were incubated for 30 min at 37°C and then directly analyzed by FACS.

Table 25: TDB buffer

quantity	ingredients
20 mM	Na_2HPO_4
2 mM	NaH_2PO_4
80 mM	NaCl
1 mM	$MgSO_4$
20 mM	Glucose
5 mM	KCl

Table 26: Propidium iodide staining solution (per ml)

quantity	ingredients
930 µl	1x PBS + 2 mM EDTA
20 µl	RNAse A (10 mg/ml)
50 µl	Propidium iodide (1 mg/ml)

4.24. Genomic DNA extraction

The cells were lysed in EB buffer in presence of RNAse A. 5 M ammonium acetate was used to precipitate proteins and cells debris. The DNA was then precipitated with isopropanol, washed with ethanol and resuspended in water.

Table 27: EB buffer

quantity	ingredients
10 mM	Tris-HCl pH 8
10 mM	NaCl
10 mM	EDTA
5%	SDS

4.25. Oligonucleotide list

Table 28: Oligonucleotide list

number	explanation	sequence
a14150	fwd; 400 nt insert of Tb927.10.14150 used for RNAi (S. Lueong)	CACGATCCGCGGATGAGCAAGACAGAGAC GG
a14150	rev; 400 nt insert of Tb927.10.14150 used for RNAi (S. Lueong)	CACGATTCTAGAGCCTCGTCGAACTCCTC
CZ2351	rev; for colony PCR of p2T7 with insert	CCGCTCTAGAACTAGTGGA
CZ2698	rev; CAT probe	GAAAGACGGTGAGCTGGT
CZ3634	fwd; CAT	GCCGCTGGCGATTCAG
CZ3634	fwd; CAT, also for CmR fwd in empty Y2H plasmids	GCCGCTGGCGATTCAG
CZ3798	fwd; Tubulin for qPCR	
CZ3799	rev; Tubulin for qPCR	
CZ4049	fwd; beta-Tubulin	AGGCTGGCCAATGCGGTAAC
CZ4051	rev; beta-Tubulin	CTCCGTGCACCAACTTGTGG
CZ4490	spliced leader for 5'end labelling	CAATATAGTACAGAAACTGTTCTAATAATAGC GTTAGT
CZ4615	fwd; CAT NB probe	ATGGAGAAAAAAATCATCGGATAT
CZ4650	fwd; pray (AD) for sequencing	
CZ4651	fwd; bait for sequencing	
CZ5184	rev; pHD1146 sequencing	TCATCCAACAAATTAAACTGCAG
CZ5496	fwd; Tb927.7.2780	ATGTCTAAAGCTCCTTCGCAA
CZ5497	rev; Tb927.7.2780	CTGAGACAATCCATTAACTCC
CZ5598	fwd; pray (AD) for sequencing	
CZ5711	rev; Actin	GGCATAGGGCTGAGTACAGGCACCAC
CZ6020	fwd; 400 nt Tb927.10.14150 mini-ORF for TAP-tagging, HindIII-site	GATCAAGCTTCCATGAGCAAGACAGAGACG
CZ6021	rev; 400 nt Tb927.10.14150 mini-ORF for TAP-tagging, ApaI-site	GATCGGGCCCGCCTCGTCGAACTCCTC
CZ6022	fwd; 400 nt Tb927.10.14150 5'UTR for Tap-tagging, SacI-site	GATCGAGCTCTTGCCACCATTTGAATTTA
CZ6023	rev; 400 nt Tb927.10.14150 5'UTR for Tap-tagging, NdeI-site	GATCCATATGTGCTGCTGATGCTTTTTG
CZ6093	fwd; AttL1universal primer for 2nd PCR, Gateway	CCCCGATGAGCAATGCTTTTTTATAATGCCAA CTTTGTACAAAAAAGCAGGCTCCAT
CZ6094	rev; AttL1universal primer for 2nd PCR, Gateway	GGGGGATAGCAATGCTTTCTTATAATGCCA ACTTTGTACAAGAAAGCTGGGT
CZ6183	fwd; 300 nt Tb927.10.14150 5'UTR, XhoI-site	GATCCTCGAGTCCCTTTTTTTTTCTGCTCT

CZ6184	rev; 300 nt Tb927.10.14150 5'UTR, HindIII-site	GATCAAGCTTTGCTGCTGATGCTTTTTGG
CZ6185	fwd; 300 nt Tb927.10.14150 3'UTR, EcoRI-site	GATCGAATTCAATAATTGTGTCGTGAGGTA
CZ6186	rev; 300 nt Tb927.10.14150 3'UTR, SacII-site	GATCCCGCGGCCCAATAATATTCAACGTGA
CZ6233	fwd; 5'UTR 14150 to test the replacement of Tb927.10.14150 by BLA	CCGACGATAGCGCCATGTGTTTG
CZ6234	rev; BLA	ATGATATACATTGACACCAGTGAAGATGC
CZ6235	fwd; BLA	CTGACTTGTATCGTCGCGATCGGAAATG
CZ6236	rev; 3'UTR 14150 to test the replacement of Tb927.10.14150 by BLA	CACCTTCTCACTAACCAGTTGATTGTTAT
CZ6252	fwd; 300 bp end of Tb927.10.14150 CDS, KpnI-site	GATCGGTACCCAAGCCAATAAAGCAAAGAGG GATTC
CZ6253	rev; 300 bp end of Tb927.10.14150 CDS, XhoI-site	GATCCTCGAGCGCAAAATCCTCCCCCTCTCC
CZ6254	fwd; 300 bp beginning of Tb927.10.14150 3'UTR, BamHI-site	GATCGGATCCATAATTGTGTCGTGAGGTAGG GATGTTGC
CZ6255	rev; 300 bp beginning of Tb927.10.14150 3'UTR, NotI-site	CCCCAATAATATTCAACGTGATATTCTTTTTT CCGCGGCCGCGATC
CZ6379	rev; Tb927.10.14150 5'UTR for Tap-tagging, NdeI-site plus ClaI-site	GATCCATATGGATCATCGATTGCTGCTGATG CTTTTTG
CZ6415	rev; 400 nt Tb927.10.14150 5'UTR for Tap-tagging, NdeI-site plus NotI-site	GATCCATATGGATCGCGGCCGCTGCTGCTGA TGCTTTTTG
CZ6417	fwd; 400 nt Tb927.10.14150 5'UTR for Tap-tagging, SacI-site plus HpaI-site	GATCGAGCTCGATCGTTAACTTGCCACCATTT GAATTTA
CZ6418	fwd; RNAi stemloop Tb927.10.14150	GATCAAGCTTAGATCTCAGCGCAAAGATCAC CAGTA
CZ6419	rev; RNAi stemloop Tb927.10.14150	GATCGTCGACGAATTCATTGGCTTGTTGCTT CTGCT
CZ6497	fwd; RT-PCR Tb927.4.1860 CDS	CCTGAGATTACAACTCGCGC
CZ6498	rev; RT-PCR Tb927.4.1860 CDS	ACTTTGCCTTTGTCTCAGCG
CZ6505	fwd; RT-PCR Tb927.9.7590 CDS	CTGACACAGGTTCCTTTGGC
CZ6506	rev; RT-PCR Tb927.9.7590 CDS	TTCATGGCCTCCTCCTTACG
CZ6512	fwd; Puro resistance, HindIII-site	GATCAAGCTTATGACCGAGTACAAGCCCA
CZ6513	rev; Puro resistance, EcoRI-site	GATCGAATTCATCAGGCACCGGGCTT
CZ6538	fwd; BLA V5, XbaI-site	TCTAGAATGGCCAAGCCTTTGTCT
CZ6539	rev; BLA V5, XhoI-site	CTCGAGCGTAGAATCGAGACCGAGGA
CZ6548	fwd; Tb927.10.14150, beginning of CDS, EcoRI-site	GATCGAATTCATGAGCAAGACAGAGA
CZ6551	rev; ZC3H5	CCGCGGAGACACCCACCA
CZ6557	fwd; Tb927.10.14150 beginning of CDS, ApaI-site	GATCGGGCCCATGAGCAAGACAGAGAC
CZ6570	fwd; Tb927.7.3040 ORF for Northern probe	CTGTCACGGAGGAGGTGGAC
CZ6571	rev; Tb927.7.3040 ORF for Northern probe	ACCTTGTTGTTCTCCGGGGT
CZ6572	fwd; Tb927.11.10340 ORF for Northern probe	GGAGGAACAGGCGAGACTCA
CZ6573	rev; Tb927.11.10340 ORF for Northern probe	TATTGTGGGACACGCCAATG
CZ6574	fwd; Tb927.3.1660 ORF for Northern probe	GTCGAGCGGAGTGGTAGTGG
CZ6575	rev; Tb927.3.1660 ORF for Northern probe	GCAACCACGAAAGCGAAAAG
CZ6578	rev; Tb927.7.3040	CGACTCCACTGGTTTGCT
CZ6580	rev; Tb927.11.4900	GCTCATTGTCAGGAAGCTG
CZ6639	rev; Tb927.10.14150, end of CDS, BamHI-site	GATCGGATCCCGCAAAATCCTCCCCCC
CZ6640	fwd; RT-PCR Tb927.11.14020	TTGTCAGTGCTCAGATCCGT
CZ6641	rev; RT-PCR Tb927.11.14020	CCTTCACAAACTTCTGCGCT
CZ6642	fwd; RT-PCR Tb927.9.3920	GAGTCCCCGTCACAAGAAGA

CZ6643	rev; RT-PCR Tb927.9.3920	ACTTGTAGACGTGGGCAAGA
CZ6644	fwd; RT-PCR Tb927.7.2340	CGAACATTACACAGGAGCGG
CZ6645	rev; RT-PCR Tb927.7.2340	TGGCCAATCATCTCACCCTT
CZ6683	fwd; AttL1-Tb927.7.3040 primer for 1st PCR, Gateway	AAAGCAGGCTCCATGATTGACCCATTTCGC
CZ6684	rev; AttL1-Tb927.7.3040 primer for 1st PCR, Gateway	GTACAAGAAAGCTGGGTTCAGTCTCTCACGG AATACA
CZ6685	fwd; AttL1-Tb927.11.4900 primer for 1st PCR, Gateway	AAAGCAGGCTCCATGATGCGAATCAAGGTAG ACG
CZ6686	rev; AttL1-Tb927.11.4900 primer for 1st PCR, Gateway	GTACAAGAAAGCTGGGTTTACGCAGCATAC GCTTC
CZ6691	fwd; AttL1-Tb927.10.14150 primer for 1st PCR, Gateway	AAAAAGCAGGCTCAATGAGCAAGACAGAGAC
CZ6723	rev; AttL1-Tb927.10.14150 primer for 1st PCR, Gateway	AGAAAGCTGGGTTACGCAAAATCCTCCCCCT
CZ6724	rev; Tb927.10.14150, NotI-site	GATCGCGGCCGCTGCAGCTCCTGCAAGTA
CZ6725	fwd; Hygro resistance, SacI-site	GATCGAGCTCATGAAAAAGCCTGAACTCA
CZ6726	rev; Hygro resistance, BstBI-site	GATCTTCGAACTATTCCTTTGCCCTCGGA
CZ6727	fwd; Hygro resistance, NdeI-site	GATCCATATGATGAAAAAGCCTGAACTCA
CZ6732	fwd; Puro resistance, NdeI-site	GATCCATATGACCGAGTACAAGCCCAC
CZ6733	rev; Puro resistance, BstBI-site	GATCTTCGAATCAGGCACCGGGCTTGC
CZ6764	fwd; end of Tubulin	AGGTACTAGCACCACTAAC
CZ6782	fwd; AttL1-Tb927.11.4900 primer for 1st PCR, Gateway	AAAAAGCAGGCTCAATGCGAATCAAGGTA GAC
CZ6783	rev; AttL1-Tb927.11.4900 primer for 1st PCR, Gateway	AGAAAGCTGGGTTTACGCAGCATACGCTT
CZ6784	fwd; AttL1-Tb927.7.3040 primer for 1st PCR, Gateway	AAAAAGCAGGCTCAATGATTGACCCATTTC GCT
CZ6785	rev; AttL1-Tb927.7.3040 primer for 1st PCR, Gateway	AGAAAGCTGGGTTCAGTCTCTCACGGAATAC
CZ6786	fwd; AttL1-Tb927.8.1500 primer for 1st PCR, Gateway	AAAAAGCAGGCTCAATGGCAATGCAGTTAT TTA
CZ6787	rev; AttL1-Tb927.8.1500 primer for 1st PCR, Gateway	AGAAAGCTGGGTTCAGTCATCGCAGATTC
CZ6823	fwd; G418	CACCCGCGCTGGGTGGAAAGCTAGCTTTA ATTTGTTGGATGAGCTATTTCATTAATTTTT TTG
CZ6824	rev; G418	CACCTGTGGCGCCGGTGATGCCGGCAATA CTGCATAGATAACAAACGCATC
CZ6825	fwd; Tb927.8.1500 5'UTR	CGACGTTGTAAAACGACGGCCAGTGAATT CCTTTAAGCATCTCTCTGTAGTAGGGTTATG
CZ6826	rev; Tb927.8.1500 5'UTR	ATTCTTCTTGAGACAAAGGCTTGGCCATTGC CCAGAAGATGCTGTATCC
CZ6827	fwd; Tb927.8.1500 ORF	CCCTCTCCTCGGTCTCGATTCTACGATGGCA ATGCAGTTATTTACCTTTGG
CZ6828	rev; Tb927.8.1500 ORF	AACAGCTATGACCATGATTACGCCAAGCTTC CGCGTAAGCAAGCAGTT
CZ6829	fwd; Tb927.11.4900 5'UTR	TTGTAAAACGACGGCCAGTGAATTCATTTTT GTTCTTTTCCATCATC
CZ6830	rev; Tb927.11.4900 5'UTR	GCTTGGCCATTTCCCTTCCTTATTTCCTTG
CZ6831	fwd; BLA-V5-Tb927.11.4900	AGGAAGGGAAATGGCCAAGCCTTTGTCTC
CZ6832	rev; BLA-V5-Tb927.11.4900	TGATTCGCATCGTAGAATCGAGACCGAG
CZ6833	fwd; Tb927.11.4900 ORF	CGATTCTACGATGCGAATCAAGGTAGACG
CZ6834	rev; Tb927.11.4900 ORF	CTATGACCATGATTACGCCAAGCTTCTGCA ATCCAGACGACGC
CZ6835	fwd; Tb927.7.3040 5'UTR	TTGTAAAACGACGGCCAGTGAATTCACCGA CTGACCGAAGTTTAG
CZ6836	rev; Tb927.7.3040 5'UTR	GCTTGGCCATCTAGGACAGTTTCCTACTT

		GAC
CZ6837	fwd; BLA-V5-Tb927.7.3040	ACTGTCCTAGATGGCCAAGCCTTTGTCTC
CZ6838	rev; BLA-V5-Tb927.7.3040	GGTCAATCATCGTAGAATCGAGACCGAG
CZ6839	fwd; Tb927.7.3040 ORF	CGATTCTACGATGATTGACCCATTTCGC
CZ6840	rev; Tb927.7.3040 ORF	CTATGACCATGATTACGCCAAGCTTATACA TTTTGCAACACATTC
CZ6843	fwd; Tb927.7.3040 ORF, AttB site	GGGACAAGTTTGTACAAAAAAGCAGGCTCCT GTCACGGAGGAGGTGGAC
CZ6845	fwd; Tb927.11.4900 ORF, AttB site	GGGACAAGTTTGTACAAAAAAGCAGGCTCCC ATGCACGTTATACCGGAC
CZ6847	fwd; Tb927.8.1500 ORF, AttB site	GGGACAAGTTTGTACAAAAAAGCAGGCTCCC CCATCGCCTTGAAGATTG
CZ6849	fwd; ZC3H5 ORF, AttB site	GGGACAAGTTTGTACAAAAAAGCAGGCTCGG GGAACACAAACAGGTAGC
CZ6850	rev; ZC3H5 ORF, AttB site	ACCCAGCTTTCTTGTACAAAGTGGTCCCCCA TTCACCTGCACTGTTCC
CZ6855	fwd; Tb927.2.4550, NB probe	CGGATGGTTTTACGCACGAT
CZ6856	rev; Tb927.2.4550, NB probe	CCTTCGTCTTCCCCTTCACT
CZ6871	rev; Tb927.7.3040 ORF, AttB site	GGGACCACTTTGTACAAGAAAGCTGGGTAC CCCGGAGAACAACAAGGT
CZ6872	rev; Tb927.11.4900 ORF, AttB site	GGGACCACTTTGTACAAGAAAGCTGGGTCG CGTGAACATATCCAACGT
CZ6873	rev; Tb927.8.1500 ORF, AttB site	GGGACCACTTTGTACAAGAAAGCTGGGTTT GCTGCTATTGTTGCCGTT
CZ6876	rev; ZC3H5 ORF, AttB site	GGGACCACTTTGTACAAGAAAGCTGGGTCCA TTCACCTGCACTGTTCC
CZ6877	fwd; ZC3H5, NB probe	GGGGAACACAAACAGGTAGC
CZ6878	rev; ZC3H5, NB probe	CCATTCACCTGCACTGTTCC
CZ6879	fwd; BLA-V5-Tb927.8.1500	GGATACAGCATCTTCTGGGCAATGGCCAAG CCTTTGTCT
CZ6880	rev; BLA-V5-Tb927.8.1500	GTAAATAACTGCATTGCCATCGTAGAATCGA GACCGAGGA
CZ6907	fwd; primer for sequencing of pGL 2084	AAAGTAGCGCTTACGGCGT
CZ6908	rev; primer for sequencing of pGL 2084	ATTCCTGCAGGGGCCCT
CZ6984	pre-18S rRNA	TCAAGTGTAAGCGCGTGATCCGCTGTGG
CZ6985	pre-5.8S rRNA	CCATCGCGACACGTTGTGGGAGCCG
CZ6990	fwd; Puromycin, EcoRI-site	GATCGAATTCATGACCGAGTACAAGCCC
CZ6991	rev; Puromycin, NcoI-site	GATCCCATGGATCAGGCACCGGGCTTG
CZ6992	fwd; G418, EcoRI-site	GATCGAATTCATGATTGAACAAGATGGA
CZ6993	rev; G418, NcoI-site	GATCCCATGGTCAGAAGAACTCGTCAAG
CZ7032	fwd; V5-Tb927.11.4900, BamHI-site	GATCGGATCCATGGGTAAGCCTATCCCTAA CCCTCTCCTCGGTCTCGATTCTACGATGCG AATCAAGGTAGACG
CZ7033	rev; Tb927.11.4900, EcoRI-site	GATCGAATTCCGCAGCATACGCTTCAGTC
CZ7053	rev; G418, SalI-site	GATCGTCGACTCAGAAGAACTCGTCAAG
CZ7079	rev; RNAi Tb927.7.3040, BamHI-site	GATCGGATCCACCCCGGAGAACAACAA
CZ7080	fwd; RNAi Tb927.7.3040, XhoI-site	GATCCTCGAGCTGTCACGGAGGAGGTG
CZ7129	fwd; Tb927.10.10350 for qPCR	TGCACCGGGACATCAAGGGG
CZ7130	rev; Tb927.10.10350 for qPCR	CCGTAACCTCCAGCCTCGCC
CZ7133	fwd; Tb927.10.7880 for qPCR	GCGGGAGGAAGATGCGGAGG
CZ7134	rev; Tb927.10.7880 for qPCR	CGGAGGGCCTCGATGAGACG
CZ7135	fwd; Tb927.11.10900 for qPCR	GAGGCCCTCAACGAGCGACA
CZ7136	rev; Tb927.11.10900 for qPCR	CATCACACACTGCGCGCGTT
CZ7137	fwd; Tb927.9.15050 for qPCR	GTGACCGGGACAAGGCGGAG
CZ7138	rev; Tb927.9.15050	CAAGGGTCCGTTCGGCCTCA
CZ7141	fwd; Tb927.11.5810 for qPCR	TCCCAACAGCAGTCGACCGC
CZ7142	rev; Tb927.11.5810 for qPCR	ACCACCTCCACGCTGCCCTA
G1887	fwd; ß-Globin (AG Stoecklin)	TCAGATCGCCTGGAGACG

G1888	rev; ß-Globin (AG Stoecklin)	CTTGAGCATCTGACTTCTGGCT
N1Tb927.10.14150	fwd; for N-terminal V5-tagging of Tb927.10.14150 (S. Lueong)	TAGATTGAGGAAGTTGCCTCATTAATTTGCACGATTAGTGTACCAAGGGGGAAAGGTTCTGCCAAAAAGCATCAGCAGCAATGGCCAAGCCTTTGTCTCAAG
N1Tb927.10.14150	rev; for N-terminal V5-tagging of Tb927.10.14150 (S. Lueong)	GGAGCTGTCATCCACCGCGGTTGAGGACGCCTTTCCGTCGGCGCGGGCGCAGCTTCAGGTGCCGTCTCTGTCTTGCTCATCGTAGAATCGAGACCGAGGAGAGG
	rev; T7	TAATACGACTCACTATAGGG
	fwd; T3	AATTAACCCTCACTAAAGGG
	fwd; T7	TAATACGACTCACTATAGGG
	fwd; AttB universal primer	GGGACAAGTTTGTACAAAAAAGCAGGCTC
	rev; AttB universal primer	GGGACCACTTTGTACAAGAAAGCTGGGT
	fwd; lambda	ACCAAAAAGTAAAATTCACAAGCTTATGGACGCACAAACACGACGAC
	fwd; ß-Globin for NB probe	GGTGAAGGCTCATGGCAAG
	fwd; RNAi Tb927.11.4900	GGGGACAAGTTTGTACAAAAAAGCAGGCTCAGAGGGTTCTAAAAAACAATCG
	rev; RNAi Tb927.11.4900	GGGGACCACTTTGTACAAGAAAGCTGGGTATAGTTTTTGTCTGTTTCCAACTC
	fwd; Tb927.11.4900 over expression	ACCAAAAAGTAAAATTCACAAGCTTATGGGCGGCCGATCGAGATCC
	rev; Tb927.11.4900 over expression	AAAGCCAACTAAATGGGCAGGATCCTTACGCAGCATACGCTTCAGTCTCAGG

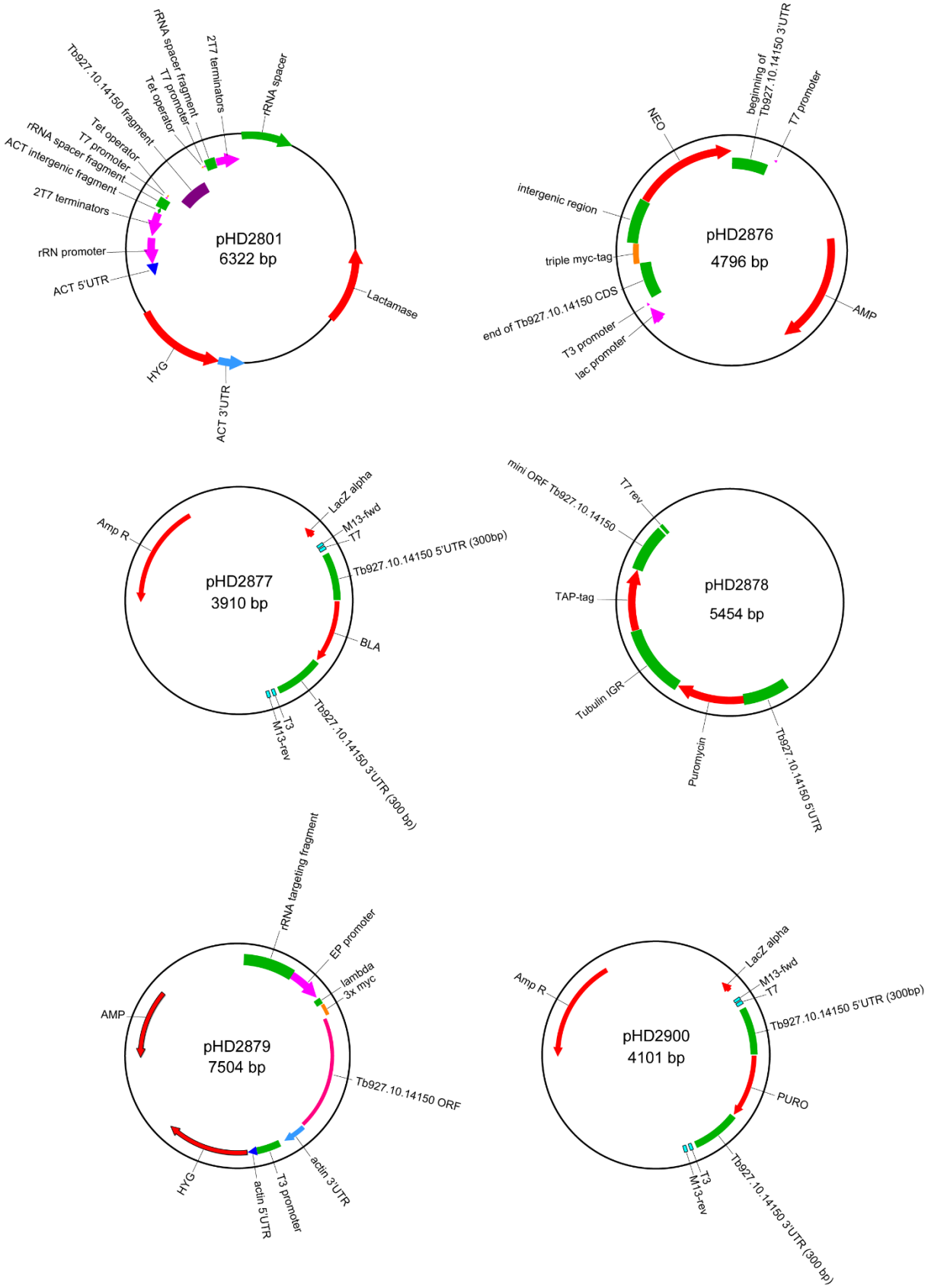
4.26. Plasmid list

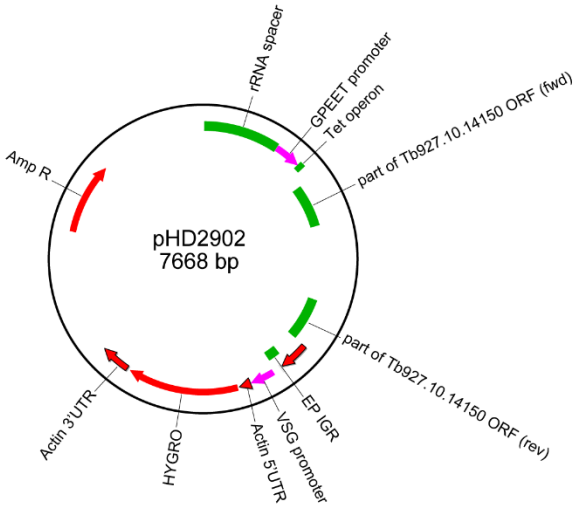
Table 29: Plasmid list

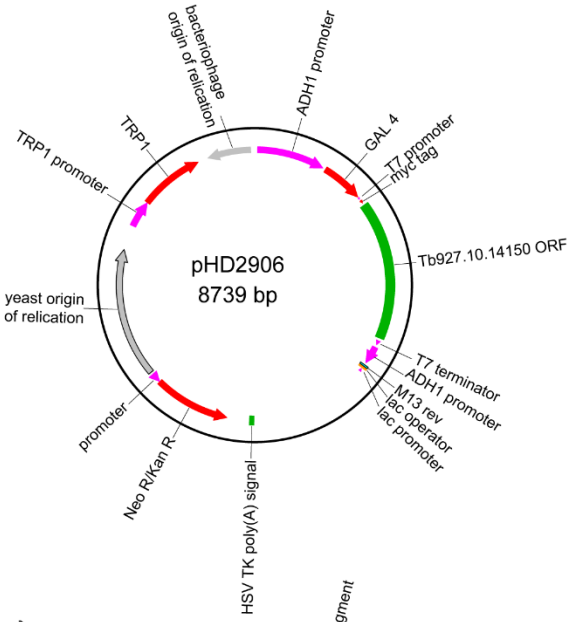
pHD number	plasmid	selection marker
2211	N-terminal V5-Caf1; endogenous	Blasticidin
2268	C-terminal Caf1-myc; overexpression	Hygromycin
2801	p2T7 Tb927.10.14150 RNAi	Hygromycin
2876	pMOtag33M + Tb927.10.14150; C-terminal myc-tag	G418
2877	for SKO of Tb927.10.14150	Blasticidin
2878	TAP-Tb927.10.14150	Puromycin
2879	tethering, lambda myc-Tb927.10.14150	Hygromycin
2899	pDonr + Tb927.10.14150; Gateway donor vector	
2900	for SKO of Tb927.10.14150	Puromycin
2902	stem loop RNAi Tb927.10.14150	Hygromycin
2906	Y2H; pGBKT7 + Tb927.10.14150	Kanamycin
2907	C-terminal myc-tagged Tb927.101.4150; overexpression	Hygromycin
2928	eYFP-ZC3H5	Puromycin
2932	TAP-ZC3H5	Blasticidin
2941	pRPa+6myc+ZC3H5	Hygromycin
2942	pHD617GW+lambda myc+Tb927.11.4900	Hygromycin
2943	pHD2907+Tb927.10.14150	Hygromycin
2945	pHD2944+ZC3H5	Hygromycin
2946	pHD2944+Tb927.7.3040	Hygromycin
2948	eYFP + Tb927.10.14150	Blasticidin

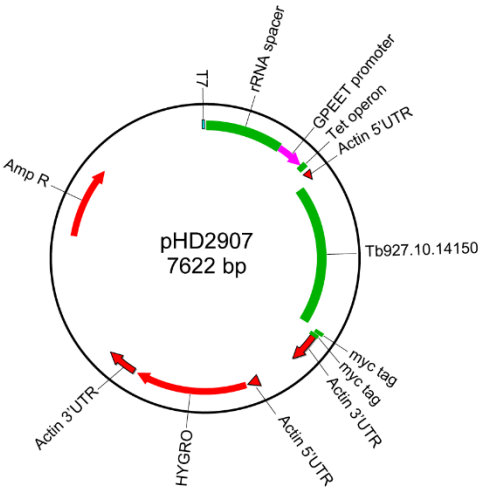
2949	Y2H; pGBKCg+Tb927.10.14150; C-terminal	
2964	pHD1991+ZC3H5 RNAi	G418
2965	pBD + Tb927.8.1500	
2969	pBD+TB927.7.3040	
2971	pAD-GW-Tb927.8.1500	
2972	pAD-GW-Tb927.7.3040	
2991	pBD-GW-Tb927.11.4900	
2992	pAD-GW-Tb927.11.4900	
2993	pHD617-lambda-myc-Tb927.8.1500	Hygromycin
2996	V5-Tb927.11.4900	Blasticidin
2998	pGL2084+G418	G418
2999	pUC19+Bla-V5+Tb927.7.3040	Blasticidin
3020	pUC19+Bla-V5+Tb927.8.1500	Blasticidin
3024	ZC3H5 RNAi	Hygromycin
3025	Tb927.11.4900 RNAi	Hygromycin
3026	Tb927.8.1500 RNAi	Hygromycin
3035	TAP-ZC3H5	Puromycin
3036	Tb927.8.1500 RNAi	G418
3050	TAP-ZC3H5	G418
p2370	pTET-7Bx; for *in vitro* transcription of beta-Globin; from AG Stoecklin	
p2829	eYFP + Dhh1; from Susanne Kramer	Blasticidin
p2845	mCherry- Dhh1; from Susanne Kramer	Blasticidin
p3295	PABP2+eYFP; from Susanne Kramer	G418
3137	pAD-GW-ZC3H5	
	pAD + T control	
2248	pBD-Caf1	
2251	pAD-Caf1	
2342	pHD1743+Caf1	
3138	pBD-GW-ZC3H5	
	pBD-Lamin (Y2H control)	
	pGBK p53	
3139	pHD1146 + ZC3H5	Hygromycin
	Tb927.7.3040 RNAi	Hygromycin
	pBS-BLA-V5	Blasticidin
3140	BP14150 pDonor	

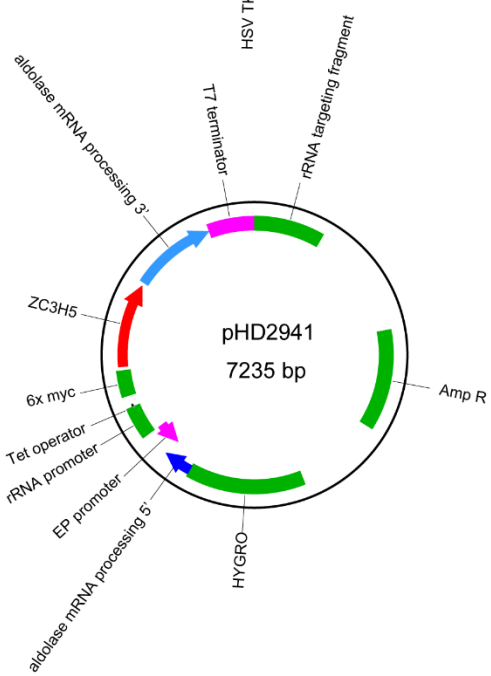
4.27. Plasmid maps

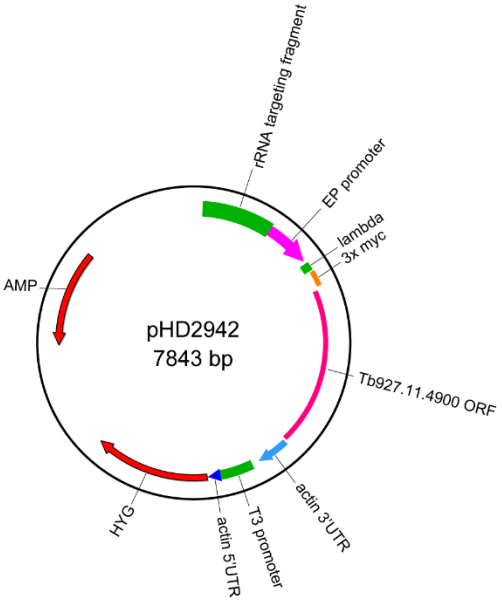


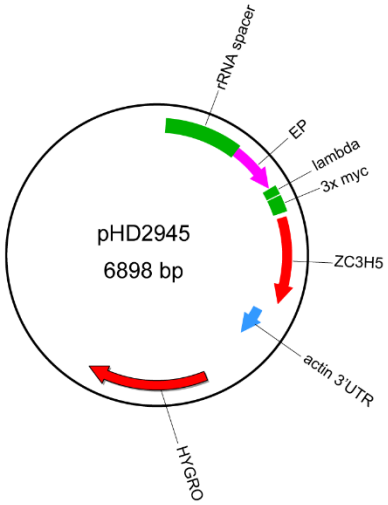

pHD2902
7668 bp
rRNA spacer
GPEET promoter
Tet operon
part of Tb927.10.14150 ORF (fwd)
part of Tb927.10.14150 ORF (rev)
EP IGR
VSG promoter
Actin 5'UTR
HYGRO
Actin 3'UTR
Amp R

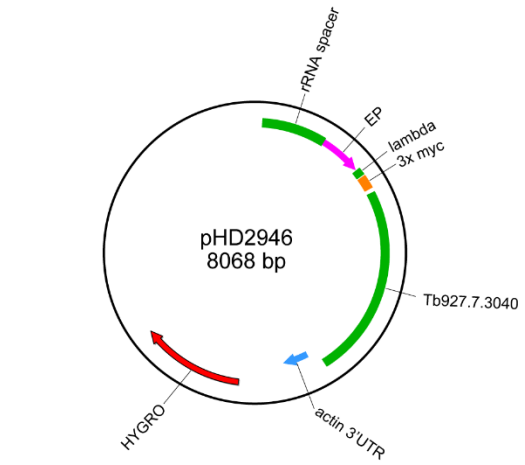

pHD2906
8739 bp
bacteriophage origin of relication
ADH1 promoter
GAL 4
T7 promoter
myc tag
Tb927.10.14150 ORF
T7 terminator
ADH1 promoter
M13 rev
lac operator
lac promoter
HSV TK poly(A) signal
Neo R/Kan R
promoter
yeast origin of relication
TRP1
TRP1 promoter

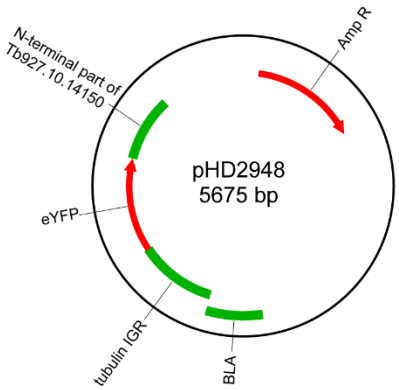

pHD2907
7622 bp
T7
rRNA spacer
GPEET promoter
Tet operon
Actin 5'UTR
Tb927.10.14150
myc tag
myc tag
Actin 3'UTR
Actin 5'UTR
HYGRO
Actin 3'UTR
Amp R

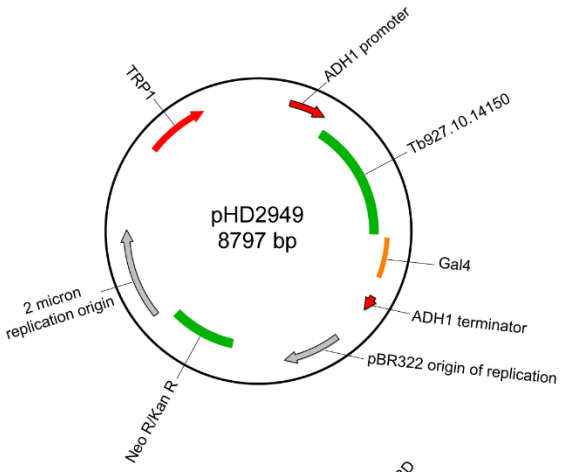

pHD2941
7235 bp
aldolase mRNA processing 3'
T7 terminator
rRNA targeting fragment
Amp R
HYGRO
aldolase mRNA processing 5'
EP promoter
rRNA promoter
Tet operator
6x myc
ZC3H5

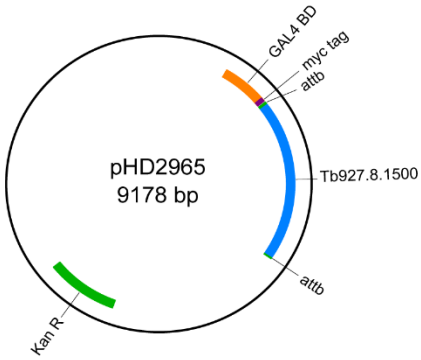

pHD2942
7843 bp
rRNA targeting fragment
EP promoter
lambda
3x myc
Tb927.11.4900 ORF
actin 3'UTR
T3 promoter
actin 5'UTR
HYG
AMP

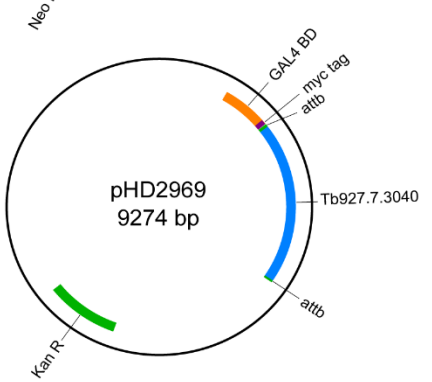

pHD2945
6898 bp
rRNA spacer
EP
lambda
3x myc
ZC3H5
actin 3'UTR
HYGRO

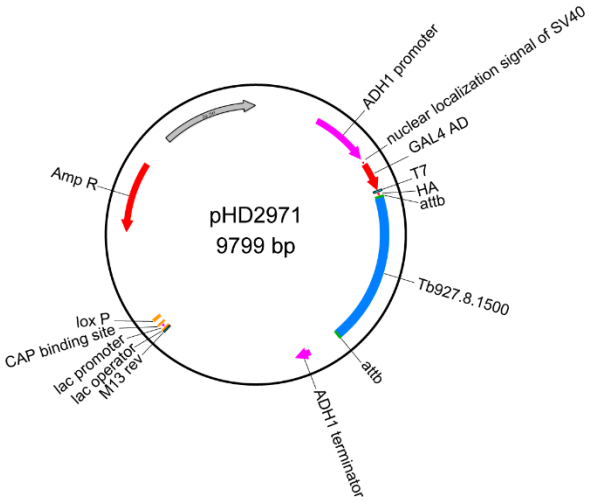

pHD2946
8068 bp
rRNA spacer
EP
lambda
3x myc
Tb927.7.3040
actin 3'UTR
HYGRO

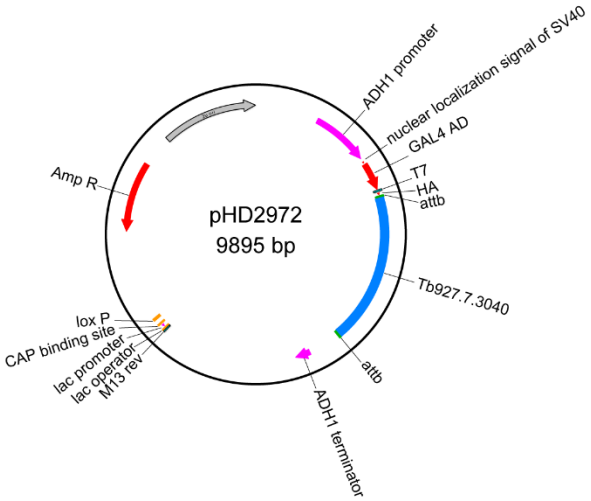

pHD2948
5675 bp
Amp R
N-terminal part of
Tb927.10.14150
eYFP
tubulin IGR
BLA

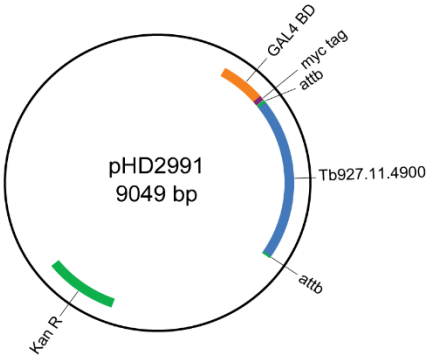

pHD2949
8797 bp
TRP1
ADH1 promoter
Tb927.10.14150
Gal4
ADH1 terminator
pBR322 origin of replication
Neo R/Kan R
2 micron
replication origin


pHD2965
9178 bp
GAL4 BD
myc tag
attb
Tb927.8.1500
attb
Kan R


pHD2969
9274 bp
GAL4 BD
myc tag
attb
Tb927.7.3040
attb
Kan R

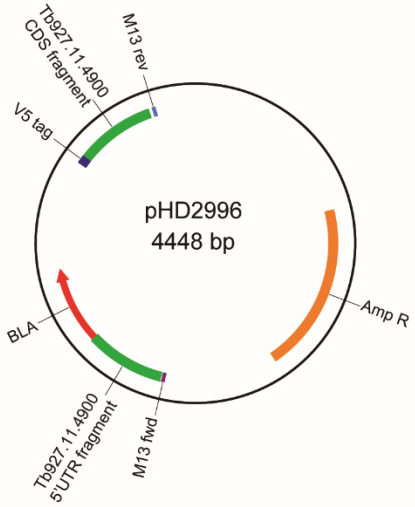

pHD2971
9799 bp
ADH1 promoter
nuclear localization signal of SV40
GAL4 AD
T7
HA
attb
Amp R
Tb927.8.1500
lox P
CAP binding site
lac promoter
lac operator
M13 rev
attb
ADH1 terminator

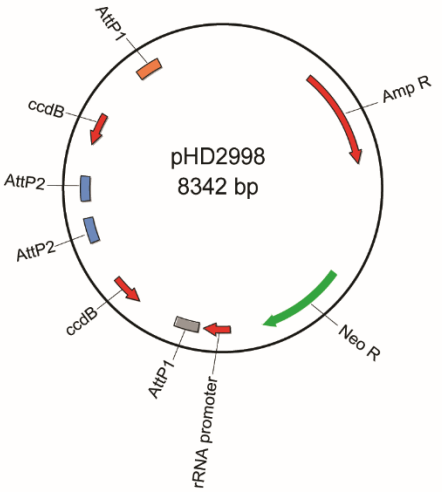

pHD2972
9895 bp
ADH1 promoter
nuclear localization signal of SV40
GAL4 AD
T7
HA
attb
Amp R
Tb927.7.3040
lox P
CAP binding site
lac promoter
lac operator
M13 rev
attb
ADH1 terminator

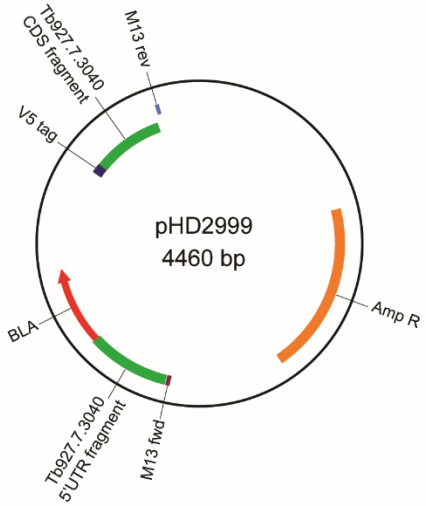

pHD2991
9049 bp
GAL4 BD
myc tag
attb
Tb927.11.4900
attb
Kan R

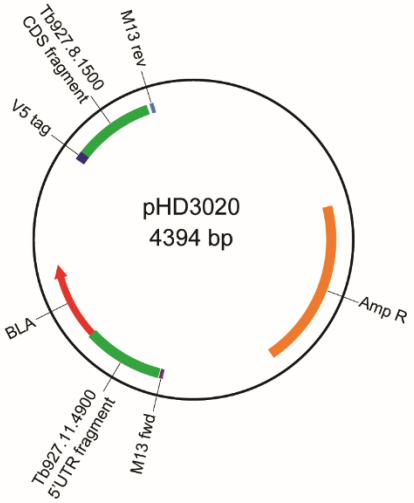

pHD2992
9670 bp
ADH1 promoter
nuclear localization signal of SV40
GAL4 AD
T7
HA
attb
Tb927.11.4900
attb
ADH1 terminator
lox P
CAP binding site
lac promoter
lac operator
M13 rev
Amp R

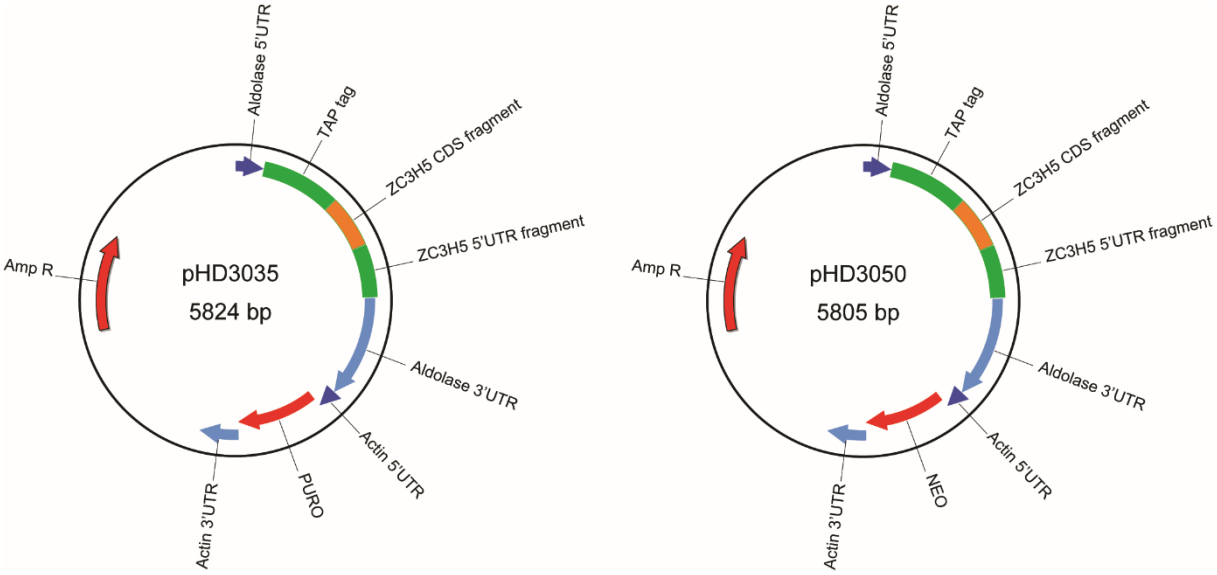

pHD2993
7879 bp
rRNA spacer
lambda
3x myc tag
attB
Tb927.8.1500
attB
actin 3'UTR
T3 promoter
actin 5'UTR
HYGRO
Amp R


pHD2996
4448 bp
Tb927.11.4900
CDS fragment
M13 rev
V5 tag
Amp R
BLA
Tb927.11.4900
5'UTR fragment
M13 fwd


pHD2998
8342 bp
AttP1
Amp R
ccdB
AttP2
AttP2
ccdB
Neo R
AttP1
rRNA promoter


pHD2999
4460 bp
Tb927.7.3040
CDS fragment
M13 rev
V5 tag
Amp R
BLA
Tb927.7.3040
5'UTR fragment
M13 fwd


pHD3020
4394 bp
Tb927.8.1500
CDS fragment
M13 rev
V5 tag
Amp R
BLA
Tb927.11.4900
5'UTR fragment
M13 fwd



4.28. Web resources

TriTrypDB	http://tritrypdb.org
GeneDB	http://www.genedb.org
BLAST NCBI	http://blast.ncbi.nlm.nih.gov/Blast.cgi
TrypTag	http://tryptag.org/?pageType=landing
MEME Motif Discovery	http://meme-suite.org/tools/dreme

References

Adl, S.M., Bass, D., Lane, C.E., Lukes, J., Schoch, C.L., Smirnov, A., Agatha, S., Berney, C., Brown, M.W., Burki, F.*, et al.* (2019). Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes. J Eukaryot Microbiol *66*, 4-119.

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

Anand, B., Verma, S.K., and Prakash, B. (2006). Structural stabilization of GTP-binding domains in circularly permuted GTPases: implications for RNA binding. Nucleic Acids Res *34*, 2196-2205.

Angelopoulos, E. (1970). Pellicular microtubules in the family Trypanosomatidae. The Journal of protozoology *17*, 39-51.

Antwi, E.B., Haanstra, J.R., Ramasamy, G., Jensen, B., Droll, D., Rojas, F., Minia, I., Terrao, M., Merce, C., Matthews, K.*, et al.* (2016). Integrative analysis of the Trypanosoma brucei gene expression cascade predicts differential regulation of mRNA processing and unusual control of ribosomal protein expression. BMC Genomics *17*, 306.

Archer, S.K., Inchaustegui, D., Queiroz, R., and Clayton, C. (2011). The cell cycle regulated transcriptome of Trypanosoma brucei. PloS one *6*, e18425.

Archer, S.K., Luu, V.D., de Queiroz, R.A., Brems, S., and Clayton, C. (2009). Trypanosoma brucei PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle. PLoS pathogens *5*, e1000565.

Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics (Oxford, England) *27*, 1653-1659.

Benz, C., Mulindwa, J., Ouna, B., and Clayton, C. (2011). The Trypanosoma brucei zinc finger protein ZC3H18 is involved in differentiation. Mol Biochem Parasitol *177*, 148-151.

Blattner, J., Helfert, S., Michels, P., and Clayton, C. (1998). Compartmentation of phosphoglycerate kinase in Trypanosoma brucei plays a critical role in parasite energy metabolism. Proceedings of the National Academy of Sciences of the United States of America *95*, 11596-11600.

Brecht, M., and Parsons, M. (1998). Changes in polysome profiles accompany trypanosome development. Mol Biochem Parasitol *97*, 189-198.

Bringaud, F., Riviere, L., and Coustou, V. (2006). Energy metabolism of trypanosomatids: adaptation to available carbon sources. Mol Biochem Parasitol *149*, 1-9.

Broadhead, R., Dawe, H.R., Farr, H., Griffiths, S., Hart, S.R., Portman, N., Shaw, M.K., Ginger, M.L., Gaskell, S.J., McKean, P.G.*, et al.* (2006). Flagellar motility is required for the viability of the bloodstream trypanosome. Nature *440*, 224-227.

Buhlmann, M., Walrad, P., Rico, E., Ivens, A., Capewell, P., Naguleswaran, A., Roditi, I., and Matthews, K.R. (2015). NMD3 regulates both mRNA and rRNA nuclear export in African trypanosomes via an XPOI-linked pathway. Nucleic Acids Res *43*, 4491-4504.

Butter, F., Bucerius, F., Michel, M., Cicova, Z., Mann, M., and Janzen, C.J. (2013). Comparative proteomics of two life cycle stages of stable isotope-labeled Trypanosoma

brucei reveals novel components of the parasite's host adaptation machinery. Mol Cell Proteomics *12*, 172-179.

Caro, F., Bercovich, N., Atorrasagasti, C., Levin, M.J., and Vazquez, M.P. (2006). Trypanosoma cruzi: analysis of the complete PUF RNA-binding protein family. Experimental parasitology *113*, 112-124.

Cassola, A. (2011). RNA Granules Living a Post-transcriptional Life: the Trypanosomes' Case. Current chemical biology *5*, 108-117.

Cassola, A., De Gaudenzi, J.G., and Frasch, A.C. (2007). Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. Molecular microbiology *65*, 655-670.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2013). System-wide identification of RNA-binding proteins by interactome capture. Nat Protoc *8*, 491-500.

Clayton, C., and Estevez, A. (2010). The exosomes of trypanosomes and other protists. Adv Exp Med Biol *702*, 39-49.

Clayton, C., Hausler, T., and Blattner, J. (1995). Protein trafficking in kinetoplastid protozoa. Microbiological reviews *59*, 325-344.

Clayton, C., Schwede, A., Stewart, M., Robles, A., Benz, C., Po, J., Wurst, M., Queiroz, R., and Archer, S. (2008). Control of mRNA degradation in trypanosomes. Biochemical Society transactions *36*, 520-521.

Clayton, C., and Shapira, M. (2007). Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. Mol Biochem Parasitol *156*, 93-101.

Clayton, C.E. (2014). Networks of gene expression regulation in Trypanosoma brucei. Mol Biochem Parasitol *195*, 96-106.

Clery, A., Blatter, M., and Allain, F.H. (2008). RNA recognition motifs: boring? Not quite. Current opinion in structural biology *18*, 290-298.

Das, A., Morales, R., Banday, M., Garcia, S., Hao, L., Cross, G.A., Estevez, A.M., and Bellofatto, V. (2012). The essential polysome-associated RNA-binding protein RBP42 targets mRNAs involved in Trypanosoma brucei energy metabolism. RNA (New York, NY) *18*, 1968-1983.

de Freitas Nascimento, J., Kelly, S., Sunter, J., and Carrington, M. (2018). Codon choice directs constitutive mRNA levels in trypanosomes. eLife *7*.

De Gaudenzi, J., Frasch, A.C., and Clayton, C. (2005). RNA-binding domain proteins in Kinetoplastids: a comparative analysis. Eukaryot Cell *4*, 2106-2114.

Dean, S., Sunter, J.D., and Wheeler, R.J. (2017). TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. Trends Parasitol *33*, 80-82.

Dejung, M., Subota, I., Bucerius, F., Dindar, G., Freiwald, A., Engstler, M., Boshart, M., Butter, F., and Janzen, C.J. (2016). Quantitative Proteomics Uncovers Novel Factors Involved in Developmental Differentiation of Trypanosoma brucei. PLoS pathogens *12*, e1005439.

Dostalova, A., Kaser, S., Cristodero, M., and Schimanski, B. (2013). The nuclear mRNA export receptor Mex67-Mtr2 of Trypanosoma brucei contains a unique and essential zinc finger motif. Molecular microbiology *88*, 728-739.

Droll, D., Archer, S., Fenn, K., Delhi, P., Matthews, K., and Clayton, C. (2010). The trypanosome Pumilio-domain protein PUF7 associates with a nuclear cyclophilin and is involved in ribosomal RNA maturation. FEBS letters *584*, 1156-1162.

Droll, D., Minia, I., Fadda, A., Singh, A., Stewart, M., Queiroz, R., and Clayton, C. (2013). Post-transcriptional regulation of the trypanosome heat shock response by a zinc finger protein. PLoS pathogens *9*, e1003286.

Erben, E., Chakraborty, C., and Clayton, C. (2014a). The CAF1-NOT complex of trypanosomes. Front Genet *4*, 299.

Erben, E.D., Fadda, A., Lueong, S., Hoheisel, J.D., and Clayton, C. (2014b). A genome-wide tethering screen reveals novel potential post-transcriptional regulators in Trypanosoma brucei. PLoS pathogens *10*, e1004178.

Fadda, A., Farber, V., Droll, D., and Clayton, C. (2013). The roles of 3'-exoribonucleases and the exosome in trypanosome mRNA degradation. RNA (New York, NY) *19*, 937-947.

Fadda, A., Ryten, M., Droll, D., Rojas, F., Farber, V., Haanstra, J.R., Merce, C., Bakker, B.M., Matthews, K., and Clayton, C. (2014). Transcriptome-wide analysis of trypanosome mRNA decay reveals complex degradation kinetics and suggests a role for co-transcriptional degradation in determining mRNA levels. Molecular microbiology *94*, 307-326.

Farber, V., Erben, E., Sharma, S., Stoecklin, G., and Clayton, C. (2013). Trypanosome CNOT10 is essential for the integrity of the NOT deadenylase complex and for degradation of many mRNAs. Nucleic Acids Res *41*, 1211-1222.

Field, M.C., Horn, D., Fairlamb, A.H., Ferguson, M.A., Gray, D.W., Read, K.D., De Rycker, M., Torrie, L.S., Wyatt, P.G., Wyllie, S.*, et al.* (2017). Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. Nature reviews Microbiology *15*, 217-231.

Franco, J.R., Cecchi, G., Priotto, G., Paone, M., Diarra, A., Grout, L., Mattioli, R.C., and Argaw, D. (2017). Monitoring the elimination of human African trypanosomiasis: Update to 2014. PLoS neglected tropical diseases *11*, e0005585.

Freire, E.R., Sturm, N.R., Campbell, D.A., and de Melo Neto, O.P. (2017). The Role of Cytoplasmic mRNA Cap-Binding Protein Complexes in Trypanosoma brucei and Other Trypanosomatids. Pathogens (Basel, Switzerland) *6*.

Fremont, S., and Echard, A. (2018). Membrane Traffic in the Late Steps of Cytokinesis. Current biology : CB *28*, R458-r470.

Friend, K., Campbell, Z.T., Cooke, A., Kroll-Conner, P., Wickens, M.P., and Kimble, J. (2012). A conserved PUF-Ago-eEF1A complex attenuates translation elongation. Nat Struct Mol Biol *19*, 176-183.

Fritz, M., Vanselow, J., Sauer, N., Lamer, S., Goos, C., Siegel, T.N., Subota, I., Schlosser, A., Carrington, M., and Kramer, S. (2015). Novel insights into RNP granules by employing the trypanosome's microtubule skeleton as a molecular sieve. Nucleic Acids Res *43*, 8013-8032.

Fu, C., Wehr, D.R., Edwards, J., and Hauge, B. (2008). Rapid one-step recombinational cloning. Nucleic Acids Res *36*, e54.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nature reviews Genetics *15*, 829-845.

Greber, B.J. (2016). Mechanistic insight into eukaryotic 60S ribosomal subunit biogenesis by cryo-electron microscopy. RNA (New York, NY) *22*, 1643-1662.

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

Hammarton, T.C., Clark, J., Douglas, F., Boshart, M., and Mottram, J.C. (2003). Stage-specific differences in cell cycle control in Trypanosoma brucei revealed by RNA interference of a mitotic cyclin. The Journal of biological chemistry *278*, 22877-22886.

Hammarton, T.C., Lillico, S.G., Welburn, S.C., and Mottram, J.C. (2005). Trypanosoma brucei MOB1 is required for accurate and efficient cytokinesis but not for exit from mitosis. Molecular microbiology *56*, 104-116.

Hammarton, T.C., Monnerat, S., and Mottram, J.C. (2007). Cytokinesis in trypanosomatids. Current opinion in microbiology *10*, 520-527.

Hanson, G., and Coller, J. (2018). Codon optimality, bias and usage in translation and mRNA decay. Nature reviews Molecular cell biology *19*, 20-30.

Hartmann, C., Benz, C., Brems, S., Ellis, L., Luu, V.D., Stewart, M., D'Orso, I., Busold, C., Fellenberg, K., Frasch, A.C.*, et al.* (2007). Small trypanosome RNA-binding proteins TbUBP1 and TbUBP2 influence expression of F-box protein mRNAs in bloodstream trypanosomes. Eukaryot Cell *6*, 1964-1978.

Helser, T.L., Baan, R.A., and Dahlberg, A.E. (1981). Characterization of a 40S ribosomal subunit complex in polyribosomes of Saccharomyces cerevisiae treated with cycloheximide. Molecular and cellular biology *1*, 51-57.

Hendriks, E.F., and Matthews, K.R. (2005). Disruption of the developmental programme of Trypanosoma brucei by genetic ablation of TbZFP1, a differentiation-enriched CCCH protein. Molecular microbiology *57*, 706-716.

Hendriks, E.F., Robinson, D.R., Hinkins, M., and Matthews, K.R. (2001). A novel CCCH protein which modulates differentiation of Trypanosoma brucei to its procyclic form. The EMBO journal *20*, 6700-6711.

Hernandez, R., and Cevallos, A.M. (2014). Ribosomal RNA gene transcription in trypanosomes. Parasitology research *113*, 2415-2424.

Heyer, E.E., and Moore, M.J. (2016). Redefining the Translational Status of 80S Monosomes. Cell *164*, 757-769.

Hoek, M., Zanders, T., and Cross, G.A. (2002). Trypanosoma brucei expression-site-associated-gene-8 protein interacts with a Pumilio family protein. Mol Biochem Parasitol *120*, 269-283.

Hogan, D.J., Riordan, D.P., Gerber, A.P., Herschlag, D., and Brown, P.O. (2008). Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS biology *6*, e255.

Hovel-Miner, G., Mugnier, M., Papavasiliou, F.N., Pinger, J., and Schulz, D. (2015). A Host-Pathogen Interaction Reduced to First Principles: Antigenic Variation in T. brucei. Results and problems in cell differentiation *57*, 23-46.

Hunziker, W., Whitney, J.A., and Mellman, I. (1992). Brefeldin A and the endocytic pathway. Possible implications for membrane traffic and sorting. FEBS letters *307*, 93-96.

Ivanov, P., Kedersha, N., and Anderson, P. (2011). Stress puts TIA on TOP. Genes & development *25*, 2119-2124.

Jackson, C.L., and Kepes, F. (1994). BFR1, a multicopy suppressor of brefeldin A-induced lethality, is implicated in secretion and nuclear segregation in Saccharomyces cerevisiae. Genetics *137*, 423-437.

Jagannathan, S., Reid, D.W., Cox, A.H., and Nicchitta, C.V. (2014). De novo translation initiation on membrane-bound ribosomes as a mechanism for localization of cytosolic protein mRNAs to the endoplasmic reticulum. RNA (New York, NY) *20*, 1489-1498.

Jeacock, L., Faria, J., and Horn, D. (2018). Codon usage bias controls mRNA and protein abundance in trypanosomatids. eLife *7*.

Jensen, B.C., Brekken, D.L., Randall, A.C., Kifer, C.T., and Parsons, M. (2005). Species specificity in ribosome biogenesis: a nonconserved phosphoprotein is required for formation of the large ribosomal subunit in Trypanosoma brucei. Eukaryot Cell *4*, 30-35.

Klausner, R.D., Donaldson, J.G., and Lippincott-Schwartz, J. (1992). Brefeldin A: insights into the control of membrane traffic and organelle structure. The Journal of cell biology *116*, 1071-1080.

Kohl, L., Robinson, D., and Bastin, P. (2003). Novel roles for the flagellum in cell morphogenesis and cytokinesis of trypanosomes. The EMBO journal *22*, 5336-5346.

Kolev, N.G., Ramey-Butler, K., Cross, G.A., Ullu, E., and Tschudi, C. (2012). Developmental progression to infectivity in Trypanosoma brucei triggered by an RNA-binding protein. Science (New York, NY) *338*, 1352-1353.

Kolev, N.G., Ullu, E., and Tschudi, C. (2014). The emerging role of RNA-binding proteins in the life cycle of Trypanosoma brucei. Cell Microbiol *16*, 482-489.

Kotani, T., Yasuda, K., Ota, R., and Yamashita, M. (2013). Cyclin B1 mRNA translation is temporally controlled through formation and disassembly of RNA granules. The Journal of cell biology *202*, 1041-1055.

Kramer, S. (2014). RNA in development: how ribonucleoprotein granules regulate the life cycles of pathogenic protozoa. Wiley interdisciplinary reviews RNA *5*, 263-284.

Kramer, S. (2017). The ApaH-like phosphatase TbALPH1 is the major mRNA decapping enzyme of trypanosomes. PLoS pathogens *13*, e1006456.

Kramer, S., and Carrington, M. (2011). Trans-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. Trends Parasitol *27*, 23-30.

Kramer, S., Kimblin, N.C., and Carrington, M. (2010). Genome-wide in silico screen for CCCH-type zinc finger proteins of Trypanosoma brucei, Trypanosoma cruzi and Leishmania major. BMC Genomics *11*, 283.

Kruger, T., Hofweber, M., and Kramer, S. (2013). SCD6 induces ribonucleoprotein granule formation in trypanosomes in a translation-independent manner, regulated by its Lsm and RGG domains. Mol Biol Cell *24*, 2098-2111.

Kumar, P., and Wang, C.C. (2006). Dissociation of cytokinesis initiation from mitotic control in a eukaryote. Eukaryot Cell *5*, 92-102.

Kurasawa, Y., Hu, H., Zhou, Q., and Li, Z. (2018). A trypanosome-specific protein cooperates with the CIF1 protein to promote cytokinesis in Trypanosoma brucei. The Journal of biological chemistry.

Lacomble, S., Vaughan, S., Gadelha, C., Morphew, M.K., Shaw, M.K., McIntosh, J.R., and Gull, K. (2009). Three-dimensional cellular architecture of the flagellar pocket and associated cytoskeleton in trypanosomes revealed by electron microscope tomography. Journal of cell science *122*, 1081-1090.

Lang, B.D., and Fridovich-Keil, J.L. (2000). Scp160p, a multiple KH-domain protein, is a component of mRNP complexes in yeast. Nucleic Acids Res *28*, 1576-1584.

Lang, B.D., Li, A., Black-Brewster, H.D., and Fridovich-Keil, J.L. (2001). The brefeldin A resistance protein Bfr1p is a component of polyribosome-associated mRNP complexes in yeast. Nucleic Acids Res *29*, 2567-2574.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature methods *9*, 357-359.

Langousis, G., and Hill, K.L. (2014). Motility and more: the flagellum of Trypanosoma brucei. Nature reviews Microbiology *12*, 505-518.

Lapointe, C.P., Wilinski, D., Saunders, H.A., and Wickens, M. (2015). Protein-RNA networks revealed through covalent RNA marks. Nature methods *12*, 1163-1170.

Leiss, K., and Clayton, C. (2016). DESeqUI-Trypanosome RNAseq analysis made easy. Zenodo.

Leiss, K., Merce, C., Muchunga, E., and Clayton, C. (2016). TrypRNAseq-A easy to use pipeline for Trypanosoma RNAseq data. Zenodo.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England) *25*, 2078-2079.

Li, H., and Tschudi, C. (2005). Novel and essential subunits in the 300-kilodalton nuclear cap binding complex of Trypanosoma brucei. Molecular and cellular biology *25*, 2216-2226.

Ling, A.S., Trotter, J.R., and Hendriks, E.F. (2011). A zinc finger protein, TbZC3H20, stabilizes two developmentally regulated mRNAs in trypanosomes. The Journal of biological chemistry *286*, 20152-20162.

Lippincott-Schwartz, J. (1993). Bidirectional membrane traffic between the endoplasmic reticulum and Golgi apparatus. Trends in cell biology *3*, 81-88.

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

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology *15*, 550.

Lueong, S., Merce, C., Fischer, B., Hoheisel, J.D., and Erben, E.D. (2016). Gene expression regulatory networks in Trypanosoma brucei: Insights into the role of the mRNA-binding proteome. Molecular microbiology.

Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nature reviews Molecular cell biology *8*, 479-490.

Luu, V.D., Brems, S., Hoheisel, J.D., Burchmore, R., Guilbride, D.L., and Clayton, C. (2006). Functional analysis of Trypanosoma brucei PUF1. Mol Biochem Parasitol *150*, 340-349.

Maier, R., Brandner, C., Hintner, H., Bauer, J., and Onder, K. (2008). Construction of a reading frame-independent yeast two-hybrid vector system for site-specific recombinational cloning and protein interaction screening. BioTechniques *45*, 235-244.

Manful, T., Fadda, A., and Clayton, C. (2011). The role of the 5'-3' exoribonuclease XRNA in transcriptome-wide mRNA degradation. RNA (New York, NY) *17*, 2039-2047.

Mani, J., Guttinger, A., Schimanski, B., Heller, M., Acosta-Serrano, A., Pescher, P., Spath, G., and Roditi, I. (2011). Alba-domain proteins of Trypanosoma brucei are cytoplasmic RNA-binding proteins that interact with the translation machinery. PloS one *6*, e22463.

Martin, M. (2011). Next Generation Sequencing Data Analysis: Cutadapt removes dapter sequences from high-throughput sequencing reads. EMBnetjournl.

Matthews, K.R. (1999). Developments in the differentiation of Trypanosoma brucei. Parasitology today (Personal ed) *15*, 76-80.

Matthews, K.R. (2005). The developmental cell biology of Trypanosoma brucei. Journal of cell science *118*, 283-290.

McCahill, A., Warwicker, J., Bolger, G.B., Houslay, M.D., and Yarwood, S.J. (2002). The RACK1 scaffold protein: a dynamic cog in cell response mechanisms. Molecular pharmacology *62*, 1261-1273.

McKean, P.G. (2003). Coordination of cell cycle and cytokinesis in Trypanosoma brucei. Current opinion in microbiology *6*, 600-607.

McKean, P.G., Baines, A., Vaughan, S., and Gull, K. (2003). Gamma-tubulin functions in the nucleation of a discrete subset of microtubules in the eukaryotic flagellum. Current biology : CB *13*, 598-602.

Meyuhas, O., and Kahan, T. (2015). The race to decipher the top secrets of TOP mRNAs. Biochim Biophys Acta *1849*, 801-811.

Michaeli, S. (2011). Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome. Future microbiology *6*, 459-474.

Michels, P.A., Bringaud, F., Herman, M., and Hannaert, V. (2006). Metabolic functions of glycosomes in trypanosomatids. Biochim Biophys Acta *1763*, 1463-1477.

Minia, I., Merce, C., Terrao, M., and Clayton, C. (2016). Translation Regulation and RNA Granule Formation after Heat Shock of Procyclic Form Trypanosoma brucei: Many Heat-Induced mRNAs Are also Increased during Differentiation to Mammalian-Infective Forms. PLoS neglected tropical diseases *10*, e0004982.

Mugo, E., and Clayton, C. (2017). Expression of the RNA-binding protein RBP10 promotes the bloodstream-form differentiation state in Trypanosoma brucei. PLoS pathogens *13*, e1006560.

Mugo, E., Egler, F., and Clayton, C. (2017). Conversion of procyclic-form Trypanosoma brucei to the bloodstream form by transient expression of RBP10. Mol Biochem Parasitol *216*, 49-51.

Noll, M., Hapke, B., Schreier, M.H., and Noll, H. (1973). Structural dynamics of bacterial ribosomes. I. Characterization of vacant couples and their relation to complexed ribosomes. Journal of molecular biology *75*, 281-294.

Ogbadoyi, E., Ersfeld, K., Robinson, D., Sherwin, T., and Gull, K. (2000). Architecture of the Trypanosoma brucei nucleus during interphase and mitosis. Chromosoma *108*, 501-513.

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

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

Ouna, B.A., Stewart, M., Helbig, C., and Clayton, C. (2012). The Trypanosoma brucei CCCH zinc finger proteins ZC3H12 and ZC3H13. Mol Biochem Parasitol *183*, 184-188.

Overath, P., and Engstler, M. (2004). Endocytosis, membrane recycling and sorting of GPI-anchored proteins: Trypanosoma brucei as a model system. Molecular microbiology *53*, 735-744.

Panigrahi, A.K., Ernst, N.L., Domingo, G.J., Fleck, M., Salavati, R., and Stuart, K.D. (2006). Compositionally and functionally distinct editosomes in Trypanosoma brucei. RNA (New York, NY) *12*, 1038-1049.

Pelham, H.R. (1991). Multiple targets for brefeldin A. Cell *67*, 449-451.

Ralston, K.S., Lerner, A.G., Diener, D.R., and Hill, K.L. (2006). Flagellar motility contributes to cytokinesis in Trypanosoma brucei and is modulated by an evolutionarily conserved dynein regulatory system. Eukaryot Cell *5*, 696-711.

Reid, D.W., and Nicchitta, C.V. (2012). Primary role for endoplasmic reticulum-bound ribosomes in cellular translation identified by ribosome profiling. The Journal of biological chemistry *287*, 5518-5527.

Reid, D.W., and Nicchitta, C.V. (2015). Diversity and selectivity in mRNA translation on the endoplasmic reticulum. Nature reviews Molecular cell biology *16*, 221-231.

Robinson, D.R., and Gull, K. (1991). Basal body movements as a mechanism for mitochondrial genome segregation in the trypanosome cell cycle. Nature *352*, 731-733.

Robinson, D.R., Sherwin, T., Ploubidou, A., Byard, E.H., and Gull, K. (1995). Microtubule polarity and dynamics in the control of organelle positioning, segregation, and cytokinesis in the trypanosome cell cycle. The Journal of cell biology *128*, 1163-1172.

Rotenberg, M.O., Moritz, M., and Woolford, J.L., Jr. (1988). Depletion of Saccharomyces cerevisiae ribosomal protein L16 causes a decrease in 60S ribosomal subunits and formation of half-mer polyribosomes. Genes & development *2*, 160-172.

Sakyiama, J., Zimmer, S.L., Ciganda, M., Williams, N., and Read, L.K. (2013). Ribosome biogenesis requires a highly diverged XRN family 5'->3' exoribonuclease for rRNA processing in Trypanosoma brucei. RNA (New York, NY) *19*, 1419-1431.

Schumann Burkard, G., Kaser, S., de Araujo, P.R., Schimanski, B., Naguleswaran, A., Knusel, S., Heller, M., and Roditi, I. (2013). Nucleolar proteins regulate stage-specific gene expression and ribosomal RNA maturation in Trypanosoma brucei. Molecular microbiology *88*, 827-840.

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

Schwede, A., Manful, T., Jha, B.A., Helbig, C., Bercovich, N., Stewart, M., and Clayton, C. (2009). The role of deadenylation in the degradation of unstable mRNAs in trypanosomes. Nucleic Acids Res *37*, 5511-5528.

Shapiro, T.A., and Englund, P.T. (1995). The structure and replication of kinetoplast DNA. Annual review of microbiology *49*, 117-143.

Sherwin, T., and Gull, K. (1989). Visualization of detyrosination along single microtubules reveals novel mechanisms of assembly during cytoskeletal duplication in trypanosomes. Cell *57*, 211-221.

Shi, H., Ramey-Butler, K., and Tschudi, C. (2018). A single-point mutation in the RNA-binding protein 6 generates Trypanosoma brucei metacyclics that are able to progress to bloodstream forms in vitro. Mol Biochem Parasitol.

Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S., and Cross, G.A. (2009). Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. Genes & development *23*, 1063-1076.

Siegel, T.N., Hekstra, D.R., Wang, X., Dewell, S., and Cross, G.A. (2010). Genome-wide analysis of mRNA abundance in two life-cycle stages of Trypanosoma brucei and identification of splicing and polyadenylation sites. Nucleic Acids Res *38*, 4946-4957.

Simpson, C.E., Lui, J., Kershaw, C.J., Sims, P.F., and Ashe, M.P. (2014). mRNA localization to P-bodies in yeast is bi-phasic with many mRNAs captured in a late Bfr1p-dependent wave. Journal of cell science *127*, 1254-1262.

Simpson, L. (1987). The mitochondrial genome of kinetoplastid protozoa: genomic organization, transcription, replication, and evolution. Annual review of microbiology *41*, 363-382.

Singh, A., Minia, I., Droll, D., Fadda, A., Clayton, C., and Erben, E. (2014). Trypanosome MKT1 and the RNA-binding protein ZC3H11: interactions and potential roles in post-transcriptional regulatory networks. Nucleic Acids Res *42*, 4652-4668.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731-745.

Subota, I., Rotureau, B., Blisnick, T., Ngwabyt, S., Durand-Dubief, M., Engstler, M., and Bastin, P. (2011). ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation. Mol Biol Cell *22*, 4205-4219.

Terrao, M., Marucha, K.K., Mugo, E., Droll, D., Minia, I., Egler, F., Braun, J., and Clayton, C. (2018). The suppressive cap-binding complex factor 4EIP is required for normal differentiation. Nucleic Acids Res.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature methods *13*, 731-740.

Uilenberg, G. (1998). A field guide for the Diagnosis, treatment and prevention of african animal trypanosomosis. Rome, FAO.

Unsworth, H., Raguz, S., Edwards, H.J., Higgins, C.F., and Yague, E. (2010). mRNA escape from stress granule sequestration is dictated by localization to the endoplasmic reticulum. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *24*, 3370-3380.

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

Vanhamme, L., and Pays, E. (1995). Control of gene expression in trypanosomes. Microbiological reviews *59*, 223-240.

Vasquez, J.J., Hon, C.C., Vanselow, J.T., Schlosser, A., and Siegel, T.N. (2014). Comparative ribosome profiling reveals extensive translational complexity in different Trypanosoma brucei life cycle stages. Nucleic Acids Res *42*, 3623-3637.

Wang, X., McLachlan, J., Zamore, P.D., and Hall, T.M. (2002). Modular recognition of RNA by a human pumilio-homology domain. Cell *110*, 501-512.

Wedel, C., Forstner, K.U., Derr, R., and Siegel, T.N. (2017). GT-rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes. The EMBO journal.

Weidner, J., Wang, C., Prescianotto-Baschong, C., Estrada, A.F., and Spang, A. (2014). The polysome-associated proteins Scp160 and Bfr1 prevent P body formation under normal growth conditions. Journal of cell science *127*, 1992-2004.

Wittinghofer, A., and Vetter, I.R. (2011). Structure-function relationships of the G domain, a canonical switch motif. Annual review of biochemistry *80*, 943-971.

Woodward, R., and Gull, K. (1990). Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of Trypanosoma brucei. Journal of cell science *95 (Pt 1)*, 49-57.

Wurst, M., Seliger, B., Jha, B.A., Klein, C., Queiroz, R., and Clayton, C. (2012). Expression of the RNA recognition motif protein RBP10 promotes a bloodstream-form transcript pattern in Trypanosoma brucei. Molecular microbiology *83*, 1048-1063.

Yoshihama, M., Uechi, T., Asakawa, S., Kawasaki, K., Kato, S., Higa, S., Maeda, N., Minoshima, S., Tanaka, T., Shimizu, N.*, et al.* (2002). The human ribosomal protein genes: sequencing and comparative analysis of 73 genes. Genome Res *12*, 379-390.

Zamudio, J.R., Mittra, B., Campbell, D.A., and Sturm, N.R. (2009). Hypermethylated cap 4 maximizes Trypanosoma brucei translation. Molecular microbiology *72*, 1100-1110.

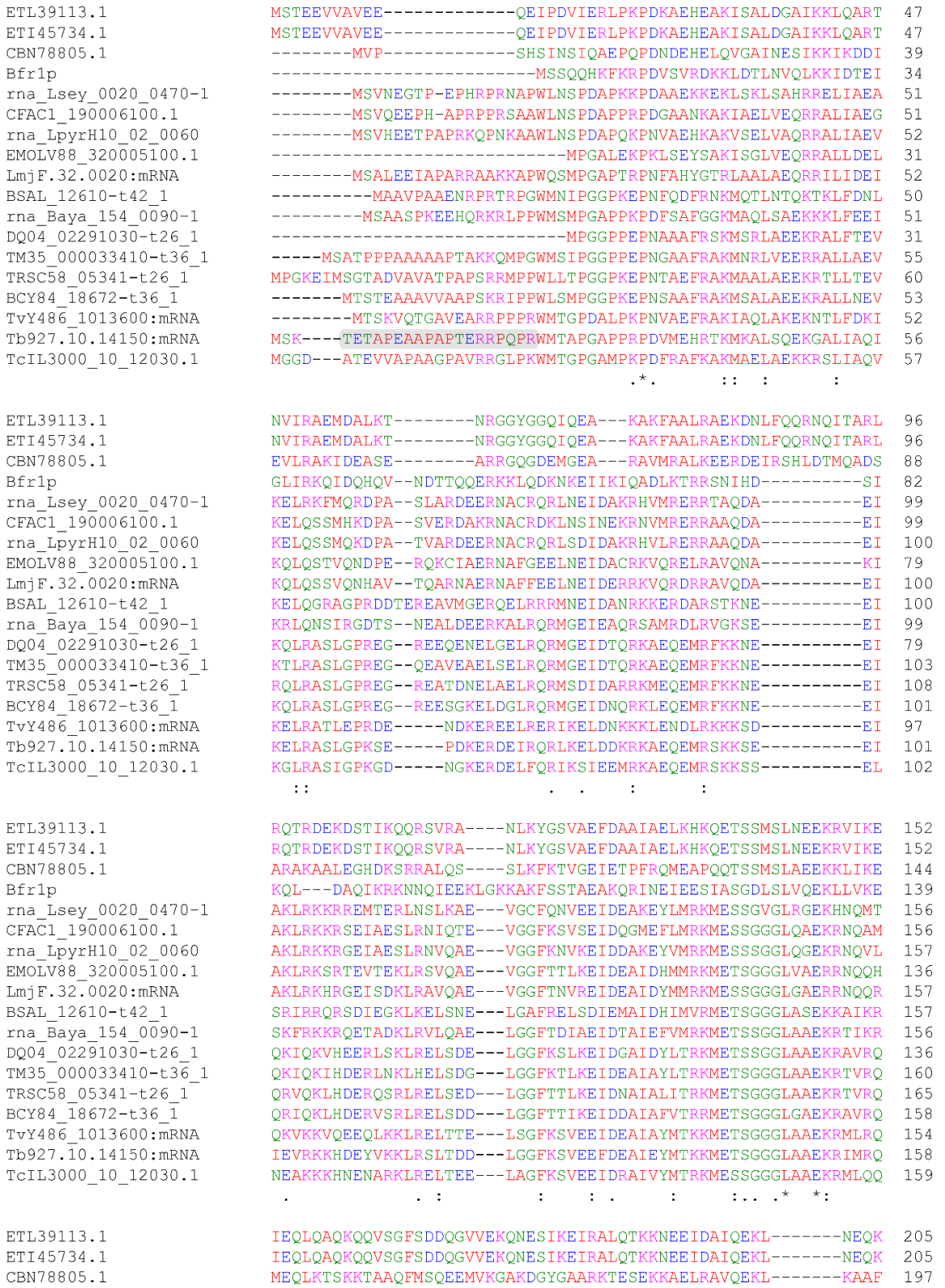
Zhou, Q., Gu, J., Lun, Z.R., Ayala, F.J., and Li, Z. (2016). Two distinct cytokinesis pathways drive trypanosome cell division initiation from opposite cell ends. Proceedings of the National Academy of Sciences of the United States of America *113*, 3287-3292.

Ziegelbauer, K., and Overath, P. (1990). Surface antigen change during differentiation of Trypanosoma brucei. Biochemical Society transactions *18*, 731-733.

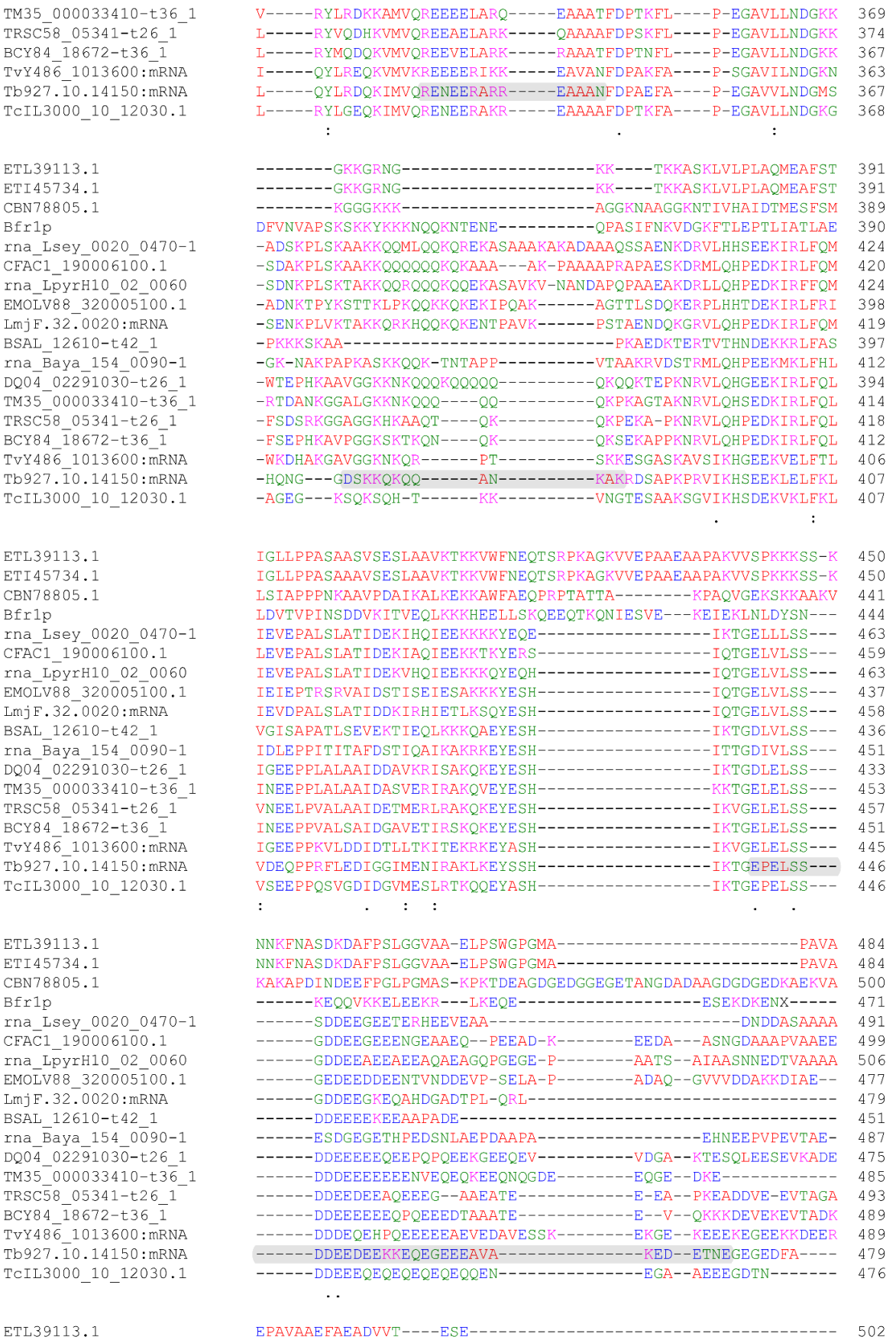
Zinoviev, A., Manor, S., and Shapira, M. (2012). Nutritional stress affects an atypical cap-binding protein in Leishmania. RNA Biol *9*, 1450-1460.

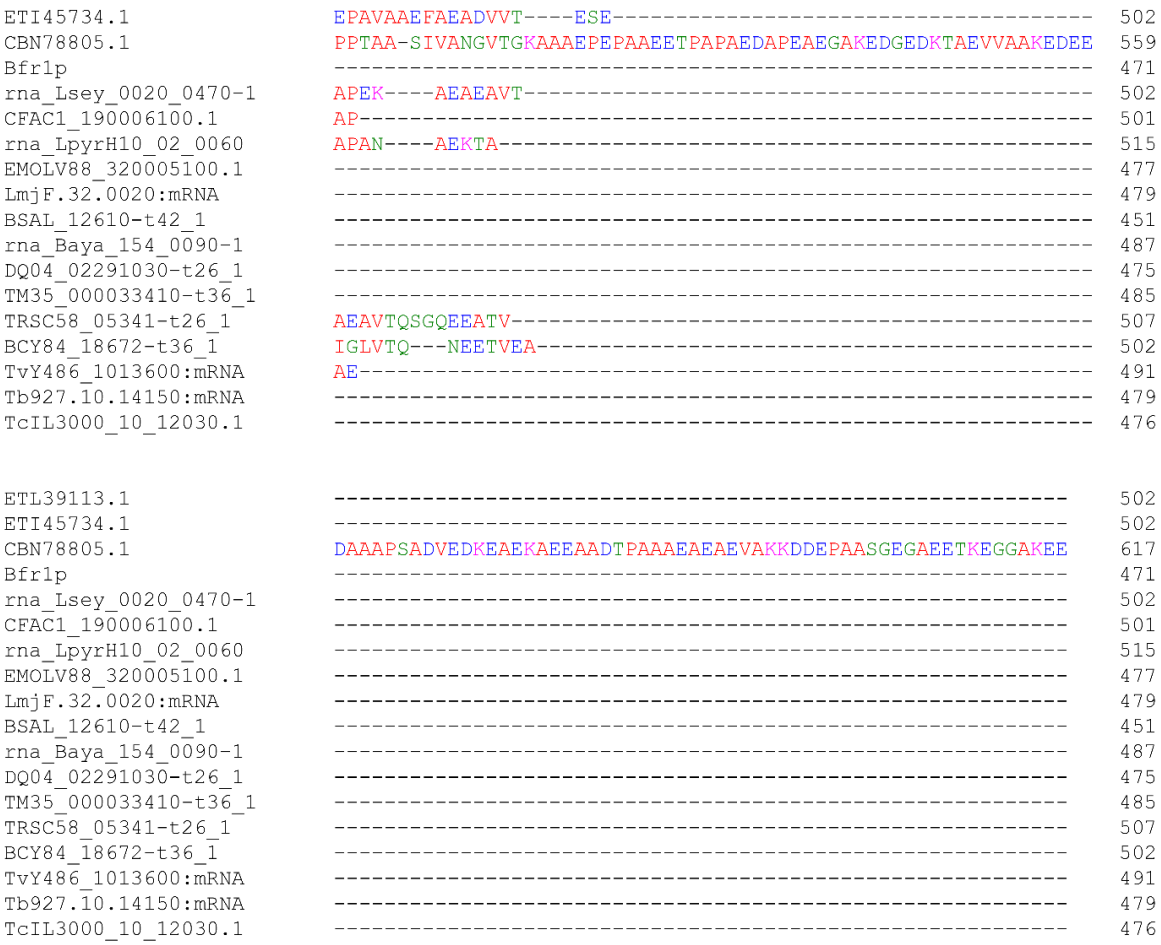
Supplementary material

Supplementary 1



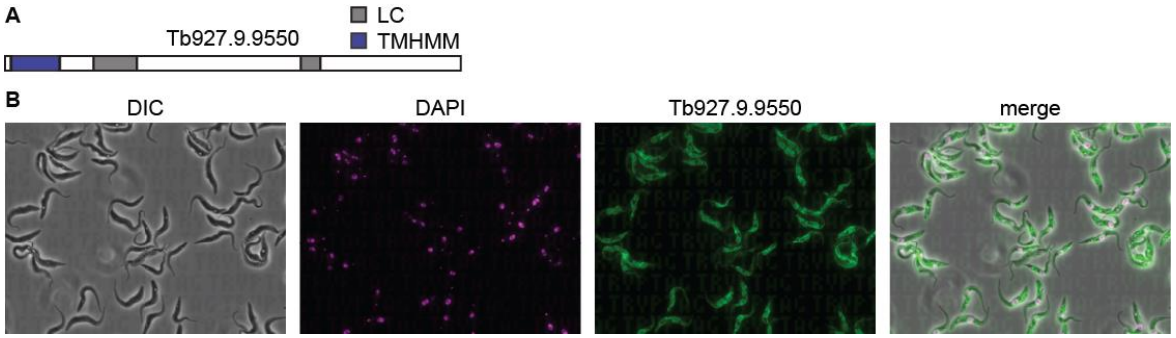
```
Bfr1p                   MQSLNKLIKDLVNIEPIRKSVDADKA-------KINQLKEELNGLNPK--DVSNQFEENQ  190
rna_Lsey_0020_0470-1    LKKLDDAKAQFLKLYPLVEAIREVTE-------QETVLQQEYVAISEQVGIHNKEYDEEM  209
CFAC1_190006100.1       LHRLEEAKKHLSQLQPLTEAMKEVAD-------QETILQQEYLAISEQIGIYNKEYDEEM  209
rna_LpyrH10_02_0060     LNKLEQAKGQLSQLQPLTEAMKEVAD-------QETVLQQEYLAISEQIGIHNKEYDEEM  210
EMOLV88_320005100.1     LQKLEEAKMHLQKLQPLTEAIKEITE-------EEVILQQEYLAICEKIGILNGEYEEKL  189
LmjF.32.0020:mRNA       LHKLEDAKTHLLRLQPLADAIKQITE-------QEVILQQEYHAICEQIGILNREYEEKL  210
BSAL_12610-t42_1        LSQLEEAKSLLLQLQPLQEAITDADD-------REASLQQEYREIHERIGALNKDYDEQY  210
rna_Baya_154_0090-1     LQQLEEVKSLLLQLQPLTEAIQEAGD-------REVMLQQEHREIHERIGALNKKYEEEL  209
DQ04_02291030-t26_1     LNKLEEAKRYLLELQPITEAITEAKH-------REAMLQREWQEINERIRNLGTEYNEQR  189
TM35_000033410-t36_1    LSKLEEAKRFLLEMQPLTEAITEAKH-------REALLQREWQEINERVRNLKSEYNEQR  213
TRSC58_05341-t26_1      LSKLEEAKRYLVELQPLTEAITEAKH-------REAMLQREWQEINDRIRSLWSEYNEQR  218
BCY84_18672-t36_1       LSKLEEAKRYLVELQPLTEAITEAKH-------REAMLQREWQEINERIRNLRAEYNEQR  211
TvY486_1013600:mRNA     LSQLEDAKRYLQELQPVNEAIAEAKH-------REALLQKEFDEINERIRGLTTACNEQR  207
Tb927.10.14150:mRNA     IGQLEDAKRYLQELQPVSEAIAEAKH-------CEATLQREIQEINERIRGLNEEYKGQR  211
TcIL3000_10_12030.1     LAQLEGAKRYLQELQPLSEAVDEAKH-------HEATLLRELQEINERIRDLNSKYKDQR  212
                        :  *.       :      :                  *   :  :        .

ETL39113.1              QALDELYR-----LNEEENKKDKFPALAKERKEIKEQLDEKFTAIKTLRKEFKEANDKYY  260
ETI45734.1              QALDELYR-----LNEEENKKDKFPALAKERKEIKEQLDEKFTAIKTLRKEFKEANDKYY  260
CBN78805.1              TVLEEMGA-----KRDE--EGGGVNALYEQKKALHEEINVKIEEIKDLRSAHRAKLDHWY  250
Bfr1p                   QKLNDIHSKTQGVYDKRQTLFNKRAALYKKRDELYSQIRQIR-------ADFDNEFKSFR  243
rna_Lsey_0020_0470-1    QKKRALDK-------EAHNGSAERSNLYKKFTEVSTKIDEISLTIDALHADFQKSMTEYD  262
CFAC1_190006100.1       QNKRTLDK-------EAQGDNAARTVIYKKRTEISARIDEISQAIDATYAAHQKARAEYE  262
rna_LpyrH10_02_0060     QKKRALDK-------EAQHSNAGRSDIYKQRMEINSKINDISQTIDSMSVKFQKAMSEYD  263
EMOLV88_320005100.1     KQKNAKHK-------EAQEDGAHRAEVYKKCDALRVRIAEISQTVESLRAERDRLSSEWH  242
LmjF.32.0020:mRNA       QKKRAKDK-------EAQADGANRADVYKQCDELRTRVSEITQSMDSLRAEREKVSSEWD  263
BSAL_12610-t42_1        TVKQSKDK-------EAQKTSVDRTVLIKERDDLRQKITKLNEEMTKLREGFNTEKESWE  263
rna_Baya_154_0090-1     HTKQATEK-------NLHTTFQQRTSVYKQCDELREKINKLNEEMDTYRASHNDAVQKWE  262
DQ04_02291030-t26_1     SMKQQREQ-------EMHSTGANRAEVYKKCDAISAKINKLNEEMNTLREEHNKALEAWN  242
TM35_000033410-t36_1    MTKQQKEQ-------ELRKTGVNRQEVYKKCDEISAKIKSLSDEMNTLREEHNKAVEAWN  266
TRSC58_05341-t26_1      ATKQQKEQ-------EIRITGVNRSEVYKKCEEISAKINKLSEEMNRMREEHNKAMEAWN  271
BCY84_18672-t36_1       ATKQQKEQ-------EIRSSGVNRAEIYKKCEEINGKITKLSGEMDRLREEHKKAMDAWN  264
TvY486_1013600:mRNA     STKMEKDQ-------LVRGKGASRQEVYAKCSELHAKITGIVEKMNTLREEHSKAMEAWN  260
Tb927.10.14150:mRNA     STKMEKDD-------KMRSTGANRQEVFKKCDEISAKITSITKEMNALSEDFQKAMEVWK  264
TcIL3000_10_12030.1     SVKMQKDE-------EIRSAGANRQEVFKKCDEINANIAKIAQEMNSLSEAHQKAMEKWN  265
                                              :  :   :  .:                 :

ETL39113.1              NNIRLVRKKKEL-------ERQKEEEARKAEYEAKLAD-YEKEMAKIHPYQDEMDLCDAL  312
ETI45734.1              NNIRLVRKKKEL-------ERQKEEEARKAEYEAKLAD-YEKEMAKIHPYQDEMDLCDAL  312
CBN78805.1              NCRRERTKFKRA-------LHKAEWLERVAAKNAHTKA-LEEEEAKKIPYESEIVLCDTL  302
Bfr1p                   AKLDKERLK------REEEQRLSKLLEQKDVDMGKLQEKLTHAKI--PAFTYEIGAIENS  295
rna_Lsey_0020_0470-1    AWYKQAREKYFAKQEEVREQRRREYEERI--NAHKIAEKKARAVKRQNPYVMEISTCAML  320
CFAC1_190006100.1       AWAKEARERYFAKQEEQRQQRRREYEERI--NAHKIAAKKERAVKRQNPYVMEISACSTL  320
rna_LpyrH10_02_0060     TWYREARDKYFAKQAEQREQRRREYEERI--NAHKIAKKQARAVKRQNPYVMEISACGTL  321
EMOLV88_320005100.1     AWSREARTKYYAQLEQRREQRRKEYEEIR--NAHKIAAKRERAAKRQNPYVTEISACATL  300
LmjF.32.0020:mRNA       AWNKEARKMYFEHLEQQRKRRKEYEERR--NAPKIAAKLARAAKRQNPYAAEISACSML  321
BSAL_12610-t42_1        AWREEAIKKYGEKMEAERKERERRRNEYM--NAEKIARKQARATKRQNPHETQIGACSTL  321
rna_Baya_154_0090-1     TWCAEARERKYTAKMETERQERYKRELERH--TNAKLTEKLNRAQRRRNPYEMEISACDTL  320
DQ04_02291030-t26_1     AWREEARTKYAAKVEAERKERHLRYLERK--NAAKLEEKRARAMRRQNPYEAEVDACSVL  300
TM35_000033410-t36_1    AWREEARAKYAAKVEAERKERQRRYLEYK--NAAKMEEKRARALRRQNPYEAEADACSTL  324
TRSC58_05341-t26_1      AWREEARAKYLAKMEAERKERQRRYLERK--NAAKLEEKRARALRRQNPYEVEIEACKTL  329
BCY84_18672-t36_1       TWREEARAKYIAKMEEERKERRRRYLERK--NAAKIEAKRARALRRQNPYEAEIDACKTL  322
TvY486_1013600:mRNA     AWRAEAFAKHQAKMEEIRKERERRIHEKN--NAAKLEEKRARALRRMNPYEVEIAACDTL  318
Tb927.10.14150:mRNA     SWCDEARAKHMAKMEEVRKERHRRFLERK--NAPKLAEKRERALRRMNPYEVEIAACDTL  322
TcIL3000_10_12030.1     AWCDEARAKHIAKMEEIQKDRQRRIMERK--NAAKLEEKRARALRRMNPYEVELAACDTL  323
                                      :  .          :      .       .  :

ETL39113.1              VSFLEKTYAKELKE-----EQDEKA--AETTAAPLELDGMKPL--QRKEEDFMMLGGGKK  363
ETI45734.1              VSFLEKTYAKELKE-----EQDEKA--AETTAAPLELDGMKPL--QRKEEDFMMLGGGKK  363
CBN78805.1              VSYLETTFMNVKGE-----APKAAAVVAPQAATNDEFKGMKMVMKKRQDEEFIVMGGGRK  357
Bfr1p                   LLVLDPTYVKPKKNILPDLSS-----------NALETKPARKV----VADDLVLVTPKKD  340
rna_Lsey_0020_0470-1    I-----QYLKQKKVMALQDAEERKKR-----EAAAHFDPSQVA----P-AGCVVLNESKW  365
CFAC1_190006100.1       I-----QYLKHKKVMVLQDEEDRKKR-----EAAAHFDPTQVA----P-AGCVVLNEGKW  365
rna_LpyrH10_02_0060     I-----HYLKQKKVMLLQDEEERKKR-----EAAAHFDPSQVA----P-AGFVVLNEGKW  366
EMOLV88_320005100.1     I-----QYLKQKKVMLEQEEQERKKR-----EAAAHFDPSQTA----P-AGCVVLNDSKW  345
LmjF.32.0020:mRNA       I-----EYLKQKKMMVQLDEEDRKRR-----EAAAHFDPSQMA----P-AGSVVVSDSKW  366
BSAL_12610-t42_1        V-----RYLRDRIVMSQRDEEERKRR-----VAMASFDPSASA----P-SGFALAAPIEL  366
rna_Baya_154_0090-1     S-----QYLIDKKHMVTREEEERVKR-----EAAATFDPAKAV----P-AGCVVLNEGKW  365
DQ04_02291030-t26_1     V-----RYLRDHKMMVQREEEELARK-----NAAATFDPTKFL----P-EGAVVLNDGKK  345
```


TM35_000033410-t36_1 V-----RYLRDKKAMVQREEEELARQ-----EAAATFDPTKFL----P-EGAVLLNDGKK 369
TRSC58_05341-t26_1 L-----RYVQDHKVMVQREEAELARK-----QAAAAFDPSKFL----P-EGAVLLNDGKK 374
BCY84_18672-t36_1 L-----RYMQDQKVMVQREEVELARK-----RAAATFDPTNFL----P-EGAVLLNDGKK 367
TvY486_1013600:mRNA I-----QYLREQKVMVKREEEERIKK-----EAVANFDPAKFA----P-SGAVILNDGKN 363
Tb927.10.14150:mRNA L-----QYLRDQKIMVQRENEERARR-----EAAANFDPAEFA----P-EGAVVLNDGMS 367
TcIL3000_10_12030.1 L-----RYLGEQKIMVQRENEERAKR-----EAAAAFDPTKFA----P-EGAVLLNDGKG 368
ETL39113.1 --------GKKGRNG----------KK----TKKASKLVLPLAQMEAFST 391
ETI45734.1 --------GKKGRNG----------KK----TKKASKLVLPLAQMEAFST 391
CBN78805.1 --------KGGGKKK----------AGGKNAAGGKNTIVHAIDTMESFSM 389
Bfr1p DFVNVAPSKSKKYKKKNQQKNTENE----------QPASIFNKVDGKFTLEPTLIATLAE 390
rna_Lsey_0020_0470-1 -ADSKPLSKAAKKQQMLQQKQREKASAAAKAKADAAAQSSAENKDRVLHHSEEKIRLFQM 424
CFAC1_190006100.1 -SDAKPLSKAAKKQQQQQQKQKAAA---AK-PAAAAPRAPAESKDRMLQHPEDKIRLFQM 420
rna_LpyrH10_02_0060 -SDNKPLSKTAKKQQRQQQKQQEKASAVKV-NANDAPQPAAEAKDRLLQHPEDKIRFFQM 424
EMOLV88_320005100.1 -ADNKTPYKSTTKLPKQQKKQKEKIPQAK------AGTTLSDQKERPLHHTDEKIRLFRI 398
LmjF.32.0020:mRNA -SENKPLVKTAKKQRKHQQKQKENTPAVK------PSTAENDQKGRVLQHPEDKIRLFQM 419
BSAL_12610-t42_1 -PKKKSKAA----------PKAEDKTERTVTHNDEKKRLFAS 397
rna_Baya_154_0090-1 -GK-NAKPAPKASKKQQK-TNTAPP----------VTAAKRVDSTRMLQHPEEKMKLFHL 412
DQ04_02291030-t26_1 -WTEPHKAAVGGKKNKQQQKQQQQQ----------QKQQKTEPKNRVLQHGEEKIRLFQL 394
TM35_000033410-t36_1 -RTDANKGGALGKKNKQQQ----QQ----------QKPKAGTAKNRVLQHSEDKIRLFQL 414
TRSC58_05341-t26_1 -FSDSRKGGAGGKHKAAQT----QK----------QKPEKA-PKNRVLQHPEDKIRLFQL 418
BCY84_18672-t36_1 -FSEPHKAVPGGKSKTKQN----QK----------QKSEKAPPKNRVLQHPEDKIRLFQL 412
TvY486_1013600:mRNA -WKDHAKGAVGGKNKQR------PT----------SKKESGASKAVSIKHGEEKVELFTL 406
Tb927.10.14150:mRNA -HQNG---GDSKKQKQQ------AN----------KAKRDSAPKPRVIKHSEEKLELFKL 407
TcIL3000_10_12030.1 -AGEG---KSQKSQH-T------KK----------VNGTESAAKSGVIKHSDEKVKLFKL 407
ETL39113.1 IGLLPPASAASVSESLAAVKTKKVWFNEQTSRPKAGKVVEPAAEAAPAKVVSPKKKSS-K 450
ETI45734.1 IGLLPPASAAAVSESLAAVKTKKVWFNEQTSRPKAGKVVEPAAEAAPAKVVSPKKKSS-K 450
CBN78805.1 LSIAPPPNKAAVPDAIKALKEKKAWFAEQPRPTATTA--------KPAQVGEKSKKAAKV 441
Bfr1p LDVTVPINSDDVKITVEQLKKKHEELLSKQEEQTKQNIESVE---KEIEKLNLDYSN--- 444
rna_Lsey_0020_0470-1 IEVEPALSLATIDEKIHQIEEKKKKYEQE----------IKTGELLLSS--- 463
CFAC1_190006100.1 LEVEPALSLATIDEKIAQIEEKKTKYERS----------IQTGELVLSS--- 459
rna_LpyrH10_02_0060 IEVEPALSLATIDEKVHQIEEKKKQYEQH----------IQTGELVLSS--- 463
EMOLV88_320005100.1 IEIEPTRSRVAIDSTISEIESAKKKYESH----------IQTGELVLSS--- 437
LmjF.32.0020:mRNA IEVDPALSLATIDDKIRHIETLKSQYESH----------IQTGELVLSS--- 458
BSAL_12610-t42_1 VGISAPATLSEVEKTIEQLKKKQAEYESH----------IKTGDLVLSS--- 436
rna_Baya_154_0090-1 IDLEPPITITAFDSTIQAIKAKRKEYESH----------ITTGDIVLSS--- 451
DQ04_02291030-t26_1 IGEEPPLALAAIDDAVKRISAKQKEYESH----------IKTGDLELSS--- 433
TM35_000033410-t36_1 INEEPPLALAAIDASVERIRAKQVEYESH----------KKTGELELSS--- 453
TRSC58_05341-t26_1 VNEELPVALAAIDETMERLRAKQKEYESH----------IKVGELELSS--- 457
BCY84_18672-t36_1 INEEPPVALSAIDGAVETIRSKQKEYESH----------IKTGELELSS--- 451
TvY486_1013600:mRNA IGEEPPKVLDDIDTLLTKITEKRKEYASH----------IKVGELELSS--- 445
Tb927.10.14150:mRNA VDEQPPRFLEDIGGIMENIRAKLKEYSSH----------IKTGEPELSS--- 446
TcIL3000_10_12030.1 VSEEPPQSVGDIDGVMESLRTKQQEYASH----------IKTGEPELSS--- 446
ETL39113.1 NNKFNASDKDAFPSLGGVAA-ELPSWGPGMA----------PAVA 484
ETI45734.1 NNKFNASDKDAFPSLGGVAA-ELPSWGPGMA----------PAVA 484
CBN78805.1 KAKAPDINDEEFPGLPGMAS-KPKTDEAGDGEDGGEGETANGDADAAGDGDGEDKAEKVA 500
Bfr1p ------KEQQVKKELEEKR---LKEQE----------ESEKDKENX----- 471
rna_Lsey_0020_0470-1 ------SDDEEGEETERHEEVEAA----------DNDDASAAAA 491
CFAC1_190006100.1 ------GDDEEGEEENGEAAEQ--PEEAD-K--------EEDA---ASNGDAAAPVAAEE 499
rna_LpyrH10_02_0060 ------GDDEEAEEAEEAQAEAGQPGEGE-P--------AATS--AIAASNNEDTVAAAA 506
EMOLV88_320005100.1 ------GEDEEDDEENTVNDDEVP-SELA-P--------ADAQ--GVVVDDAKKDIAE-- 477
LmjF.32.0020:mRNA ------GDDEEGKEQAHDGADTPL-QRL---------- 479
BSAL_12610-t42_1 ------DDEEEEKEEAAPADE---------- 451
rna_Baya_154_0090-1 ------ESDGEGETHPEDSNLAEPDAAPA----------EHNEEPVPEVTAE- 487
DQ04_02291030-t26_1 ------DDEEEEEQEEPQPQEEKGEEQEV----------VDGA--KTESQLEESEVKADE 475
TM35_000033410-t36_1 ------DDDEEEEEEENVEQEQKEEQNQGDE--------EQGE--DKE---------- 485
TRSC58_05341-t26_1 ------DDEEDEEAQEEEG--AAEATE----------E-EA--PKEADDVE-EVTAGA 493
BCY84_18672-t36_1 ------DDEEEEEEQPQEEEDTAAATE----------E--V--QKKKDEVEKEVTADK 489
TvY486_1013600:mRNA ------DDDEQEHPQEEEEEAEVEDAVESSK--------EKGE--KEEEKEGEEKKDEER 489
Tb927.10.14150:mRNA ------DDEEDEEKKEQEGEEEAVA----------KED--ETNEGEGEDFA---- 479
TcIL3000_10_12030.1 ------DDEEEQEQEQEQEQEQQEN----------EGA--AEEEGDTN------- 476
ETL39113.1 EPAVAAEFAEADVVT----ESE---------- 502



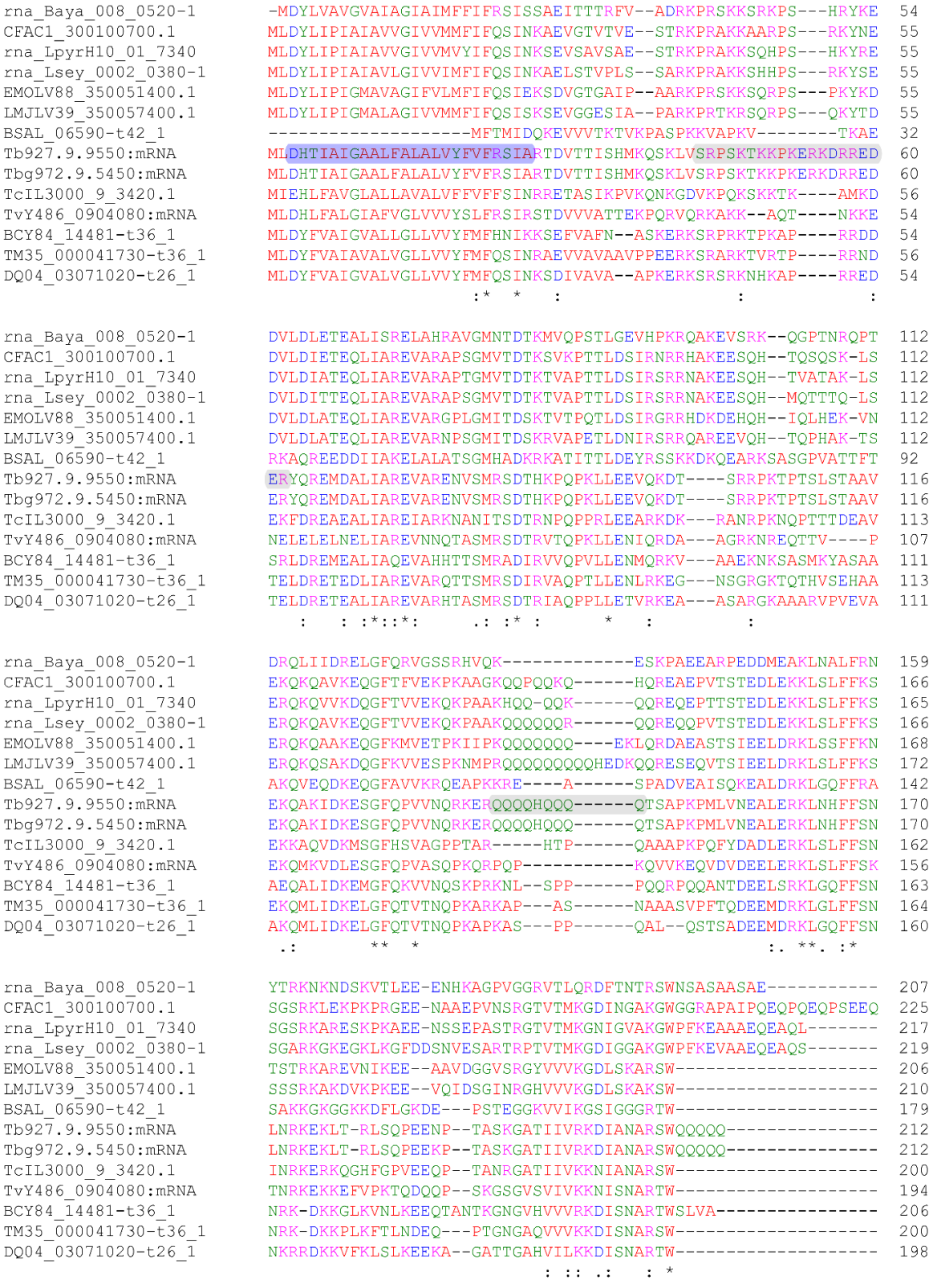
Supplementary Figure 1: Sequence analysis of Tb927.10.14150. Sequences of Tb927.10.14150 homologues were compared using Clustal Omega. Low complexity regions are marked in grey. The asterisk (*) marks fully conserved residue. Colon (:) marks conservation between amino acid groups with similar properties. Period (.) marks conservation between amino acid groups of weakly similar properties.

Supplementary 2



Supplementary Figure 2: Domains and localization of Tb927.9.9550. A. Domains of Tb927.9.9550. LC: Low complexity region. TMHMM: Transmembrane domain. **B.** TrypTag images of C-terminally GFP-tagged Tb927.9.9550 (green). Nucleus and Kinetoplast were stained with Hoechst (cyan) (Dean et al., 2017).

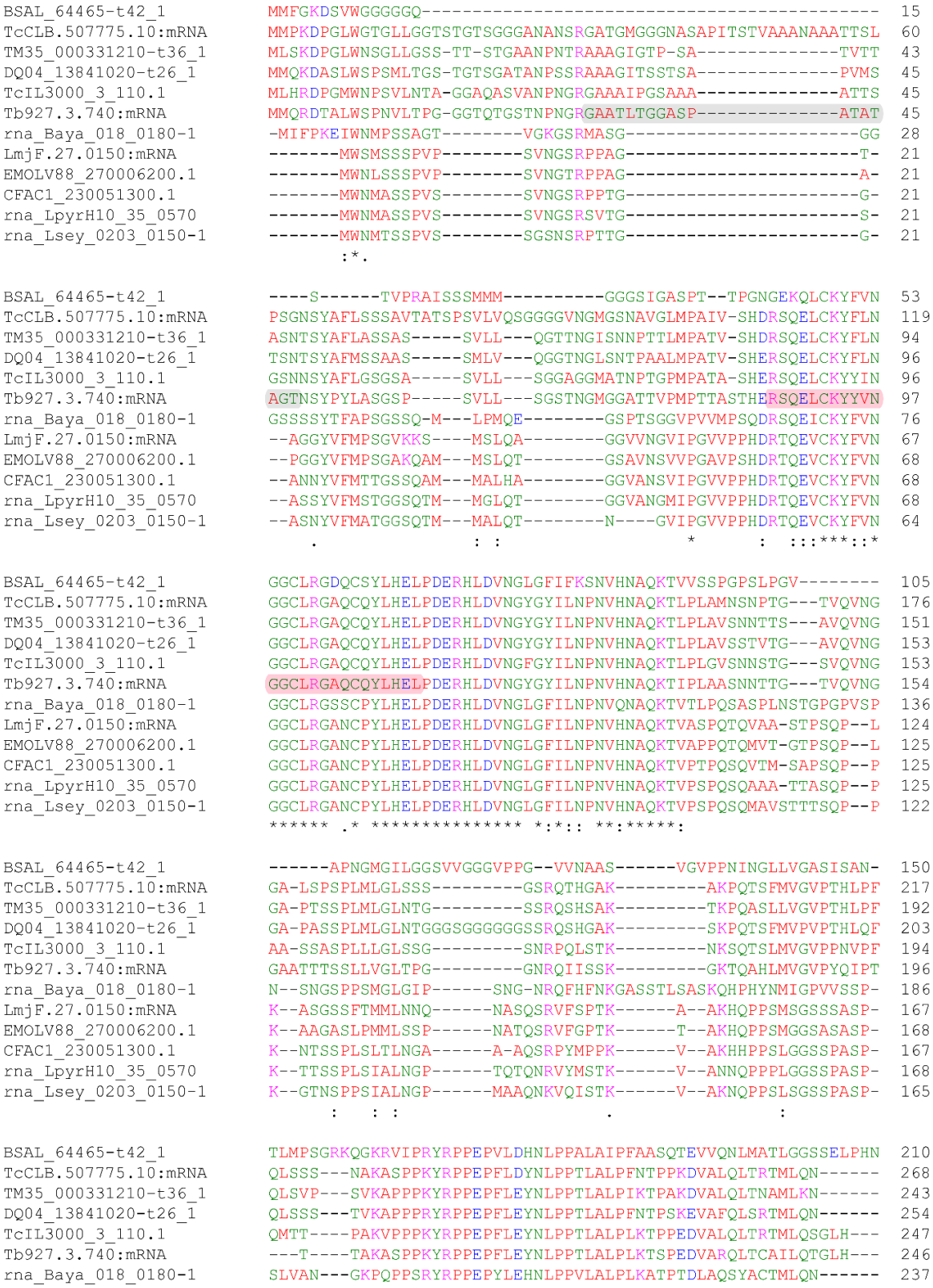
Supplementary 3



```
rna_Baya_008_0520-1       ------        207
CFAC1_300100700.1         VWEAEE        231
rna_LpyrH10_01_7340       ------        217
rna_Lsey_0002_0380-1      ------        219
EMOLV88_350051400.1       ------        206
LMJLV39_350057400.1       ------        210
BSAL_06590-t42_1          ------        179
Tb927.9.9550:mRNA         ------        212
Tbg972.9.5450:mRNA        ------        212
TcIL3000_9_3420.1         ------        200
TvY486_0904080:mRNA       ------        194
BCY84_14481-t36_1         ------        206
TM35_000041730-t36_1      ------        200
DQ04_03071020-t26_1       ------        198
```

Supplementary Figure 3: Sequence analysis of Tb927.9.9550. Sequences of Tb927.9.9550 homologues in Kinetoplastida were compared using Clustal Omega. Low complexity regions are marked in grey and transmembrane domain is marked in blue. The asterisk (*) marks fully conserved residue. Colon (:) marks conservation between amino acid groups with similar properties. Period (.) marks conservation between amino acid groups of weakly similar properties.

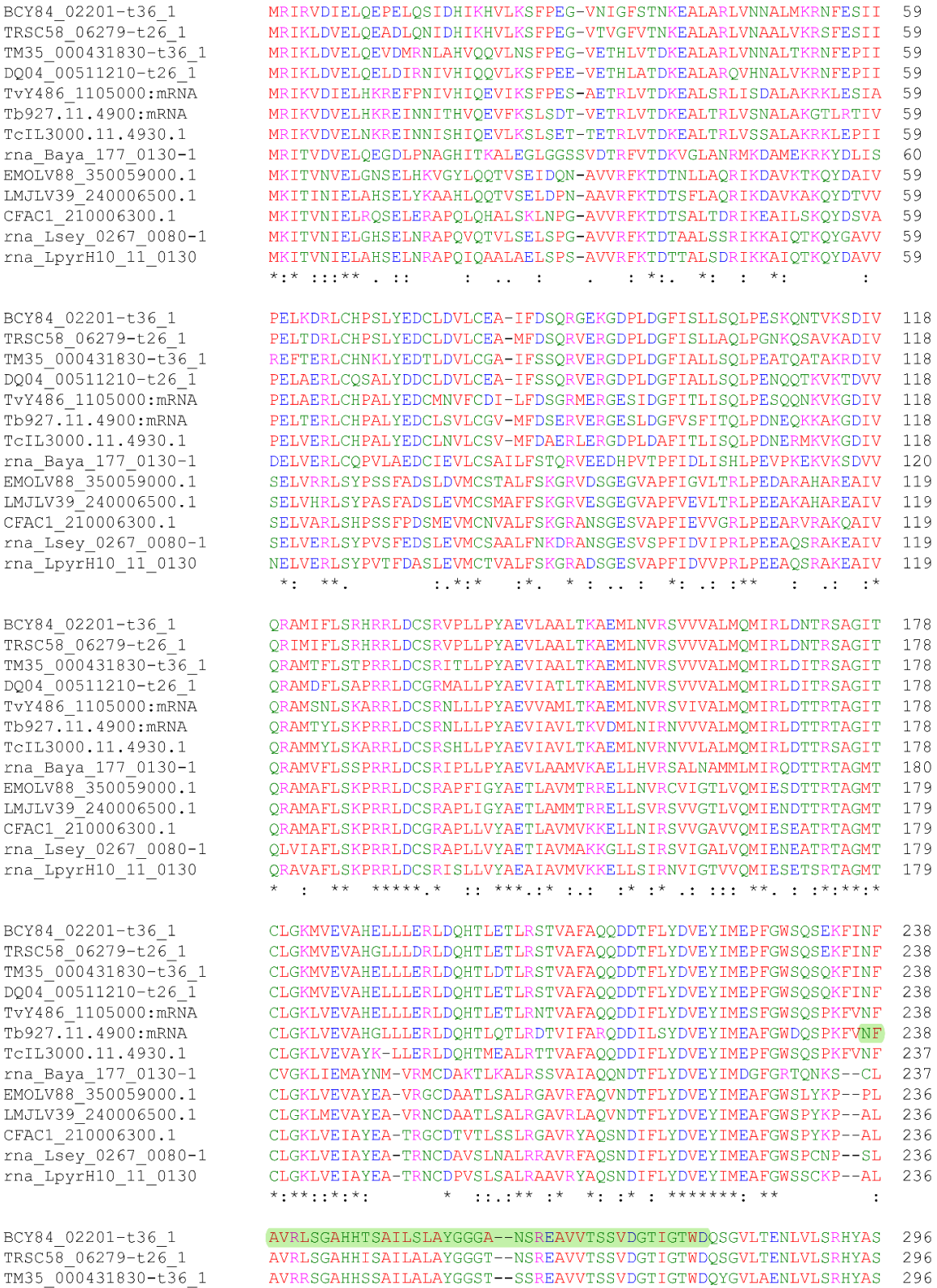
Supplementary 4



```
LmjF.27.0150:mRNA         QTMQM---GKLPAPPKYNPPEPYLELNLPPALAFPLKVPAKDTTATLTHLMLQN------  218
EMOLV88_270006200.1       QMVPM---GKLPAPPKYNPPEPYLELNLPPALAFPLRVSTKDTTATLTHLMLQN------  219
CFAC1_230051300.1         QTAPA---GKLPPPPRYNPPEPYLELNLPPALVFSLKAPVKDTAALLASAMLLN------  218
rna_LpyrH10_35_0570        QTVPN---GKLPPPPKYNPPEPYLELNLPPALAFPLKAPAKDTAASLARAMLQI------  219
rna_Lsey_0203_0150-1       QTAPN---GKLPPPPRYNPPEPYLELNLPPALAFPLKAPTKDTATSLSHAMLQS------  216
                                     :*.**** *: ****.*.: : .   : .      :

BSAL_64465-t42_1          IIGGILRV      218
TcCLB.507775.10:mRNA      --------      268
TM35_000331210-t36_1      --------      243
DQ04_13841020-t26_1       --------      254
TcIL3000_3_110.1          --------      247
Tb927.3.740:mRNA          --------      246
rna_Baya_018_0180-1       --------      237
LmjF.27.0150:mRNA         --------      218
EMOLV88_270006200.1       --------      219
CFAC1_230051300.1         --------      218
rna_LpyrH10_35_0570       --------      219
rna_Lsey_0203_0150-1      --------      216
```

Supplementary Figure 4: Sequence analysis of ZC3H5. Sequences of ZC3H5 homologues in Kinetoplastida were compared using Clustal Omega. Low complexity regions are marked in grey and zinc finger domain is marked in red. The asterisk (*) marks fully conserved residue. Colon (:) marks conservation between amino acid groups with similar properties. Period (.) marks conservation between amino acid groups of weakly similar properties.

Supplementary 5

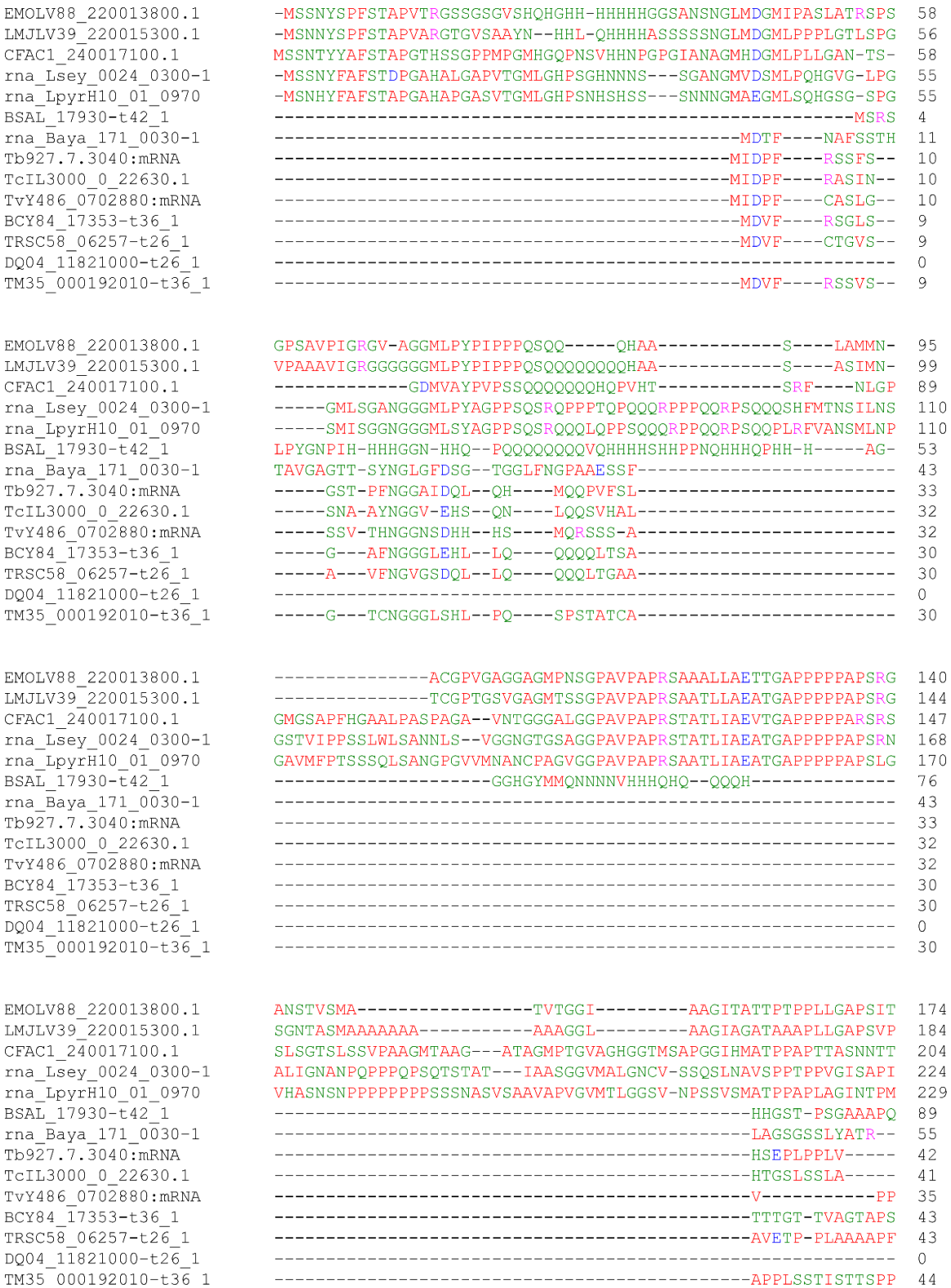


```
DQ04_00511210-t26_1       AVRRSGTHHTSAILTLAYGGSS--GSREAVVTSSVDGTIGTWDHSGVLAENLVLSRHYAS  296
TvY486_1105000:mRNA       VVRRSGTHHNSAILALAYGGGN--GSREAVVTSSVDGTIGTWDHSGVLTENLVLSRHYAS  296
Tb927.11.4900:mRNA        SVRHSGVHHKSAILTLAYSGGN--GSREVVVTSSVDGTIGTWDHSGALTENLVLSRHYAS  296
TcIL3000.11.4930.1        AVQCSRTHHTSAILALAYSGNS--GSREVVVTSSVDGTIGTWDQAGVLTENLVLSRHYAS  295
rna_Baya_177_0130-1       AISRFVSHHTHPILSLAYCGGSGGNWREAVVTSSVDGTISTWDYTGTLTESIMLSRHYAS  297
EMOLV88_350059000.1       SLRRSSSHHEYPILSLAYCGGA--NHREVVVSSSSDGTIGTWDGVGVLLQNVLLSRHYAS  294
LMJLV39_240006500.1       SLRRSSCHHELPILSLAYCGGA--NHREVVVSSSSDGTIGTWDGVGVLLQNVLLSRHYAS  294
CFAC1_210006300.1         TLRRSTGHHENPILSLAYCGGA--NCREVVVTSSSDGTIGTWDGTGILLQNVLLSRHYAS  294
rna_Lsey_0267_0080-1      SLRRSIPHHESSILSLVYFGGV--NCRELVVTSSSDGTIGTWNGVGVLQQNVLLSRHYAS  294
rna_LpyrH10_11_0130       SLRRSTSHHENPILTLAYCGGA--NCREVVVTSSSDGTIGTWDGVGVLLQNVLLSRHYAS  294
                           :     **   **:*.* *.   . ** **:** ****.**:  * * :.::*******

BCY84_02201-t36_1         SMDFANSGRILIVGTVGRHASVAPAIVLYSTDPTYNEESNWQESGGAEPRGARFITTVRS  356
TRSC58_06279-t26_1        SMDFANSGRILIVGTVGRHAAVAPAVVIYSTDPTYNEESNWQESGGAEPRGARFITAVRS  356
TM35_000431830-t36_1      SMDLANRGRTLIVGTVGRNPTVAPAIVLYSADSPYNEESNWQESCGAEPRGARFITTVRS  356
DQ04_00511210-t26_1       SMDFANRGRTLIVGTVGRYTTVAPAIVIYTADATYNEESNWHESCGAEPRGARFITTVRS  356
TvY486_1105000:mRNA       SIDFANRGRTLIVGTVGRIASVPPAIVIYSAESTYNEESNWQENCGAEPRGARFITAVKS  356
Tb927.11.4900:mRNA        SIDFADHGRALIVGTVGRTASIPPAIVIYTAQPTYNEESNWQEKCGAEPRDAMFITTVKC  356
TcIL3000.11.4930.1        SIDFANRGRALIVGTVGRMSSVPPAIVIYTAESTYNEESNWQESCGAEPRGARFITTVKC  355
rna_Baya_177_0130-1       SLDFANRGHTLIVGTVGRHSNTPPAVVLYNEEGN-SREPHWEESGAVEPQNARFITAVRN  356
EMOLV88_350059000.1       CLDLTNRGHTLIVGAVGRYANTPPAVIFYNEDGN--RKAQWQECGGTEPDDAHFISSLKC  352
LMJLV39_240006500.1       CLDLTNRGHTLIVGAVGRYTNTPPAVIFYNEDGN--RKAQWQECGGTEPDDASFISSLKC  352
CFAC1_210006300.1         CVDLTNRGHSLIVGAVGRSANTPPAVIFYNEESN--RKAQWQECGGTEPDDATFISSLKC  352
rna_Lsey_0267_0080-1      CVGLTNRGRSLIVGAVGRYSNTPPAVIFYSEDGN--RKAQWQECGGTEPDDASFISGLKC  352
rna_LpyrH10_11_0130       CVDLTNRGHTLIVGAVGRYSNTPPAVIFYNEDGN--RKAQWQECGGTEPDDASFLSCLKC  352
                          .:.::: *: ****:***     **:::*. :    .: :*.*  ..** .* *:: ::

BCY84_02201-t36_1         LRGTGTLRYCCGVTTASANSVLLYDGTQMIQEYNDHNDIVSAMHVIPDRDHILVSGSRDC  416
TRSC58_06279-t26_1        LKGTGTLRYCCGVTTASANSVLLYDGTQMIQEYTDHNDIISAVHVIPDREHILVTGSRDC  416
TM35_000431830-t36_1      LRSAGTLRYCCGVTTNTSNTMILYDATQVIQEYNDHNDIISAMHVIPDRDNTVVTGSRDC  416
DQ04_00511210-t26_1       LRNAGTLRYCCGATTSATSRMILYDSTQVIQEYNDHNDIITAMHVISDRDNMVVTGSRDC  416
TvY486_1105000:mRNA       LRNAGTLRYCCGVTVTTANTLILYDNTQCIHEYNDHNDIISAMHVICDRENIVVTGSRDC  416
Tb927.11.4900:mRNA        LRSFSSMRYCCGATVATGSTLILYDGTQVIQEYNGHSDVISAMHVIPDRDNTVVTGSRDC  416
TcIL3000.11.4930.1        LRNHGSIRYCCGATVTTGNSLILYDGTQVIQEYSDHNDIISAMHVIPERDNTVVTGSRDC  415
rna_Baya_177_0130-1       IRTYPSLRYCVGVSTATSHLLLLYDRTQLVQEYRDHTDILTAMHIPSDRDSTVITGSRDC  416
EMOLV88_350059000.1       LKDSSGSKYCVGVQTSNSNALMIYDGQQVTQRYFDHTDIITAMHVPSDRENLVITGSRDC  412
LMJLV39_240006500.1       LKDSSGSRYCVGVQTSNANSLMIFDGQHVTQRYFDHTDIITAIHVPSDRENLVITGSRDC  412
CFAC1_210006300.1         LKDPSGSRYCVGVRTSNANPLMVYDGSMVTQRYFDHTDIITAIHVPSDRDNFVITGSRDC  412
rna_Lsey_0267_0080-1      LNDPSGSRYCVGVRTSNANPLLVYDGPCVTQRYFDHSDIITAIHVPRDRDNLIITGSRDC  412
rna_LpyrH10_11_0130       LKDPSGSRYCVGVRTSSANPLLLYDGSRVTQRYFDHSDIITAIHVPSDRDNVVVTGSRDC  412
                          :.     :** *. .     ::::*     :.* .*.*:::*:*:  :*:  :::*****

BCY84_02201-t36_1         SVMIYDLRMAQNTTTPSAYRHYSTVTSIGSCGDYLFTSGLDRRVVVHDFRMMGQGVATRE  476
TRSC58_06279-t26_1        SVMVYDLRMAQAATTPSAYRHYSTVTSIGSCGDCLFTGGLDRRVVGHDLRMMGQGVATRE  476
TM35_000431830-t36_1      SVMVYDLRIPQNATTTSAHRHYSTVTSIGSCGDCLFTSGLDRRVVVHDLRMMGQGMATRE  476
DQ04_00511210-t26_1       SVMVYDLRMVQNTTTTSAHRHYSTVTSIGSCGDYLFTSSLDRRIVAHDLRMMGQGMATRE  476
TvY486_1105000:mRNA       SVMVYDLRMPQNATTTSAHRHYSTVTSIGNCGDYLFTSGLDRRVVAHDLRMMGQGMATRE  476
Tb927.11.4900:mRNA        SVMVYDLRMPQSVTTTSAHRHYSTVTSIGSCGDFLFTSGLDRRVVVNDLRMMGQGMATQD  476
TcIL3000.11.4930.1        SVMVYDLRIPQSATTTSTHRHYSTVTSIGSCGDFLFTSGLDRRVVVHDLRMMGQGMATRE  475
rna_Baya_177_0130-1       SIMVYDLRSRQSAVTF--AHHSSTVSAIGTCGNYLFTAGLDKRVVMDDMRMIGHQPMVRD  474
EMOLV88_350059000.1       STLIYDLRDPRHVSSE--THHTNTVSCIASCEHNLFTAGLDKRIVVEDLRMLRSPTTQRD  470
LMJLV39_240006500.1       STLIYDLRDPRNVSAA--THHTNTVSCITSCDNNLFTAGLDKRVVVEDFRMLRSPNVQRD  470
CFAC1_210006300.1         STLLYDLRDPRSVNSI--AHHSNTVSCITTCDNYLFTAGLDKRLVLADLRMLRGSTAHRD  470
rna_Lsey_0267_0080-1      STLLYDLRDPRNVSSM--AHHTNTVSCITSCDNYLFTAGLDKRLVISDLRMLRGPTTIRD  470
rna_LpyrH10_11_0130       STLLYDLRDPRNVSSM--AHHNNTVSCITSCDNYLFTAGLDKRLVVADLRMLRGSTANRD  470
                          * ::****  : . :    :* .**:.* .* . ***..**:*:*  *:**:      ::

BCY84_02201-t36_1         LDSAVLSLSVNPNMVCAASTMTGIHLINFANNAM--PTCRADGGRMSPRYNAIAWNAQGN  534
TRSC58_06279-t26_1        LDSAVLSLSVNPNMVCAASTMTGIHLINFANSAM--PTCRADGGRMSPRYNAIAWNAQGN  534
TM35_000431830-t36_1      LDSAVLSLSVNPNMVCAASTMTGIHLINFANNAM--PTCRADGGRMSPRYNAIAWNAQGD  534
DQ04_00511210-t26_1       LDSAVLSISVNPNMVCAASTMTGIHLINFSNNTM--PTCRADGGKMSPRYNAIAWNAQGD  534
TvY486_1105000:mRNA       LDSAVLSISVNSNMVCAASTMTGIHLINFSNNGI--PACRADGGRMSPRYNSIAWNAQGD  534
Tb927.11.4900:mRNA        LDSAVLSISVNSSMVCAASTLTGVHLINFGNNPI--LTCRADSMKMSPRYNAVSWNAQGD  534
TcIL3000.11.4930.1        LDSAVLSLSVNSSMVCAASTMTGIHLINFGNNAL--PTCRADGGQMSPRYNAISWNVQGE  533
rna_Baya_177_0130-1       MDSAVLSMSVSSTMQCAVATMTGACLIDFAGGATMPTTVRAECGVATPRYNAIAWNSAGN  534
EMOLV88_350059000.1       MDSAVLSMSVSNSLQCAVATLTGVYLINFSSGTSVPTSSRFDRGGNASLYNAVCWNNAGT  530
LMJLV39_240006500.1       MESAVLSMSVSNSLQCAVATLTGVYLINFSSSTSVPTSSRSDCGGSASRYNAVCWNNAGT  530
CFAC1_210006300.1         MDSAVLGLSVNSALQCAVSTLTGVYMINFANGTTMPTSSRADCGGNASRYNAIAWNNAGT  530
rna_Lsey_0267_0080-1      MDSAILSISVNNALQCAVSTLTGVYVINFANVTTVATSSRADCGSGALRYNTISWNNAGN  530
```

```
rna_LpyrH10_11_0130      MDSAVLSVSVNSSLQCAVSTLTGVYVINFANGTNVPTSSHADCGGSSLRFNTVCWNNAGN  530
                         ::**:*.:**.  : **.:*:**  :*:*..      : : :    :  :*::.**  *

BCY84_02201-t36_1        VLYAGGDSNTLDLFTRSYPDTEAYIA-   560
TRSC58_06279-t26_1       VLYAGGDSNTLDLFTRSYPDTEAYTA-   560
TM35_000431830-t36_1     VLYAGGDSNTLDMFTRSYPDAETYAS-   560
DQ04_00511210-t26_1      VLYAGGDSHTLDLFTRSYPDAEAYAS-   560
TvY486_1105000:mRNA      VLYAGGDSNTLDLFTRSFPEQEAYSS-   560
Tb927.11.4900:mRNA       VLYAGGDNNTLDMFTRRFPETEAYAA-   560
TcIL3000.11.4930.1       ILYAGGDNNTLDMFTRSFADTEVYGA-   559
rna_Baya_177_0130-1      LLYAGGDNMTLDVYARPYGEADTFDAA   561
EMOLV88_350059000.1      ILYCGGDARTLDLFAPAYELN-SFEV-   555
LMJLV39_240006500.1      ILFGGGEAHTLDLFAPAYELN-SFEV-   555
CFAC1_210006300.1        ILYGGGEGRTLDLFVPPFDVN-AFEN-   555
rna_Lsey_0267_0080-1     ILFGGGEGRTLDLFVRPLETA-LFEA-   555
rna_LpyrH10_11_0130      ILYGGGEGRTLDLFAPTYDGN-VFEA-   555
                         :*: **:  ***::.        :
```

Supplementary Figure 5: Sequence analysis of Tb927.11.4900. Sequences of Tb927.11.4900 homologues in Kinetoplastida were compared using Clustal Omega. WD40 domains are marked in green. The asterisk (*) marks fully conserved residue. Colon (:) marks conservation between amino acid groups with similar properties. Period (.) marks conservation between amino acid groups of weakly similar properties.

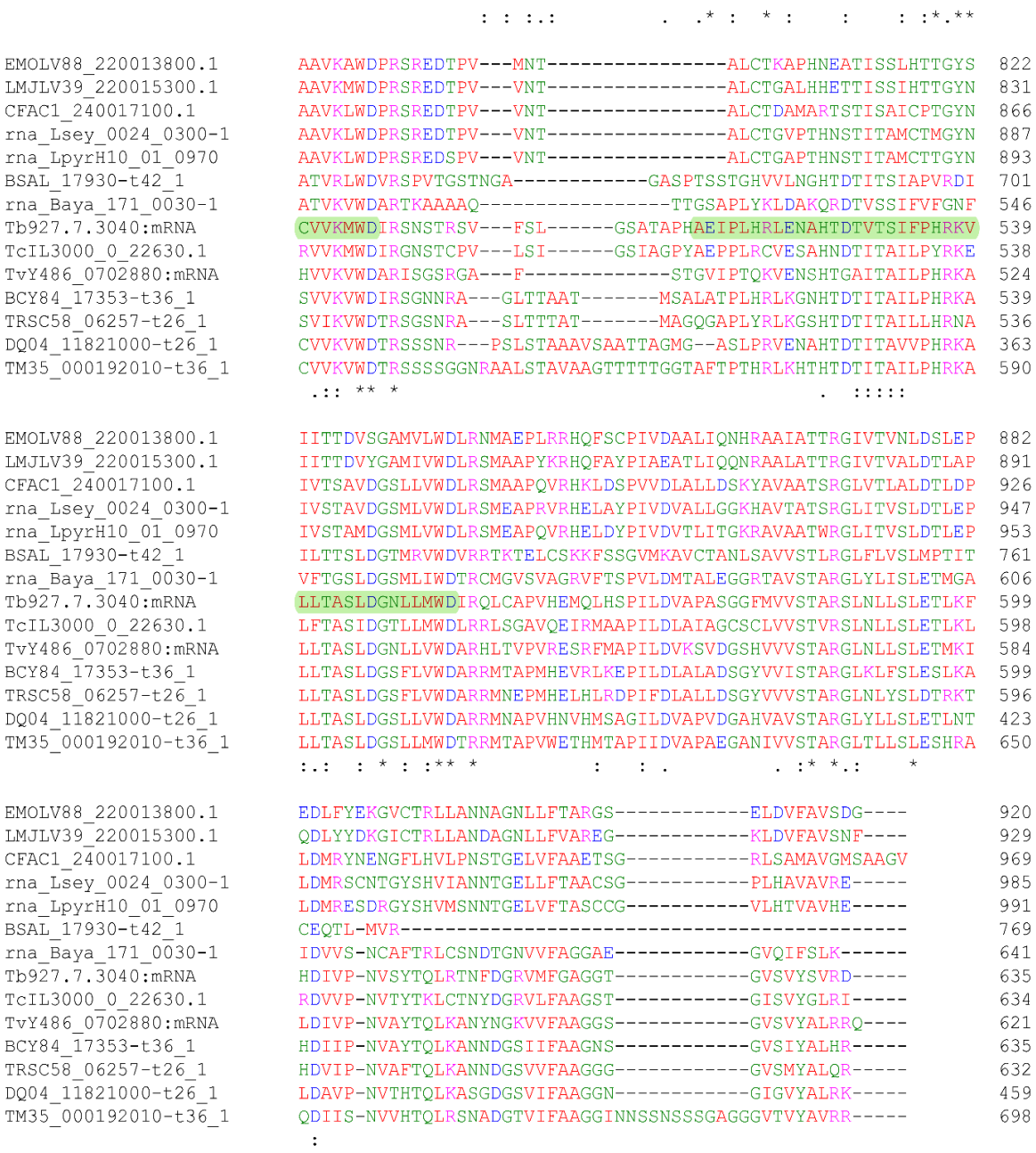
Supplementary 6



```
EMOLV88_220013800.1       TLNTPLA--------AVA------TATPEPVASS-SVISMSPSFKAFTTSASEVLSSMME  219
LMJLV39_220015300.1       TITPSLA--------ALA------ATTPESAAAASAAALVSPSFKVFTTSAAAMLSSMAE  230
CFAC1_240017100.1         GTKTPR---MLPAGGAVAAPR---T-PTGSDPATATTSLVSNSFKAFTTSAATMLSTLGN  257
rna_Lsey_0024_0300-1      GVAAAAASGVTGVGGAAVSRGGAVSPVAVENTSSFATSFMSPSFKAFTTSAATMLSTQGA  284
rna_LpyrH10_01_0970       SAAVA----AAGAGGSVALSGGSVSPVTADNASSFGASLVSPSFKAFTTSVATMLSTQGA  285
BSAL_17930-t42_1          GATQY-----------------------------------------H------------  95
rna_Baya_171_0030-1       ------------------------------------------------------------  55
Tb927.7.3040:mRNA         ------------------------------------------------------------  42
TcIL3000_0_22630.1        ------------------------------------------------------------  41
TvY486_0702880:mRNA       NT----------------------------------------------------------  37
BCY84_17353-t36_1         R-----------------------------------------------------------  44
TRSC58_06257-t26_1        R-----------------------------------------------------------  44
DQ04_11821000-t26_1       ------------------------------------------------------------  0
TM35_000192010-t36_1      NITTT-------------------------------------T------------  50

EMOLV88_220013800.1       SDWMVTPA-DSVGAQAAIPSVATA--TAA------GAATGSSSELARPPSINPHSKVEIQ  270
LMJLV39_220015300.1       AGGLADPV-EGAGGTAACSVAGAQVAAAA------AASTATPLVPTRSPAINPHSKVEIQ  283
CFAC1_240017100.1         AGASRDGA----GAASTEGSA-SS------------AGAARSCADRGVSELSPFSRVEVR  300
rna_Lsey_0024_0300-1      AAASR---MDNVSTAAVEGAA-ALLETVHSGRCSPPAATSAAV--PVATGLSPFGHVEVQ  338
rna_LpyrH10_01_0970       APDAAARGNDGTSASAANAAV-GLLETAHDGHGAPPVTTTAAAAAGACDRLSRFSHVEVE  344
BSAL_17930-t42_1          ----------------------------QHQQHQ---QQQQQTAAPAAQSAQIPADHVHIR  125
rna_Baya_171_0030-1       ----------------------------GS--------TAGGFTAPM-GGGNPPSALLEAA  79
Tb927.7.3040:mRNA         ---------------------------------------------S-NVGASHNGGVEEN  56
TcIL3000_0_22630.1        ---------------------------------------------S-SVGISHNDNIEDN  55
TvY486_0702880:mRNA       ----------------------------------------------DGPVHSGGIEGS  49
BCY84_17353-t36_1         ------------------------------------M-------------VGSGNVSAH  54
TRSC58_06257-t26_1        ------------------------------------T-------------FGGGHVSAH  54
DQ04_11821000-t26_1       ------------------------------------------------------------  0
TM35_000192010-t36_1      ---------------------------TN--------TAANTHTHA-HAYSNIMGGLRPH  74

EMOLV88_220013800.1       FTLSPRQERGGVDE------GSVFGFEDESGGEAAADQLMQVLMALKHLQSFTDPQE--E  322
LMJLV39_220015300.1       FTLPTTREHVDADE------CDAFGFEEESEDDYAADILTQVLMTIKDVDRNNDPEE--A  335
CFAC1_240017100.1         LALENCDPPAGESAAAAAVEDGAATWDPTVETAFEVDPTSLFLATFTIGDHEEE-A--FA  357
rna_Lsey_0024_0300-1      CTLPTRERG-----DTGMNESGLDPWYPTQDIEDELGQLGYIFACLSNLDDDGIPGCTTV  393
rna_LpyrH10_01_0970       LTLTGCGRR-----DTITSEDGSVQWDPAQDVDDELDELGYLLATICCDGDDDSVSYTMH  399
BSAL_17930-t42_1          FRMPLEIPL----------------TATPQERQAKISELSKLLTTLREVAKGKP-L----  164
rna_Baya_171_0030-1       TTY-F-STS----------------IEADRIDDGVLNRISRTLGALQNISEGRH-L----  116
Tb927.7.3040:mRNA         TTLHMT-VT----------------EEVDAFDEGQLEKLRRTLQLLEGISQGRH-A----  94
TcIL3000_0_22630.1        TTLHVT-VT----------------EDIDSFDEKQVEKLRRTLQLLEGISQGCH-L----  93
TvY486_0702880:mRNA       TTFHIT-AT----------------TEVDPCEEKQLEHIRRTLMLLGGISQGLH-I----  87
BCY84_17353-t36_1         TTVHVT-TS----------------TEVDPFDEKRIDEIRRTLRMLEDIGHGRH-I----  92
TRSC58_06257-t26_1        TTVHIT-TS----------------TEVDLCDEQRIDDIRTTLRMLEDIGHGRH-L----  92
DQ04_11821000-t26_1       ------------------------------------------------------------  0
TM35_000192010-t36_1      TSMHFTAGP----------------FDVGPSEAARLEELRNILRLLDDVSRGRH-H----  113

EMOLV88_220013800.1       SL--EGEAQKWYGFILD--CGNPHHAAVALKARMS--QLVDRDLSIMLEGLAAAIIHYNI  376
LMJLV39_220015300.1       DL--EMRGKEWYIFILD--SGDARQAAAALKAKMS--QLTDNELRVMLNGLASAIIYHNT  389
CFAC1_240017100.1         HLEEADQGDRERLDLID--LLNPPKSAALIGARLDDCLKRRVEIQSTLFSLTLAIVAGEG  415
rna_Lsey_0024_0300-1      STANMDEGNNHHLDLFY--NSNLRVAAAKVCKLIEENRHDTSKLRLTFREVTLAVISGGV  451
rna_LpyrH10_01_0970       SLQDMDYGRSYAVDVFT--GSDVKSAAANICKIIDKRQSDTADLRAAIQGLTFAIVSQDE  457
BSAL_17930-t42_1          --------VDWD-SIVATAVEDPVEGAKVIAQSLQ--RFPSQ-LPAHLSAVAQRINGPER  212
rna_Baya_171_0030-1       --------VDWM-SLID--VENPVTSAQRINANLA--NVPAFSINDAYRSISGCITSPRH  163
Tb927.7.3040:mRNA         --------VDWK-GLIN--PSKPVESAKNVSSMLA--RPTTL--PSAIECVAKCITTPEH  139
TcIL3000_0_22630.1        --------INWK-EIVD--AKKPAESAQHISSMLA--KPTIL--PSAVESIAKCITTPEH  138
TvY486_0702880:mRNA       --------VNWN-SLVD--PSKPVESAKSISAVLA--KPTTL--PSAIESVAKCMTTPAH  132
BCY84_17353-t36_1         --------INWT-ELVD--PSQPVESAKKISAALA--RPTTV--PGAISSVAECITTPAH  137
TRSC58_06257-t26_1        --------INWA-ELVD--PSQPVESAKKISVALA--RPTTV--PSAIASVAECITTPAH  137
DQ04_11821000-t26_1       ------------------------------------------------------------  0
TM35_000192010-t36_1      --------VVWA-ELVD--PRQPVESARKITAALV--QHPAS--SGVIEAIAACITTPAH  158

EMOLV88_220013800.1       WSVALLMGASEKKTLCHLLLRIVLRRLVSHRH----------------------------  408
LMJLV39_220015300.1       WAAVLHLGIQEKATLRHLVLRIVLKMLVAHRC----------------------------  421
CFAC1_240017100.1         YELVPLLPVRHSMPLRYSLVYHLLTRLLDARR----------------------------  447
rna_Lsey_0024_0300-1      YKVILYLPRLEAVTLRYSVVKYLLESLVHDHR----------------------------  483
```

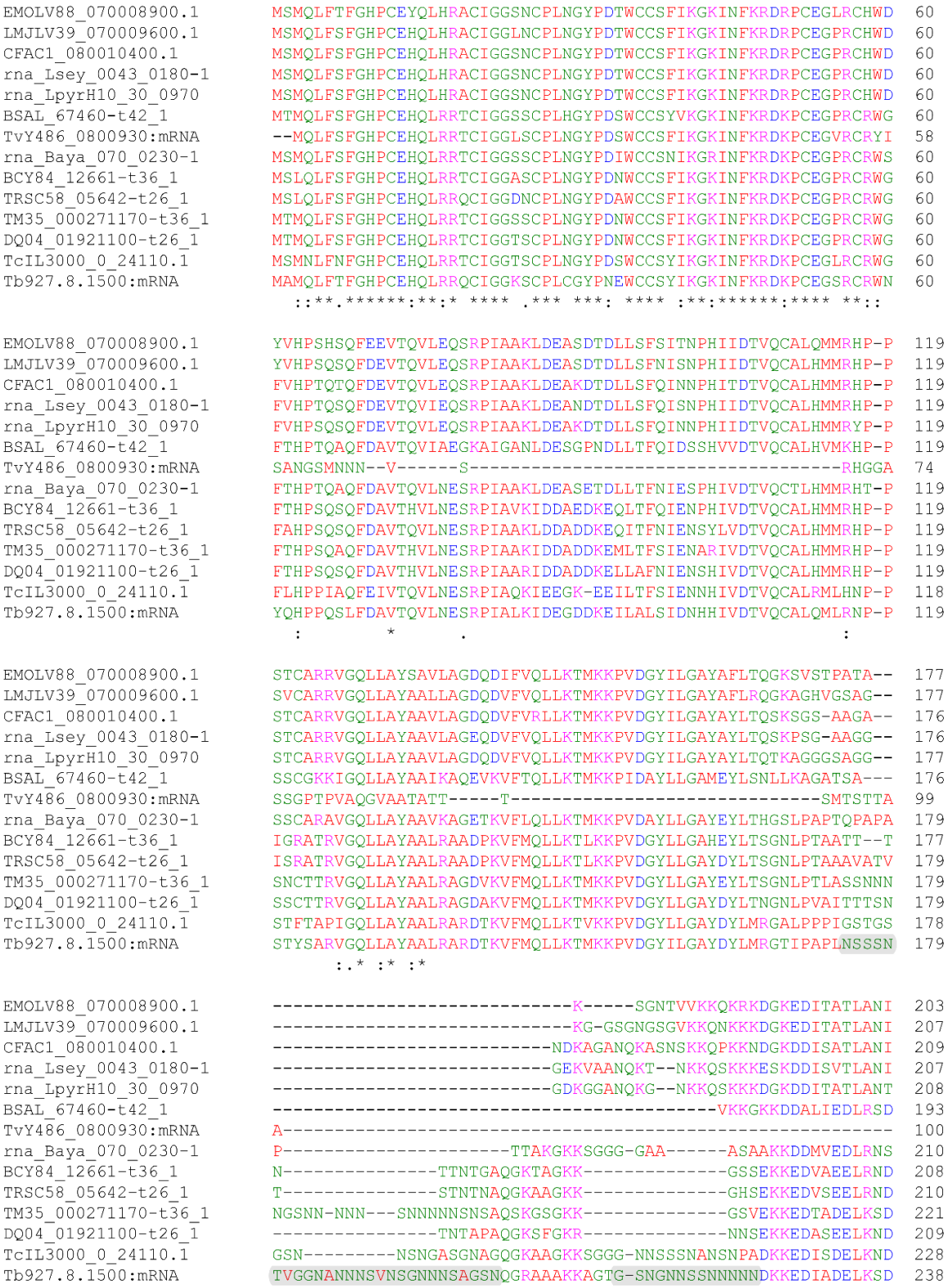
```
rna_LpyrH10_01_0970      FRVLLHLPRQEALTLRYSILKYLLERLVRDHR---------------------------  489
BSAL_17930-t42_1         LAVVALLEEPYRSKVRDEVLQGVLKALTSVTITTESSPSVTATTATPSAAAQRTSTPPPP  272
rna_Baya_171_0030-1      LAIVTNLQEKYKLEVRDVVLKDLLDRLTMVG---------------------------  194
Tb927.7.3040:mRNA        LSIVAKLDQPFRSQVRDAVLQELLNTLTTER---------------------------  170
TcIL3000_0_22630.1       LSIVAKLDPPFRVQVRDAVLRDLLNTLTTER---------------------------  169
TvY486_0702880:mRNA      LSIVANLEPPFWSQVRDAVLHDLLNTLTTER---------------------------  163
BCY84_17353-t36_1        LSIVALLDPPFLAKVRDAVLQDLLNALTMVH---------------------------  168
TRSC58_06257-t26_1       LSIVALLDPPFMAKVRDAVLHDLLNALTMAH---------------------------  168
DQ04_11821000-t26_1      ------------------------------------------------------------  0
TM35_000192010-t36_1     LAIVAALEAPFRVQVRDAVLQELLTALTTSR---------------------------  189

EMOLV88_220013800.1      --------------------------ASHDAEKYVSVLVKMLLVGWVSVKGIVDTVELC  441
LMJLV39_220015300.1      --------------------------VPCEAEKYVAVLVKMLIVGWVSMKGTIDTVELC  454
CFAC1_240017100.1        --------------------------SLSTASPYIKGLAAMLDCGSINARSYAAMVDMF  480
rna_Lsey_0024_0300-1     --------------------------REMDVEKPAKDFASLVDLGLVSAAGFAAVVDTF  516
rna_LpyrH10_01_0970      --------------------------RSADVDTFGKKFTSMMNLGLVRSMELAAVIDTF  522
BSAL_17930-t42_1         PISTVDSESSSDQQQTPHLLQQHAEEAEEGFVPLATRMMVEMIRLDLVLIRGVATTLDTL  332
rna_Baya_171_0030-1      --------------------------D-QPMPVCAEMLVEMLNLNLVRLSGVAVMIEQF  226
Tb927.7.3040:mRNA        --------------------------DPPPMPVCSEMLAEMVQLNLVVLRGVSRTLETL  203
TcIL3000_0_22630.1       --------------------------DPSPMPVCSEMLAEMVQLNLVVLRGVSRTLETL  202
TvY486_0702880:mRNA      --------------------------ESHPMPLCSEMLAEMVRLNLVVLRGVSKTLEAL  196
BCY84_17353-t36_1        --------------------------ESNPMPVCSEMLVEMVKLNLVVLRGVASTLEAL  201
TRSC58_06257-t26_1       --------------------------DLDPMPVCSEMLVEMAKLNLVVLRGVASTLEAL  201
DQ04_11821000-t26_1      -----------------------------------------MVRLNLVVLRGVATTLEAL  19
TM35_000192010-t36_1     --------------------------NPGPMPVCSEMLAEMVRLNLVVLRGVATTLEKL  222
                                                           :   . :        ::

EMOLV88_220013800.1      LQHEDLIHIGLRILIHCFSKSPNAFELRAALD---------------------------A  474
LMJLV39_220015300.1      LQHDELIHVGLRIVEHSFHKSRRVVELRAALA---------------------------T  487
CFAC1_240017100.1        LQHDDLVPVVLRLLATGLSKNRQKDNLCAALK---------------------------T  513
rna_Lsey_0024_0300-1     LSHEAMVPVALRLLTFGLGSVDMGKTVREALE---------------------------A  549
rna_LpyrH10_01_0970      LQHETMVSAALRLIALGLSSVNADKGVRAALE---------------------------G  555
BSAL_17930-t42_1         LRDRNARQAAVAVIGLLAERYAQSQSTAAASANNSKQGKSKQQQQQSSSTTTTSTNDEGI  392
rna_Baya_171_0030-1      LSEPDKCRAAIAVLGRVAEQKPGDEAVRRAIR-----------------------D---A  260
Tb927.7.3040:mRNA        LSDTNTRRAAIAVLGKLADRSRGDAVFARAVQ-----------------------N---L  237
TcIL3000_0_22630.1       LVDSTTRRAAIAVLGKLAERSRGDAVFARAVQ-----------------------C---L  236
TvY486_0702880:mRNA      LNDPGTRRAAIAVLGKLADHRRGDEVFALAVQ-----------------------N---L  230
BCY84_17353-t36_1        LSDPNSRRAAIAALGKLADQKRGNEVFLGAMQ-----------------------N---L  235
TRSC58_06257-t26_1       LGDPGSRRAAIAALGKLAEQQRGNEVFLGAIQ-----------------------N---L  235
DQ04_11821000-t26_1      LSDASTRRAAVAVLGKLADQHRGDEVFLRAVQ-----------------------N---V  53
TM35_000192010-t36_1     LDSPDTRRAAIAVLGRLADQNRGDELFLQVVQ-----------------------N---V  256
                         *         :  :                .

EMOLV88_220013800.1      QP-----RLAQRLA---ELYATNPTFEMEVMLVTRW--WRNLPPSICCS-----------  513
LMJLV39_220015300.1      RP-----RLARRLA---ELYATEPRFEVDVALVTRW--WRGTPPTIYCS-----------  526
CFAC1_240017100.1        TP-----RIAQRLR---AIYTENPSYEVDVALVTRL--WDKAVPFLNAT----------  552
rna_Lsey_0024_0300-1     VP-----RIAQQLH---ELYAHSPQYEVDVALITRL--WQREKPLLFGS-----------  588
rna_LpyrH10_01_0970      VP-----RIAQRLH---EVYARNRKYEVDVALITRL--WQRKEPVLYGC-----------  594
BSAL_17930-t42_1         LPMLNHVRSALQLAAADAAVNSCGELDYDLNAIYRYMDWRHTPPPPTAGVATLQPAAVVA  452
rna_Baya_171_0030-1      KGIESKLR----------ELYAHPEYEYDVVTITRLLEGTEAVK--------------AV  296
Tb927.7.3040:mRNA        EPLV--------------RAINDTEYEYDRVAIARLLRWGGSER-----------EASLV  272
TcIL3000_0_22630.1       EPLL--------------RNIEEPEYEYDRIAIGRLLNWNGMGN-----------EVSLV  271
TvY486_0702880:mRNA      EPLL--------------HSINEPEYDYDRIAISRLLRWCGAEN-----------DARLV  265
BCY84_17353-t36_1        EPLV--------------RMIDEPEYEYDCLAISRLLGWGGAHK-----------AASLV  270
TRSC58_06257-t26_1       EPLV--------------RTIEEPEYVYDCLTIARLLGWDGPHR-----------AASLV  270
DQ04_11821000-t26_1      EPLV--------------RAIQEPEYEYDCVAIARLLGWGGADK-----------AAALV  88
TM35_000192010-t36_1     EPLV--------------RRIKEPEYDYDCITIARLLGWSTAERTGAGGTTTTAAAPSLV  302
                                           :   : *

EMOLV88_220013800.1      ------------------VPMENPQVCFPLGPVTCMEYFNVRDEVVTGDMRGIVTLWGPP  555
LMJLV39_220015300.1      ------------------VPMENSQGDSLSNSVTCMAYFGVRDEVVTGDMSGAVTLWGPP  568
CFAC1_240017100.1        ------------------VPFETSFE-KSQCPVTCMSYFGVRDELISGHMNGSVVLWGPP  593
rna_Lsey_0024_0300-1     ------------------LQLQTTFS-SNDCPVTCMSYFGVRDELITGHMNGSVVLWGAP  629
rna_LpyrH10_01_0970      ------------------LELESNFH-RGNCPVTCLSYFGVRDELISGHMNGSVVLWGAP  635
BSAL_17930-t42_1         SSSSTNTKDSAAPHLQTMVRVPKP-LATPHEAITSMAYFACRDQLVTGGAERTISVWGSQ  511
rna_Baya_171_0030-1      S-------------LRSVC------TLKSVAPVTCLHYVRLRDELFSSRADGTTTLWGAP  337
Tb927.7.3040:mRNA        L-------------ERIIA------SQPQQQYQVTSMAYIGQRDELASATCDGSVVIWGSP  313
TcIL3000_0_22630.1       R-------------ERVIS------HVPYHCQVTSMAYVSGRDELASAACDGSVFIWGAQ  312
```

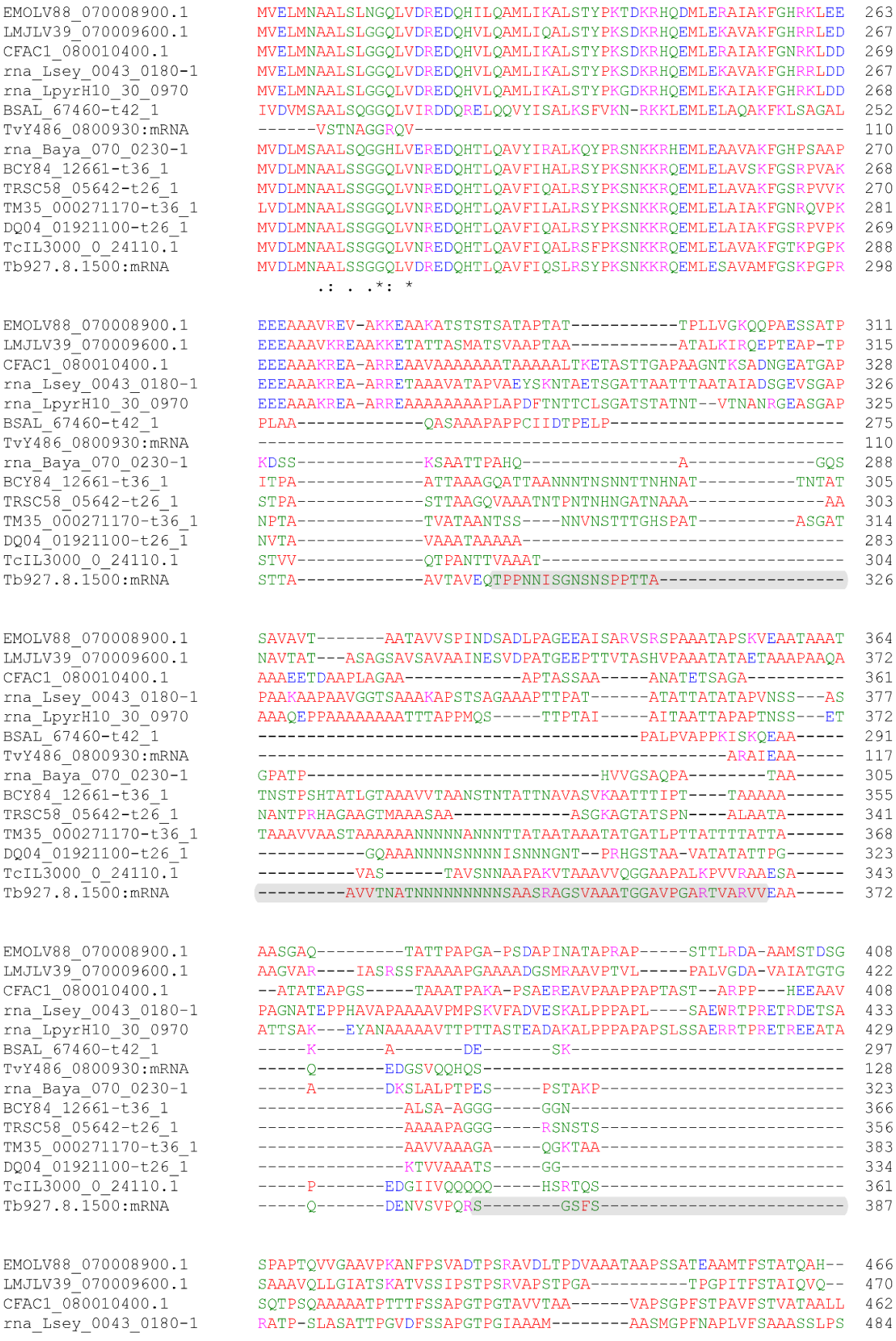
```
TvY486_0702880:mRNA       R-------------VKNI--------FLRQQCPVTCMAYFTPRDELASATCDGTVIVWGGP  305
BCY84_17353-t36_1         P-------------VK---------RMDHSSPVTAMTYFRQRDELVSATFDGTVTIWGTP  308
TRSC58_06257-t26_1        P-------------VK---------KMEQPSTVTSMAYFRQRDELVSGRCDGVVTIWGGP  308
DQ04_11821000-t26_1       P-------------VR---------TLEQPCLVTSMAYFRQRDELASATCDGSVVVWGSP  126
TM35_000192010-t36_1      P-------------VK---------TLEQSCAVTSMAYFRQRDELVSATCDGSVVIWGSP  340
                                                    :*.: *.  **:: :.       :**

EMOLV88_220013800.1       MSTRQPLAKNTLGVQHISVRPRGVVPLPMNCVPVAMSGQRLDGQYLAIAGMPLKTRKPYS  615
LMJLV39_220015300.1       GSSRQPSAKNPLGFPHVSVRPRGVVPLPMNCVPVAMAGQRLDGQYLAIAGMPFKSRKPYS  628
CFAC1_240017100.1         QMVRDTKAGSFLPQTRAVIRPRGIVPLPLDCIPVGMAGQRMDGQYLAIASMPYKSRKPYA  653
rna_Lsey_0024_0300-1      QMVREKSRGSSFSLSRPIIRPRGVVPLPLDCVPVGMAGQRIDGQYLAIASMPYKSRKPYL  689
rna_LpyrH10_01_0970       QMVREKPHGGLFAVPRPAVRPRGVVPLPLDCVPVGMAGQRIDGQYLAIASMPYKSRKPYL  695
BSAL_17930-t42_1          VS---------------GEVPSASFQLPAHTTPVSMDAS-IRGSLLAVGCVTSSS-----  550
rna_Baya_171_0030-1       NAA-------------EVVQPKPVLELQKNCVPVAMAGL-TRGYHTVVASMPCHPGSPYS  383
Tb927.7.3040:mRNA         NAVS------------KEVRPATSIDLPQNYVPVALEAP-PNGSYMVVAGMPISLASQQA  360
TcIL3000_0_22630.1        NAAS------------KDVRPATSIDLPQNYLPVALDCP-PSGDYMVIAGMPISALSQQV  359
TvY486_0702880:mRNA       HPST------------GEVRPSVMLELPQNHLPVAMDSP-SRGNYLAVAVVPIPAANMQA  352
BCY84_17353-t36_1         HPST------------GDVKAAVALDLPQYCVPVAMDGP-PRGNYLVIAGMPFPAVNVQT  355
TRSC58_06257-t26_1        HPST------------GDIKAAMTIDLPPQCVPVAMDGP-PRGNYLVIAGMPFPAANLQA  355
DQ04_11821000-t26_1       HPST------------GEVRPATTVELPQRCVPVAMDGA-SRGNYLVVAGMPFPASSISA  173
TM35_000192010-t36_1      NPIT------------GDVKPSVTIDLPQQCVPIAMDGP-PRGNYLVIAGMPFPSGNLRA  387
                                          . *     *:.:      *   .:. :

EMOLV88_220013800.1       KYSEQLSTRPSSTVASATGMTAMDGVVEGVGSA-----------GDAAGVVTGSAAGTGN  664
LMJLV39_220015300.1       KYCGRPAVRPSSTSAAATAMAAADGAVDGAGVA-----------GDAASMATGAAVGTGS  677
CFAC1_240017100.1         AYCGAPPSGAAAAGRMG------RAEAEGGDREETNAAQRRDGMAGAAMHGGGGGGGGAG  707
rna_Lsey_0024_0300-1      RFCATGEAKAPTGSKAA------D-TADS-NNA----STSADSN-STQSTYRREASNGAG  736
rna_LpyrH10_01_0970       PYCTGVEARAATGSTDS------D-AAENINKA----STFAKGN-TTPSATRREAPNGAG  743
BSAL_17930-t42_1          ---------------------------------------------------KGAV-----P  555
rna_Baya_171_0030-1       TYVEHRKRL-------------------------------------SSGKAGDEQATVD  405
Tb927.7.3040:mRNA         MMKASGSGISGHN---------------------------------WEKRRQASPSRVR  386
TcIL3000_0_22630.1        MIKATSGSGANGQ---------------------------------VDKKRPAPPGRVK  385
TvY486_0702880:mRNA       MMKEYGIST-------------------------------SSHGGGKKKQSLSSIAK  378
BCY84_17353-t36_1         MELERGSINP--NV--------------------MN------NSFNSSGKKRSSFATVPR  387
TRSC58_06257-t26_1        MEFENGSMST--NL--------------------LN------NS---SSKKRNSPAAVTK  384
DQ04_11821000-t26_1       IAGENSATGSSGN---------------------SN------SSSNNNSKRRSLSSGTSK  206
TM35_000192010-t36_1      MMNEGINGSDNNNSSNN-------------NNNNTT------SNGSGSSKKRNILSSPAK  428

EMOLV88_220013800.1       AGAIIVITCNENSLRWTKGEIVMRPPGTVLTAITAF--RNSIVAVGESAVSVVATTSGMS  722
LMJLV39_220015300.1       AGAIIVITCNENSIRWNKGEIVLRPPGIALTAITAF--RNSIVAVGESAVSAAAA----G  731
CFAC1_240017100.1         VGAIIVITCNENSLRWNKGEVILRPPGVALTAITAF--RNSIVAVGESVVKPTSRLVSAA  765
rna_Lsey_0024_0300-1      VGAIIVITCNENSIRWNKGEVIMRPAGVALTAITAF--RNSIVAVGESMVQWATPSA--A  792
rna_LpyrH10_01_0970       VGAIIVITCNANSIRWNKGEVIMRPAGVALTAITAF--RNSIVAVGESAVLRATPSA--D  799
BSAL_17930-t42_1          SSCVQAYLC-SDSGAWTVGDIIKRPDRTVLTNVKCLAVKGSVVCTTETLLA---------  605
rna_Baya_171_0030-1       GPHLRVIGFNEEWGMWSCTKTIRRGTDTTLTAVAAL--SNSTICSAESVRV---------  454
Tb927.7.3040:mRNA         APLLRFLTCNEGTGVWMSGDTTMRKENTALTAAVAL--PNMVVCTAESYPSERSS-----  439
TcIL3000_0_22630.1        GPLLRFLTLNEDTGVWTSGETAVRRDNTTLTAATAL--ANMVVCAAESCPADRAS-----  438
TvY486_0702880:mRNA       GPYLRFLTCNESTGVWSHGESISRKENSTITTVAAL--SNHVVCVAESCISERPS-----  431
BCY84_17353-t36_1         GPIVRLLTCNESTGAWTSGETITRKETVILTTAAAL--ASFVVCTAESGPSDKLS-----  440
TRSC58_06257-t26_1        GPFLRFLTCNESTGLWTNGETISRKENVTLTAVAAL--GNLAVCTAESAPTEKFS-----  437
DQ04_11821000-t26_1       GPLLRLLTCNESTGAWTNGETITRKENVTLTAVAAL--ANLVVCTAESAPSERSP-----  259
TM35_000192010-t36_1      GPVLRFLTCNESTGSWANGETITRKEHVTLTAVAAL--SNLVVCTAESIPSERLT-----  481
                             :           *   .   *     :*   .:   .  :.  *:

EMOLV88_220013800.1       GG-SGSGISSGSASAPHRLSFVDVVNGVAMRQIPKAHDNYITSLTVLEECSYVLLSGGRD  781
LMJLV39_220015300.1       GP-VAAGMLGGNVSAPHRLSFVDVMSGSVMRQIPKAHDDYITALTVLEESSYVLLSGGRD  790
CFAC1_240017100.1         AVAAAAVASSAAASSQNRLSFVDVAHGTVMRQIPKAHDDYITSLNVLEEGSYVLLSGGRD  825
rna_Lsey_0024_0300-1      AV------AAAAVASPHRLSFVDVAGGTVMRQIPKAHDDYITALNVLDDGSYVLLSGGRD  846
rna_LpyrH10_01_0970       A-------AAAAMASPHRLSFVDVACGTVLRQILSAHDDYITALNVLDEASYVLLSGGRD  852
BSAL_17930-t42_1          -----------ADQPQHQVALFNA-SGHETRVFARAHADCITALGVSADSDHLIITGSRD  653
rna_Baya_171_0030-1       -----------GAQLRHSLSLLDLEVGQVVRTFDDVHDDLTTVLSSCEDNEQVLISGSRD  503
Tb927.7.3040:mRNA         -----------EAGIQHNLVLLNGCAPQPLGVFEGAHEDYITVLRVGDESGHTLFTGSRD  488
TcIL3000_0_22630.1        -----------ESGLQHSLVLLNACTPQPSLVFDRAHEDYITTLCVGDEGGNTIFSGSRD  487
TvY486_0702880:mRNA       -----------ESGTQHDLVLINGYTGQTMGHFECAHSDHITVLRTAEDNDHMLYTGSRD  480
BCY84_17353-t36_1         -----------ESGLQHDLVLLNGHTSQFLRRFERVHDDYTTIIRAAEDGDRILFTGSRD  489
TRSC58_06257-t26_1        -----------EFGLQHDLVLLNGHTSQLLRRVEKAHDDCITVVRAAEDGDRILFTGSRD  486
DQ04_11821000-t26_1       -----------ETGLQHDLVLLNGHTTQTMRRFERVHDDYTTIVRVGEDSDHALFTGSRD  308
TM35_000192010-t36_1      -----------ESGLQHDLLLLNGHTAQPLRRFEHVHDDYTTVLRVGEEGEHVLLSGSRD  530
```



Supplementary Figure 6: Sequence analysis of Tb927.7.3040. Sequences of Tb927.7.3040 homologues in Kinetoplastida were compared using Clustal Omega. WD40 domains are marked in grey. The asterisk (*) marks fully conserved residue. Colon (:) marks conservation between amino acid groups with similar properties. Period (.) marks conservation between amino acid groups of weakly similar properties.

Supplementary 7




EMOLV88_070008900.1 MVELMNAALSLNGQLVDREDQHILQAMLIKALSTYPKTDKRHQDMLERAIAKFGHRKLEE 263
LMJLV39_070009600.1 MVELMNAALSLGGQLVDREDQHVLQAMLIQALSTYPKSDKRHQEMLEKAVAKFGHRRLED 267
CFAC1_080010400.1 MVELMNAALSLGGQLVDREDQHVLQAMLIKALSTYPKSDKRHQEMLERAIAKFGNRKLDD 269
rna_Lsey_0043_0180-1 MVELMNAALSLGGQLVDREDQHVLQAMLIKALSTYPKSDKRHQEMLEKAVAKFGHRRLDD 267
rna_LpyrH10_30_0970 MVELMNAALSLGGQLVDREDQHVLQAMLIKALSTYPKGDKRHQEMLEKAIAKFGHRKLDD 268
BSAL_67460-t42_1 IVDVMSAALSQGGQLVIRDDQRELQQVYISALKSFVKN-RKKLEMLELAQAKFKLSAGAL 252
TvY486_0800930:mRNA ------VSTNAGGRQV---------- 110
rna_Baya_070_0230-1 MVDLMSAALSQGGHLVEREDQHTLQAVYIRALKQYPRSNKKRHEMLEAAVAKFGHPSAAP 270
BCY84_12661-t36_1 MVDLMNAALSSGGQLVNREDQHTLQAVFIHALRSYPKSNKKRQEMLELAVSKFGSRPVAK 268
TRSC58_05642-t26_1 MVDLMNAALSSGGQLVNREDQHTLQAVFIQALRSYPKSNKKRQEMLELAVAKFGSRPVVK 270
TM35_000271170-t36_1 LVDLMNAALSSGGQLVNREDQHTLQAVFILALRSYPKSNKKRQEMLELAIAKFGNRQVPK 281
DQ04_01921100-t26_1 MVDLMNAALSSGGQLVNREDQHTLQAVFIQALRSYPKSNKKRQEMLELAIAKFGSRPVPK 269
TcIL3000_0_24110.1 MVDLMNAALSSGGQLVNREDQHTLQAVFIQALRSFPKSNKKRQEMLELAVAKFGTKPGPK 288
Tb927.8.1500:mRNA MVDLMNAALSSGGQLVDREDQHTLQAVFIQSLRSYPKSNKKRQEMLESAVAMFGSKPGPR 298
.: . .*: *
EMOLV88_070008900.1 EEEAAAVREV-AKKEAAKATSTSTSATAPTAT----------TPLLVGKQQPAESSATP 311
LMJLV39_070009600.1 EEEAAAVKREAAKKETATTASMATSVAAPTAA----------ATALKIRQEPTEAP-TP 315
CFAC1_080010400.1 EEEAAAKREA-ARREAAVAAAAAAATAAAAALTKETASTTGAPAAGNTKSADNGEATGAP 328
rna_Lsey_0043_0180-1 EEEAAAKREA-ARRETAAAVATAPVAEYSKNTAETSGATTAATTTAATAIADSGEVSGAP 326
rna_LpyrH10_30_0970 EEEAAAKREA-ARREAAAAAAAAPLAPDFTNTTCLSGATSTATNT--VTNANRGEASGAP 325
BSAL_67460-t42_1 PLAA----------QASAAAPAPPCIIDTPELP---------- 275
TvY486_0800930:mRNA ---------- 110
rna_Baya_070_0230-1 KDSS----------KSAATTPAHQ----------A----------GQS 288
BCY84_12661-t36_1 ITPA----------ATTAAAGQATTAANNNTNSNNTTNHNAT----------TNTAT 305
TRSC58_05642-t26_1 STPA----------STTAAGQVAAATNTPNTNHNGATNAAA----------AA 303
TM35_000271170-t36_1 NPTA----------TVATAANTSS----NNVNSTTTGHSPAT----------ASGAT 314
DQ04_01921100-t26_1 NVTA----------VAAATAAAAA---------- 283
TcIL3000_0_24110.1 STVV----------QTPANTTVAAAT---------- 304
Tb927.8.1500:mRNA STTA----------AVTAVEQTPPNNISGNSNSPPTTA---------- 326
EMOLV88_070008900.1 SAVAVT-------AATAVVSPINDSADLPAGEEAISARVSRSPAAATAPSKVEAATAAAT 364
LMJLV39_070009600.1 NAVTAT---ASAGSAVSAVAAINESVDPATGEEPTTVTASHVPAAATATAETAAAPAAQA 372
CFAC1_080010400.1 AAAEETDAAPLAGAA----------APTASSAA-----ANATETSAGA---------- 361
rna_Lsey_0043_0180-1 PAAKAAPAAVGGTSAAAKAPSTSAGAAAPTTPAT------ATATTATATAPVNSS---AS 377
rna_LpyrH10_30_0970 AAAQEPPAAAAAAAATTTAPPMQS-----TTPTAI-----AITAATTAPAPTNSS---ET 372
BSAL_67460-t42_1 ----------PALPVAPPKISKQEAA----- 291
TvY486_0800930:mRNA ----------ARAIEAA----- 117
rna_Baya_070_0230-1 GPATP----------HVVGSAQPA--------TAA----- 305
BCY84_12661-t36_1 TNSTPSHTATLGTAAAVVTAANSTNTATTNAVASVKAATTTIPT----TAAAAA------ 355
TRSC58_05642-t26_1 NANTPRHAGAAGTMAAASAA----------ASGKAGTATSPN----ALAATA------ 341
TM35_000271170-t36_1 TAAAVVAASTAAAAAANNNNNANNNTTATAATAAATATGATLPTTATTTTATTA------ 368
DQ04_01921100-t26_1 ----------GQAAANNNNSNNNNISNNNGNT--PRHGSTAA-VATATATTPG------ 323
TcIL3000_0_24110.1 ----------VAS------TAVSNNAAPAKVTAAAVVQGGAAPALKPVVRAAESA----- 343
Tb927.8.1500:mRNA ---------AVVTNATNNNNNNNNNSAASRAGSVAAATGGAVPGARTVARVVEAA----- 372
EMOLV88_070008900.1 AASGAQ---------TATTPAPGA-PSDAPINATAPRAP-----STTLRDA-AAMSTDSG 408
LMJLV39_070009600.1 AAGVAR----IASRSSFAAAAPGAAAADGSMRAAVPTVL-----PALVGDA-VAIATGTG 422
CFAC1_080010400.1 --ATATEAPGS-----TAAATPAKA-PSAEREAVPAAPPAPTAST--ARPP---HEEAAV 408
rna_Lsey_0043_0180-1 PAGNATEPPHAVAPAAAAVPMPSKVFADVESKALPPPAPL----SAEWRTPRETRDETSA 433
rna_LpyrH10_30_0970 ATTSAK---EYANAAAAAVTTPTTASTEADAKALPPPAPAPSLSSAERRTPRETREEATA 429
BSAL_67460-t42_1 -----K-------A-------DE-------SK---------- 297
TvY486_0800930:mRNA -----Q-------EDGSVQQHQS---------- 128
rna_Baya_070_0230-1 -----A-------DKSLALPTPES-----PSTAKP---------- 323
BCY84_12661-t36_1 ----------ALSA-AGGG-----GGN---------- 366
TRSC58_05642-t26_1 ----------AAAAPAGGG-----RSNSTS---------- 356
TM35_000271170-t36_1 ----------AAVVAAAGA-----QGKTAA---------- 383
DQ04_01921100-t26_1 ----------KTVVAAATS-----GG---------- 334
TcIL3000_0_24110.1 -----P-------EDGIIVQQQQQ-----HSRTQS---------- 361
Tb927.8.1500:mRNA -----Q-------DENVSVPQRS--------GSFS---------- 387
EMOLV88_070008900.1 SPAPTQVVGAAVPKANFPSVADTPSRAVDLTPDVAAATAAPSSATEAAMTFSTATQAH-- 466
LMJLV39_070009600.1 SAAAVQLLGIATSKATVSSIPSTPSRVAPSTPGA----------TPGPITFSTAIQVQ-- 470
CFAC1_080010400.1 SQTPSQAAAAATPTTTFSSAPGTPGTAVVTAA------VAPSGPFSTPAVFSTVATAALL 462
rna_Lsey_0043_0180-1 RATP-SLASATTPGVDFSSAPGTPGIAAAM--------AASMGPFNAPLVFSAAASSLPS 484

```
rna_LpyrH10_30_0970      G---------ATPKTAFSSAPGTPATAAVV--------IASTNLFNAPPLFSTSANTLPS  472
BSAL_67460-t42_1         ------------------------------------AASA---A-----TVA--------  305
TvY486_0800930:mRNA      ---------------------------------------------------NSSSVLA  135
rna_Baya_070_0230-1      -------NAVELPKKAAAASSGQAI----------SSAAAHSSSA---LASS--------  355
BCY84_12661-t36_1        ------------NNNVLASGGNPSKLSET-PEEPSTTLSHSGNY----H-SVNSNSSMV  407
TRSC58_05642-t26_1       ------------NTTLAGSGGIQPKMSGT-LEEGGAASAPTGAS---LG-ASSTSNNAA  398
TM35_000271170-t36_1     ------------TAAAAAGGGQSPKAKET-AEDDGIGNHH---------SSTSNNNNAA  420
DQ04_01921100-t26_1      ---------------NVVVVGGQPPRPIEASTEEEGTSHNHTNNY---HNNNTNSNTSTV  376
TcIL3000_0_24110.1       -------G-----MNTSSGSGSNGNYASVSNSNGSGTTAA---A-----VVA----AGNS  397
Tb927.8.1500:mRNA        -------N-----TNNSSGNSGNASYVNVCSAGAGATTAA---A-----VVAAAVAAAGA  427

EMOLV88_070008900.1      ---------QSAAASTW---------------ASVLSRQAG---APGTPDTTVSASAVPH  499
LMJLV39_070009600.1      ---------QHNVMPVP---------------INALAGQGS---AMGALGLQATTAAATP  503
CFAC1_080010400.1        PTPSMSQPAVSPAATTPIFTTAATPAFATAAPVAALGRPPALPPTATTPGTTT-------  515
rna_Lsey_0043_0180-1     TQPTLMQASPLPIATTQASAGAALAAVTTPLLSAAAARPPSQPSAVPAAGS---------  535
rna_LpyrH10_30_0970      PQQTLTQISPLATISTNAAAA---------PASAVAAVRPPSQPVAISTAANTAVFLAVPP  524
BSAL_67460-t42_1         ---------------TAATTG----KKNDAAPV----------------------------  319
TvY486_0800930:mRNA      -----------SSRRSQVSPG----TTSNPSSA-QRAKDSNLS---------------TM  164
rna_Baya_070_0230-1      --------------------AG----K-------------------------------SA  360
BCY84_12661-t36_1        -----------SNRRPQAVRG----NS-TPRAVSTRNKDSSST---------------SP  436
TRSC58_05642-t26_1       -----------SMRRPQIAGG----TS-TLTTASTRNREASST---------------SP  427
TM35_000271170-t36_1     -----------PTRRPQAVGG----TPTTPFTASARLKEMANT---------------SP  450
DQ04_01921100-t26_1      -----------STRRPQALGG----AA-TPLATNTRNKDSTGG---------------SS  405
TcIL3000_0_24110.1       -----------STRRTHPTSG----TVSPSAPAAARAKESVA-----------------  424
Tb927.8.1500:mRNA        -----------PSRRPQGTTG----STTPVTPVSTRSKEPCVP---------------TT  457

EMOLV88_070008900.1      S----------------VPPLTQPRLATVGISRLHPYLAGRAAKMQDSAPRLLPLFD--P  541
LMJLV39_070009600.1      A----------------ASTPAPARTAIVGISKLHPYLAGRAANMRDEVPGLSPIYD--P  545
CFAC1_080010400.1        ----TTTTTGSTPSTATTTATSTGAGLIVGVSKALQHLREKPDSY--WRPRLWKPKS--P  567
rna_Lsey_0043_0180-1     ------------TATTTTTAAAAAVGPIVGISKVASALRMYPDRY--DKSTLWSAPD--K  579
rna_LpyrH10_30_0970      GAQPPPLQQQQQQPCPSHIQAPAATGPIIGISKIFPTLRKYASRF--DRLSLWTFPP--T  580
BSAL_67460-t42_1         -------------------TTQQANKRVIAISKRALSLPMPTEPR------KYAPLHRVD  354
TvY486_0800930:mRNA      VT-----ASSNNSAAGVPAQIVKAQQPIVGISQLAQM-AKQVAPG------HYDPVR--T  210
rna_Baya_070_0230-1      GLAP-----------------RTVSAGVSKIVPLQLA-KSAAPAL------AYDPVR--T  394
BCY84_12661-t36_1        AKQQP--LVQESPQRQIPP--PPPPQPVVGISRVAQQ-VQQIPPR------HYEPIR--T  483
TRSC58_05642-t26_1       ARQPS--TAQESLQRQMPTQPQPQPQPVVGISRVAQQ-VRQAPAL------QYEPIR--T  476
TM35_000271170-t36_1     AKQQT--QT-----------QAPPPQPVVGISRLAMQ-SKQAPIR------HYEPVR--T  488
DQ04_01921100-t26_1      AKQQQ--QQ-----------QQQQQQPVIGISRLALQ-AKKAPAR------HYDPIR--T  443
TcIL3000_0_24110.1       -------VMPP--------SRQQHLAPIAGISQLVQP-SKQAVQG------RYDPIR--T  460
Tb927.8.1500:mRNA        TIAPTT-TVQPTRQQQQQQQQQRPPPPIVGISQLVQQ-PRQPAAC------RYEPIR--T  507
                                                   :

EMOLV88_070008900.1      RSHTSGLFICGGLFSVDSAAWSSSPLGDAALNRLVDTPSRD-HVRREEG-----------  589
LMJLV39_070009600.1      KSHSSGLFICGGQFTVEPLVWNYAPLGDAAANRRLDSPTRE-YMRIKEE-----------  593
CFAC1_080010400.1        DDRLPDLYLCGGQFTVKSHLWNPTPLGDATPLPLMDKATEN-LIEYLST-----------  615
rna_Lsey_0043_0180-1     VGPGPNLYLCGGQFSVKSQLWGSTPLGEAARQPAVDAATKR-LVKYITS-----------  627
rna_LpyrH10_30_0970      DSTAPDQYLCGGQLSVKSHLWGPTPLGEAAPQLPMDGVTQR-LAKYIAE-----------  628
BSAL_67460-t42_1         DDILQDVLILGGLLSLSSKSSCLLGFAKKSPAASTSPSDGNSPLRLLAEH--GPSILDAG  412
TvY486_0800930:mRNA      DLPLPSIALLGGLLALTPQSATWALLGLATSHKYLH-GIQSCQM----------------  253
rna_Baya_070_0230-1      DQPLPNIFLLGGLLALSPAGAAAAPFGLASTAQWTN-SGAT----------------------  434
BCY84_12661-t36_1        DQPLPNVVLFGGLFGLTPQSASWALLGRASCFNSGS-AASNAPVLAA--------------  529
TRSC58_05642-t26_1       DLPLPNVVLFGGLLGLTPQSASWALLGAASSFKSTT-AMSTSP---A--------------  519
TM35_000271170-t36_1     DLPLPNIALFGGLFGLTVQSATWALLGSATSHKTTS-LVTNTPSAVAAPSMTTATVVSGG  547
DQ04_01921100-t26_1      DLPLPNIALFGGLFGLTPQAASWALLGRATCQKPTP-SASNTPVATAAAVAA-------A  495
TcIL3000_0_24110.1       DLPLPNITVFGGLFGLTPQSATWAPLGSALSRKAPL-ATQSSVV----------------  503
Tb927.8.1500:mRNA        DLPLPNITLFGGLFGLTTQSAAWAPLGAATSRKTSP-PSLPGVV----------------  550
                              .  : ** : :          :.

EMOLV88_070008900.1      ----------------------------VLHWPTPLKGEQMLHQASAADGKPL-------S  615
LMJLV39_070009600.1      ----------------------------NRQQRREAREAQAARLASIAAGNLA-------S  619
CFAC1_080010400.1        ----------------------------AEAVMA----AADARPREEPER-----------  633
rna_Lsey_0043_0180-1     ----------------------------LRELNA----RETMRVAEGDSESLLPP------  650
rna_LpyrH10_30_0970      ----------------------------TRRDES----EAAARVAEHASLPVVSAGDAGRN  657
BSAL_67460-t42_1         NEQHSTRRNAPLDDSKSSCLLGFAKKSPAASTSPSDGNSPLRLLAEHGPSILDAGNEQHS  472
TvY486_0800930:mRNA      ----------------------------PPQAAVSLGVSPGDISREADDNTF---------  277
rna_Baya_070_0230-1      -------------------------ADTVRPPPPEADGSAY------FADTVF---------  456
BCY84_12661-t36_1        ---------------------------------PSCGLSFVEAAGEADDKVL---------  548
```

Supplementary Figure 7: Sequence analysis of Tb927.8.1500. Sequences of Tb927.8.1500 homologues in Kinetoplastida were compared using Clustal Omega. Low complexity regions are marked in grey. The asterisk (*) marks fully conserved residue. Colon (:) marks conservation between amino acid groups with similar properties. Period (.) marks conservation between amino acid groups of weakly similar properties.

Supplementary 8

Supplementary Figure 8: Tb927.11.4900 is a putative G-protein. Sequence of Tb927.11.4900 with the G-domains in the circular permutation of G4-G5-G1-G2-G3 and the WD40-domains.