Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by
M.Tech Chaitali Chakraborty
born in: Durgapur, India

Oral-examination: October 18th, 2017

Interactions of the CAF1-NOT complex
And
The role of ZC3H30 in combating stress in *Trypanosoma brucei*

Referees: Prof. Dr. Christine Clayton (ZMBH)

Prof. Dr. Luise Krauth-Siegel (BZH)

Interactions of the CAF1-NOT complex
And
The role of ZC3H30 in combating stress in *Trypanosoma brucei*

Doctoral dissertation by
Chaitali Chakraborty
Zentrum für Molekular Biologie
Universität Heidelberg
Heidelberg, Germany

Dedicated to the Creator and the women who have inspired me with their lives.

These are the few quotes that have shaped my life:
"Science, for me, gives a partial explanation for life. In so far as it goes, it is based on fact, experience and experiment."
Rosalind Franklin in a letter to her father Ellis Franklin, 1940, Excerpted from Brenda Maddox's ' The Dark Lady of DNA', 2002.

"We must not forget that when radium was discovered no one knew that it would prove useful in hospitals. The work was one of pure science. And this is a proof that scientific work must not be considered from the point of view of the direct usefulness of it. It must be done for itself, for the beauty of science, and then there is always the chance that a scientific discovery may become like the radium a benefit for mankind."

Marie Sklodowska Curie at a lecture in Vassar College, Poughkeepsie, New York (14th May, 1921), in Cambridge editorial Partnership, 'Speeches that Changed the World'.

Things are much more marvellous than the scientific method allows us to conceive.
Barabara McClintock; quoted by Evelyn Fox Keller. 1983. *A feeling for the Organism. The Life and Work of Barbara McClintock.* New York., Freeman, p203.

Table of Contents

Acknowledgements

I would like to extend my gratefulness to:

Prof. Christine Clayton for providing me this opportunity to work with her in the lab. It was under her supporting supervision that I learnt and grew as a scientist. Now that I have finished my thesis, I take with me not only the pair of critical eyes for scrutinizing data and a questioning mind that doubts ideas, but also, the philosophy of science and my commitment to it. I was more than happy to work with her, as a lady who had dedicated her life to study the neglected disease of the tropics, and have contributed enormously for the scientific development to understand the biology of the parasite causing the disease.

Prof. Luise Krauth-Siegel for her suggestions, ideas and critical comments, which helped me to question, doubt and think on the experimental aspects of my project; like what was I doing, why was I doing it, does it ask any question, will it contribute to better understanding the phenomenon I observed.

Prof. Georg Stoecklin for the support and critical suggestions during the TACs. During the RNA meetings, it was a thrill to listen to Prof. Clayton and Prof. Stöcklin, since I always learnt something new in every meeting.

I would like to extend my special gratitude to **Dr. Igor Minia, Dr. Elisha Mugo, Dr. Esteban Erben** and **Dr. Monica Terrao**. They were the ones who trained me, and the people whom I pestered with innumerable questions, and who never got bothered. They tolerated my naivity towards science and helped me develop ideas; critically question them and calmed me when I was upset.

Ute Leibfried and Claudia Helbig who offered technical assistance that has been making our work in the lab easier and faster.

Dr. Susanne Kramer and **Dr. Carine Goos** for sharing their plasmid (Tb927.8.3820-YFP) and their antibodies (α-Dhh1, α-PFR, α-SCD6), and for answering to all my emails, regarding stress granule enrichment protocol, and the outcome of it. **Prof. Elmar Schiebel** for allowing me to use his incubators for the yeast-two-hybrid screens, and for letting me use the GBP-beads for the co-immunoprecipitation pull down experiment.

Prof. Thomas Rupert, Dr. Sabine Merker, and **Dr. Bernd Heißling,** from the Mass-Spectrometry facility from ZMBH for processing and data searching for my mass-spectrometry samples in the, and **David Ibberson** for processing and analysing my deep-sequencing DNA samples and RNA-Seq samples.

Dr. Abeer Fadda, Clementine Mercé, and **Kevin Leiss,** for analysis of NGS data, building pipelines that made it easier to analyse deep sequencing and RNA-Seq data for us biologists. Also, they took great care to explain the rationale behind the codes

they had written and discussed with me as to what kind of analysis would answer my biological question best.

My parents, friends in Heidelberg and at home in India. Thank you for all the moral support, love and time. Without you it would have been extremely tough.

Summary

The trypanosomatids are a group of eukaryotic parasites (*Trypanosoma brucei, Trypanosoma cruzi and Leishmania sp*.), which cause diseases affecting humans and animals. Since most genes do not have individual promoters and regulating transcription factors, the mode of gene expression regulation is post-transcriptional: almost entirely determined by RNA binding proteins, stabilizing complexes and the mRNA decay machinery. In my PhD, I worked on two projects in *Trypanosoma brucei;* a screen to identify the genome wide interactors of the CAF1-NOT complex and the role of the RNA binding protein ZC3H30 in stress.

The decay-promoting deadenylation machinery, the CAF1-NOT complex, degrades poly (A) tails, an event that marks mRNA for degradation. I did a genome-wide yeast two-hybrid screen to discover novel proteins that might influence CAF1-NOT complex function in the parasite. I looked at the total pool of genes obtained after screening, and assessed the protein class and biological process of the candidates obtained. To assess interaction specificity I compared the yeast-two-hybrid screen data with those from other trypanosome proteins.

African trypanosomes undergo stress in their insect host in the form of heat shock or in their human host from the immune response, or during fever or intake of drugs. The parasites have developed differential gene expression to cope with stress using RNA binding proteins. The second part of the manuscript focuses on a zinc finger RNA binding protein ZC3H30, which is inessential for the growth of trypanosomes under normal growth conditions. However, conditions like heat shock, ER stress, oxidative stress and starvation stress, are lethal for trypanosomes that lack ZC3H30. The absence of ZC3H30 has no effect on stress markers, except that major cytoplasmic *HSP70* mRNA is more abundant. I have shown that ectopically expressed ZC3H30, when tethered to reporter mRNA, can reduce its transcript and encoded protein abundance. ZC3H30 is a *bona fide* component of the stress granule; in presence of heat shock, ZC3H30 protein re-localises to stress granules from the cytoplasm. It pulls down another protein, which also migrates to stress granules upon heat shock and might also have a role in stress management for trypanosomes. By transcriptome profiling, procyclic cells lacking ZC3H30 showed no differences from wild type in transcript abundance, and specific targets were not found. However, procyclic form trypanosomes lacking ZC3H30 have a higher overall translation rate than wild type cells. Possibly ZC3H30 affects the mis-reading of mRNAs during translation, since the knockouts of ZC3H30 are less sensitive than wild type to aminoglycoside translation inhibitors like Hygromycin. Nevertheless, I have not been able to identify definitively the role of ZC3H30 in stress resistance.

Zusammenfassung

Die Trypanosomatidae sind eine Gruppe von eukaryotischen Parasiten (*Trypanosoma brucei, Trypanosoma cruzi und Leishmania sp.*), die Krankheiten verursachen, welche Mensch und Tier betreffen. Da die Parasiten keine einzelnen Promotoren oder regulierende Transkriptionsfaktoren besitzen, ist die Art der Genexpressionsregulation posttranskriptionell; sie wird fast vollständig durch RNA-bindende Proteine, stabilisierende Komplexe und die mRNA-Abbau-Maschinerie bestimmt. Während meines Doktorats habe ich zwei Projekte über *Trypanosoma brucei* bearbeitet; die genomweiten Interaktionspartner des CAF1-NOT-Komplexes und die Rolle des RNA-bindenden Proteins ZC3H30 während zellulärem Stress.

Die Zerfall-fördernde Deadenylierungsmaschinerie, auch CAF1-NOT-Komplex, baut Poly (A)-Schwänze ab, ein Prozess, der mRNAs für den Abbau markiert. Ich habe einen genomweiten Hefe-Zwei-Hybrid-System-basierten Screen durchgeführt, um neue Proteine zu identifizieren, die die Funktion des CAF1-NOT-Komplexes des Parasiten beeinflussen könnten. Ich habe die Gesamtheit der Gene, die durch das Screening identifiziert wurden, betrachtet und habe die Proteinklasse und den biologischen Prozess der erhaltenen Kandidaten bestimmt. Um die Interaktionsspezifität zu beurteilen, habe ich die Hefe-Zwei-Hybrid-Sytem-Daten mit derer anderer Trypanosomen-Proteinen verglichen.

Afrikanische Trypanosomen unterliegen Stress in ihrem Insektenwirt in Form von Hitzeschock oder in ihrem menschlichen Wirt durch die Immunantwort, während Fieber oder während der Einnahme von Arzneimitteln. Die Parasiten haben eine differentielle Genexpression entwickelt, um mithilfe von RNA-bindenden Proteinen mit Stress fertig zu werden. Der zweite Teil des Manuskripts konzentriert sich auf das RNA-bindende Zinkfinger-Protein ZC3H30, das für das Wachstum von Trypanosomen unter normalen Wachstumsbedingungen irrelevant ist. Unter Bedingungen wie Hitzeschock, ER Stress, oxidativem Stress und Hunger-Stress, ist das Fehlen von ZC3H30 für Trypanosomen tödlich. Die Abwesenheit von ZC3H30 hat keine Auswirkung auf Stressmarker, mit der Ausnahme, dass die zytoplasmatische *HSP70* mRNA häufiger vorkommt. ZC3H30 ist ein *bona fide* Bestandteil von Stressgranules; in Gegenwart von Hitzeschock verteilt sich das ZC3H30 Protein vom Zytoplasma zu den Stressgranules um. Es bindet an ein anderes hypothetisches Protein, das unter Hitzeschock auch zu Stressgranules wandert und ebenfalls eine Rolle im Stressmanagement von Trypanosomen spielen könnte. Anhand von Transkriptom-Profiling konnten keine Unterschiede zwischen prozyklischen Zellen ohne ZC3H30 und Wildtyp-Zellen in Bezug auf Transkript-Häufigkeit festgestellt werden und spezifische Ziel-RNAs konnten nicht identifiziert werden. Prozyklische Trypanosomen ohne ZC3H30 haben eine allgemein höhere Translationsrate als prozyklische Wildtyp-Zellen. Möglicherweise regelt ZC3H30 das falsche Ablesen von mRNAs während der Translation, da die Knockouts von ZC3H30 weniger empfindlich gegenüber Aminoglykosid-Translationsinhibitoren wie Hygromycin sind. Nichtsdestotrotz war ich nicht in der Lage, die Rolle von ZC3H30 in der Vermittlung von Stress-Resistenz endgültig zu bestimmen.

Abbreviations

AD	GAL4 transcription factor's activating domain
APOL1	Apolipoprotein 1
ALPH1	Apa H like phosphatase
AUF1	AU binding factor 1
3-AT	3-amino triazole
BD	GAL4 transcription factor's DNA binding domain
BLA^R	Blasticidin resistance
BF/BS	Bloodstream form parasite
CAF1	Ccr4 associated factor 1
CAT	Chroramphenical acetyl transferase gene
CATT	Card agglutination test for trypanosomiasis
Ccr4/CCR4	Carbon catabolite repression 4 homolog
CAF40	Ccr4 associated factor 40
CDS	Coding DNA sequence
Cy5	Indodicarbocyanine/ Cyanine5
CSF	Cerebrospinal fluid
DAPI	4',6-diamidino-2-phenylindole
DKO	Double knockout
DKO-ectopic	Double knockout expressing ectopic copy
DTT	Dithiothreitol
DNAj	HSP40
Dhh1	DExD/H box ATP dependent RNA helicase
E	Eluate
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis-(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
EP	Glutamic acid (E), Proline (P) repeat protein (Procyclin)
ESAG	Expression site associated gene
eIF	Eukaryotic initiation factor(s)
FCS	Fetal Calf Serum
Fig	Figure
G	Guanine or granules
GFP	Green Fluorescent protein
HAT	Human African Trypanosomiasis
HMI-9	Hirumi's Modified Iscove's Medium-9
HSP	Heat Shock Protein
HSC70	Heat Shock protein 70 (constitutive)
H_2O_2	Hydrogen Peroxide
Hyg	Hygromycin
I	Input
IgG	Immunoglobulin G
Kb	Kilo base
KDa/KD	Kilo Dalton
LSM	Sm-like protein
MEM	Minimal Essential Medium of Eagle
mRNA	messenger RNA
MKT1	Maintenance of killer toxin 1
mM	milli Molar
mRNP	messenger ribonuclear protein
NB	Northern blot
NMD	Non sense mediated decay
NOT	Negative on TATA box
nt	nucleotide

PAN	poly (A) nuclease
PARN	poly (A)- specific ribonuclease
PABP	poly A binding protein
P-body	Processing body
PBS	Phosphate Bufder saline
poly (A)	poly Adenosine
poly Q	poly Glutamine
PTC	pre termination codon
PUF	Pumillo and FBF domain
PURO^{R}	Puromycin resitance
RBP	RNA binding protein
RNAi	RNA interferance
RNasin	RNase inhibiting protein/ RNase inhibitors
rpm	revolutions per minute
RRM	RNA recognition motif
S/SN/Sup	Supernatant
SCD6	Supressor of Clathrin deficiency
SL	Splice leader
TCA	Trichloroacetic acid
tet	tetracycline
T/Tot	Total
TR	Trypanothione reductase
TxNPx	Thioredoxin
Un	Unbound
UTR	Untranslated region
UV	Ultra violet
V	Volt
VSG	Variant surface glycoprotein
W	Wash
WB	Western blot
XRN	5'-3' exorionuclease
YTH	Yeast-two-hybrid
ZC3H30/20/41/(other numbers)	Zinc finger domain protein 30/29/41(other numbers)

Chapter I- Introduction

Trypanosoma brucei and African sleeping sickness

Discovery

In 1792 Thomas Winterbottom was appointed as physician to the Sierra Leone Company. From now until the next 4 years Winterbottom would spend his time in Africa. At the end of his stay, he would return to Europe and write an account of his stay and his medical perspectives of the diseases in sub-Saharan Africa. The book 'An account of the native Africans in the neighbourhood of Sierra Leone' was published first in 1803. This was the first time that African sleeping sickness was described. He observed that slave traders avoided slaves with swelling of the neck. This symptom of cervical lymphadenopathy and lethargy of the natives became his eponymous sign: "Winterbottom's sign". Another century passed with no progress on the science of the disease. Later in 1901 it was Everett Dutton who first identified the human parasite *Trypanosome gambiense* while examining blood of a patient who could not be cured of malaria. Although he and his colleagues published extensively on trypanosomes in the Memoirs of the Liverpool School, his unexpected death at the age of 29 proved tragic for scientific world working on this newly discovered parasite [1].

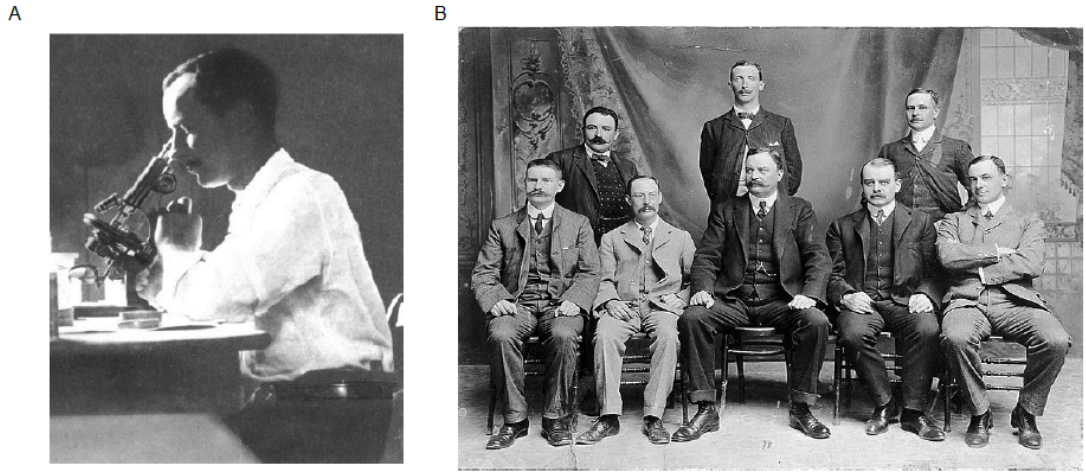


Figure 1.1 Pioneers of the human African trypanosomiasis. (A). Joseph Everett Dutton [1] from 'The Liverpool School', who identified the human form of trypanosome *Trypanosoma gambiense.* **(B).** Sir David Bruce (centre) with the Mediterranean Fever Commission (Wikipedia).

Fortunately for us David Bruce accepted the request made by the British Government to investigate the disease by joining the Sleeping Sickness Commission. Bruce was already famous for investigating and connecting the bacterial Gram-negative organism *Brucella* sp. (named in his honour) with Malta fever (now known as Brucellosis) in 1887. He now successfully established the vector of sleeping sickness as the tsetse fly, and found animal African

trypanosomiasis (Nagana). On the 29th of May 1903, the Sleeping Sickness Commission sent a report to the Royal Society (Bruce et al., 1903), which made the following conclusions:

- 'sleeping sickness is caused by ….. a species of trypanosomes'
- 'this species is probably that described by Dutton from the West Coast of Africa and called by him *Trypanosoma gambiense*'
- 'trypanosomes are transmitted by *Glossina palpalis* and by it alone'.

Epidemiology

It has been estimated that 60 million people are exposed to human African trypanosomiasis (HAT), in nearly 200 separate active foci in sub Saharan Africa. But only 4-5 million people are under surveillance. *T. b. rhodesiense* is responsible for the zoonotic HAT disease prevalent in Eastern and Southern Africa. *T. b. rhodensiense* is transmitted usually by the *Glossina moristans* and *G. fuscipes.* The insect vector for *T. b. gambiense* is *G. palpalis* and *G. fuscipes.* Figure 1.2 shows a map of endemic foci of *T. b. gambiense* and *T. b. rhodesiense*. Between the period 2000 and 2009 a total 2.2 million people were screened. A total of about 42,000 HAT cases were identified from over 6000 different locations [2].
Most countries in sub-Saharan Africa have been able to control epidemics of HAT. With the control of the vector, the chances of infection and loss of life should reduce with time, unless control is disturbed by war. Each outbreak distresses the community with life and economic losses. In a HAT outbreak in Buma in 2002 in the Democratic Republic of Congo, 57 individuals were found infected by *T.b.gambiense* [3]. A serious issue is the use of humans as a reservoir by the parasites. Recently there have been several incidences identified where the patients were either PCR positive for the 177bp satellite of Trypanosomes or possessing antibodies against the LiTaT1.3 (VSG) antigen by CATT. But the parasite could not be identified microscopically [4-6]. Also in a 15-year-long follow up study, it was found that the patients had gradual reduction in antibody titres (against VSGs) with time and was also PCR negative. This was one of studies that demonstrated that humans can also show trypanotolerance [7]. These studies support the notion that trypanosomes can use humans as a long-lasting reservoir, which can give rise to periodic resurgences of HAT in endemic foci. Recently the demonstration of extravascular parasite in skin of infected mice [8] is urging the need to look for novel reservoirs for the parasite in human tissues.

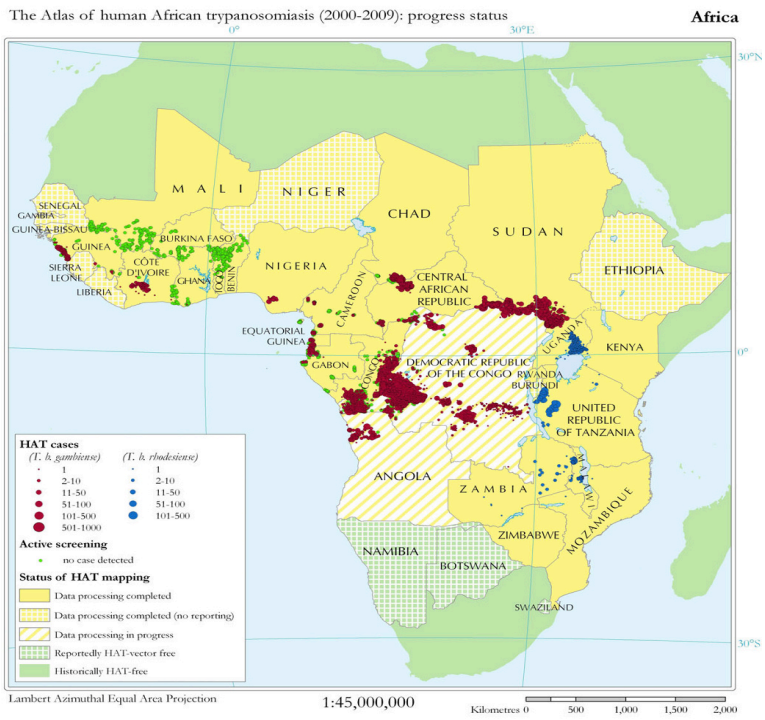


Figure 1.2: The Atlas of HAT during 200-2009 (taken from [2])

Cause of the disease

There are three subspecies of *Trypanosoma brucei*, and two of them are responsible for the disease in humans. *T. brucei brucei* infects cattle, but not humans. *T. brucei gambiense* causes a chronic sleeping sickness in central and western Africa, and *T. brucei rhodesiense* causes an acute disease in east and southern Africa. The disease is classified into 2 different stages depending on the presence of the parasite in the cerebrospinal fluid.

Symptoms of stage 1 (no parasites in the cerebrospinal fluid) include-

1. Appearance of a chancre, usually painless after 5-15 days of being bitten [9]. The chancre is usually absent at the time of diagnosis in African patients, who present after more than two weeks. In case of the European patients, infected especially by *T. brucei rhodesiense*, there is usually a high prevalence of chancre [10].
2. Intermittent fever usually 3 weeks after the bite.
3. Lymphoadenopathy, a characteristic sign of trypanosomiasis caused by *T. brucei gambiense*.
4. Facial edema.
5. Transient urticarial, erythematous or macular rashes appearing 6-8 weeks after the onset of the disease.
6. Trypanids (pale, blotchy, edematous erythematous macules all over the trunk).

Symptoms of stage 2 (late or neurologic state)-

1. Persistent headaches.

2. Daytime lethargy and sleepiness (eponymous sign), followed by insomnia in the night. There are alterations in the circadian fluctuations of cortisol, prolactin, growth hormone and plasma rennin [11]
3. Behavioural changes; mood swings and depression.
4. Appetite loss, wasting syndrome, weight loss.
5. Children undergo seizures (rare in adults).

To survive in their human host trypanosomes must escape host immunity. Each trypanosome expresses a highly antigenic variable surface glycoprotein, called the VSG (Variant surface glycoprotein). VSGs are developmentally regulated dense surface proteins, which represent at least 10% of the total protein in bloodstream trypanosomes. Antibodies raised by the host immune system against this protein results in complement mediated lysis of the parasite. This however does not clear infection, because some parasites have turned on a new VSG surface protein. Periodic destruction of the abundant antigenic variant of the parasite followed by outgrowth of a different antigenically different parasite gives cyclical waves of parasitemia [12, 13].
Besides VSG, the parasite has developed species-specific defense mechanisms. For example, the human infective *T. brucei rhodesiense* express a serum resistance associated protein, which confers resistance to lysis of the trypanosomes by human apolipoprotein L1. *T. b. gambiense* escapes lysis by a different mechanism using a specific glycoprotein with β-sheets, and an inactivating mutation of the haptoglobin-haemoglobin receptor (which is needed for uptake of relevant lipoprotein complexes) [14].
Laboratory diagnostics of the disease currently include three stages; screening, diagnostic confirmation and staging. Screening tests include Card Agglutination Test for Trypanosomiasis (CATT/ *T. b. gambiense*) and cervical lymph node palpation and puncture [3]. CATT depends on the agglutination of LiTat1.3 (VSG) antigen [15]. Although it is a very effective and widely used method for identification of trypanosome-infected individuals, patients infected by trypanosomes not expressing this antigen usually escape the screening. A newer development to this method is the *LATEX/T.b.gambiense,* which uses supended LATEX coupled with LiTat1.3, 1.5 and 1.6 (VSG proteins).
After screening, parasite detection is done microscopically, usually from the aspirate obtained from the chancre, lymph node and blood. Like microfilarial parasites, Trypanosomes can also be detected from wet, or thick dried blood films. Other methods include Quantitative buffy coat, where the blood from the patient is centrifuged to concentrate the parasites, and their nucleus and kinetoplasts are stained by Acridine orange [16]. Another very interesting and clever way to detect trypanosomes is the Mini-anion-exchange centrifugation technique. The mEACT is based on seelctive adherence of negatively charged blood cells to a positively charged column, allowing trypanosome purification [17, 18].
Finally, the stage of the disease is determined according to the presence of the parasite in the cerebrospinal fluid (CSF). Usually white blood cells are counted, and second stage treatment is prescribed to those patients with a minimum threshold of 5-20 cells/μl of the CSF [19], [15]. Also, elevated concentrations of protein or IgM in the CSF are signs of second stage trypanosomiasis [19, 20] [21] [15].

The Trypanosome cell

Kinetoplastids are protozoan parasites that probably first arrived in the earth ecosystem 300 million years ago [22]. They belong to the superphylum Excavata, phylum Euglenozoa and are grouped in the class Kinetoplastida. They are microscopic unicellular protozoan parasites that are found in humans, insects, birds, amphibians, fish and plants. They belong to the family Kinetoplastidae. The trypanosome cell is elongated and has a polarized microtubule cytoskeleton [23]. The microtubule minus ends are anterior and the plus ends are posterior [24]. The flagellum, mitochondrion, kinetoplast and flagellar pocket are present as single copies in G1-phase cells and are positioned within the cytoskeletal corset between the posterior end and the centre of the cell. The flagellar pocket is the most posterior structure, and is the only site for endo- and exocytosis [20]. This site is extremely important in the bloodstream form, as GPI-anchored VSG recycling and other receptor-mediated endocytosis occurs there ([25]). Depending on the life cycle stage of the parasite, the parasite can vary from 1.3-3.5μm in diameter and or 8-50μm in length [26]. The mitochondrion is a long elongated structure, which runs from the posterior to the anterior end of the parasite. The absence of cristae in the bloodstream form mitochondrion correlates with the absence of mitochondrial respiration in this stage; generation of energy is by glycolytic reactions that mostly occur in specialized compartments called glycosomes [12]. In the procyclic form the mitochondrion is a highly active organelle. The mitochondrion contains a unique catenated DNA called the kDNA, which consist of maxicircles and minicircles [27]. The maxicircles (~ 50 copies/kinetoplast) encode mitochondrial proteins and the minicircles (~ 10,000 copies/kinetoplast) encodes guide RNAs, which are involved in editing of mitochondrial mRNA [28, 29]. The most obvious morphological difference between parasites of different stages is in the positioning of the nucleus and the kinetoplast. The kinetoplast is at the far posterior end in the bloodstream form, whereas in the procyclic form it lies midway between the cell nucleus and the posterior end. In the epimastigote the kinetoplast lies anterior to the cell nucleus [30]. The nuclear genome consists of 11 megabasepair chromosomes, and minichromosomes that are usually about 50-100 Kb in size. The genes in the chromosomes are organised in polycistrionic arrays [31]. The minichromosomes are usually silent DNA, except that some have VSG genes at the subtelomeric regions. (Only kinetoplastids that undergo antigenic variation by VSG switching contain minichromosomes [32, 33].)

Evolution of Trypanosomes

Trypanosomes are parasites infecting all metazoan living organisms, from aquatic animals like fishes, leeches, flies to man. One popular hypothesis suggests that the early form of the parasite infested the gut of aquatic vertebrates, subsequently occupied the gut of haematophagus insects and this was followed by transmission to other vertebrates including man [34]. In the other hypothesis it is believed that the trypanosomes evolved in the haematophagus insect and were subsequently transmitted to other hosts while the fly was feeding. 18S rRNA sequences were used to construct the first phylogenetic tree of the trypanosomes [34]. This tree indicated

that the human infective *T.brucei* had diverged before the evolution of *T.cruzi*; and the *Leishmania spp.* and the *Leptomonas spp*. were the last to evolve. This model lacked the root organism from which trypanosomes and bodonids evolved. Therefore models using comparison of the kinetoplast DNA, heat shock proteins, and the segregation of chromosomes during replication were used later to understand the evolution of trypanosomes. One study group combined small subunit rRNA analyses with HSP90 analyses and demonstrated that bodonids and kinetoplastids are classified into 4 main clades, and trypanosomatids occupy the 4^{th} clade [35]. Although this study provided robust evidence that the trypanosomatids were a close relative of bodonids, but it failed to assign all kinetoplastids a position in the final phylogenetic tree. In another study, the maxicircle genome of *Trypanosoma lewsi* was compared to that of *T. brucei*, *T. cruzi*, *T. rangeli* and *Leishmania tarentolae*. It was found that T. lewsi maxicircle genome was similar to that of *T. brucei*, *T.cruzi* and *T. rangeli*, and they shared similar editing partners, and these editing partners were different from *Leishmania tarentolae* [36]. In spite of these numerous attempts to study evolution, the origin of the trypanosomatids is still not clear. Recently, there have been suggestions for using minicircle conformation as a tool to study the evolution of trypanosomatids [35]. Figure 1.3 shows the phylogenetic tree constructed for the parasites trypanosomatidae and bodonidae based on 18S rRNA homology and chromosome segregation proteins.

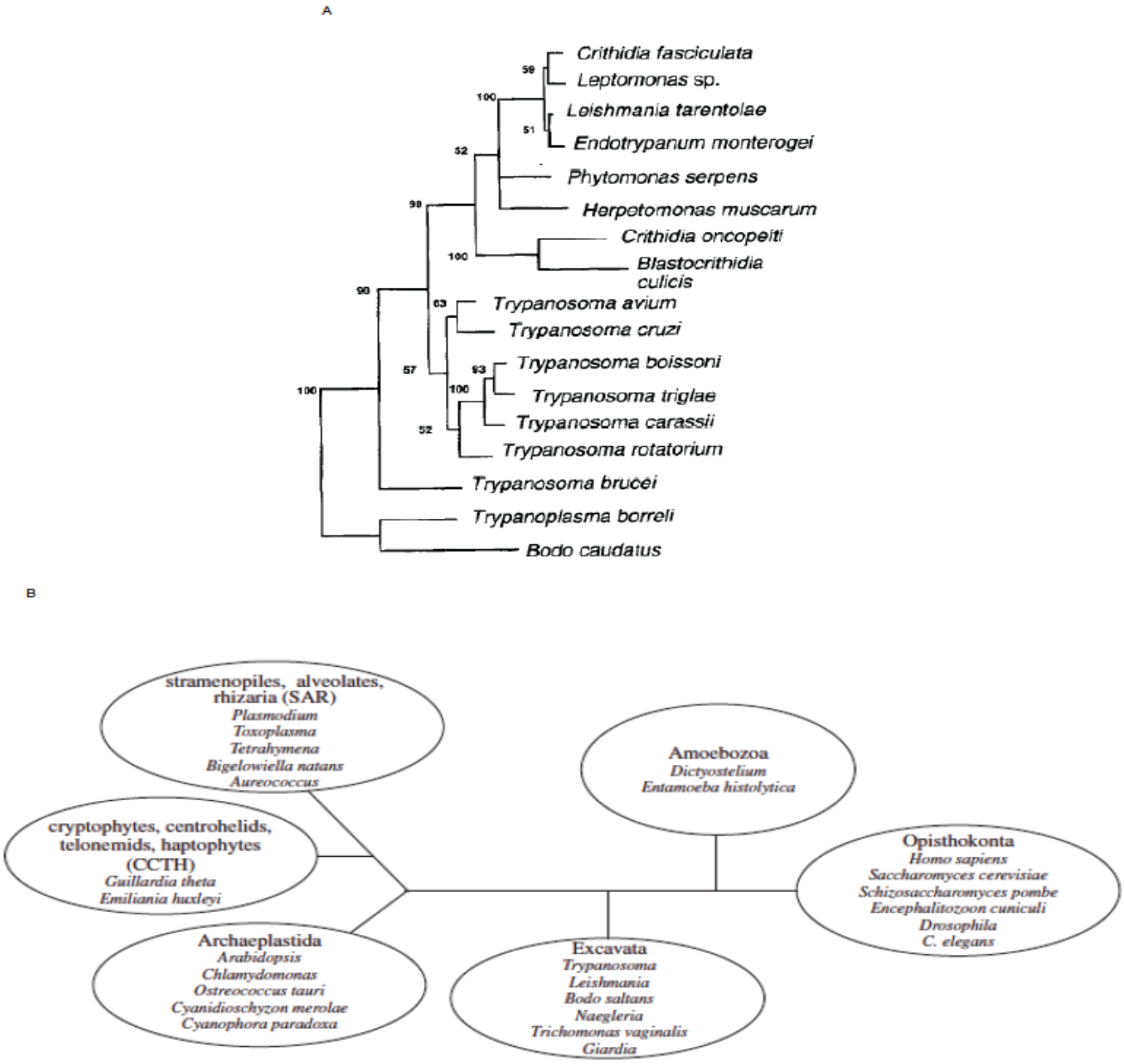


Figure 1.3: Unrooted phylogenetic trees. (A). A kinetoplastid phylogenetic tree constructed from small subunit rRNA (18S rRNA) dataset [34]. (B). Current phylogenetic tree constructed from genome sequences for selective organisms [37]

Life cycle of *Trypanosoma brucei*

Figure 1.4 depicts the life cycle of the parasite. In the blood of the mammals the parasite lives either as a slender cell or stumpy form. The slender form of the parasite maintains infection in the mammal. Unchecked proliferation of the bloodstream form parasite can eventually kill the host [38]. As noted above, bloodstream form parasites are coated with VSG. It is a homodimer that covers the entire cell. About 5.5×10^6 homodimers are present in one bloodstream form cell [39]. They are embedded in the plasma membrane by an anchoring at the C-terminal end

by a glycosylphosphatidylinositol linker in the outer leaflet of the plasma membrane [40]. The VSG genes are a cohort maintaining antigenic variation in the bloodstream form trypanosomes. There are about 2000 VSG genes, which are expressed from 20 expression sites. Their peptide sequences are extremely unique, with >50-52% sequence identity in close homologs [41]. The *VSG* mRNA is expressed from 20 expression sites, with expression site associated genes (ESAG) a 70bp repeat region. The expression site is transcribed polycistronically by RNA polymerase I. The *VSG* gene cluster undergoes allelic exclusion and gene rearrangement to maintain the antigenic cohort of the bloodstream form trypanosomes [42-44]. Gene recombination, creating new sequences, makes their cohort even more complex.

Upon high parasitemia, the slender parasites stop dividing and become morphologically more rounded. These are the fly-infective stumpy form. On uptake of a blood meal from an infected individual, the parasites move into the midgut of the fly. Here, they experience serious acidic and proteolytic stress. The stumpy parasites are resilient to these stresses, whereas the slender form is rapidly killed [45, 46]. The tsetse fly shows high resistance to trypanosomal infections. In a study to investigate, tsetse infection quantitatively, uninfected tstetse flies maintained under sterile laboratory conditions were given infected bloodmeal. It was found that only 15% of the experimental population and 1% of the natural population of the fly were infected by *Trypanosoma spp.* [8]. Successful development of the parasite in the fly is important for efficient transmission in the mammalian host. In the midgut of the fly, anti-microbial peptides (*defensin, attacin, cecropin*) challenge parasite survival. RNAi against such peptides increase the establishment of the parasite in the tsetse midgut [47]. In the alimentary tract, the generation of reactive oxygen species (ROS) can induce a systemic immune response in the fat cells [47]. Also, production of hydrogen peroxide in the proventriculus region right after blood meal ingestion may reduce the parasite load in the initial stages of infection [48].

The fly form of the trypanosome expresses surface glycoproteins called procyclins. The GPEET and EP procyclins vary in their internal pentapeptide (GPEET) or dipeptide (EP) repeats. GPEET procyclin is highly expressed in the tsetse midgut within the first 4 days of infection, and its expression is repressed within 4-7 days post infection; whereas the glycosylated EP procyclin is expressed in later procyclics [49, 50]. Two closely related genes express GPEET and about 6-7 genes express EP [51]. Like VSG, their expression (using RNA polymerase I) is both transcriptionally and post-transcriptionally regulated [52]. From the fly midgut the parasite travels to the hypopharynx of the tsetse fly. First, epimastigotes expressing BARP are formed. Then these turn into metacylic forms expressing metacyclic VSGs. The metacyclic VSGs have unique promoters active only in the salivary gland stage of trypanosome [53, 54]. Upon uptake of bloodmeal, the fly transmits the metacyclics in the bloodstream of the host, where it shuts off the expression of the metacyclic VSGs and starts expressing normal bloodstream form VSGs again, thus keeping the cycle in continuum.

The life cycle of *Trypanosoma brucei*

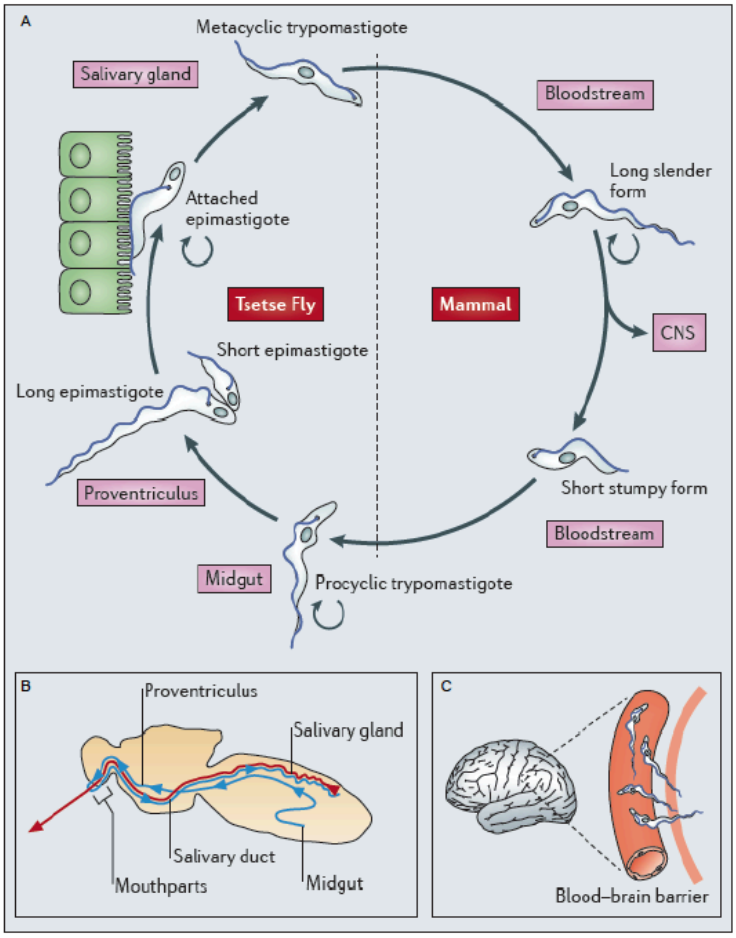


Figure 1.4: A simple representation of the life cycle of Trypanosoma brucei and infection in the tsetse fly and humans. A. Diagram of the life cycle of *Trypanosoma brucei* and its different cellular forms. B. Diagram to depict the infestation and infection of trypanosomes in fly. C. Diagram of second stage infection by trypanosomes by crossing blood brain barrier (taken from [55]).

Energy metabolism of trypanosomes

Trypanosomes adapt their energy metabolism according to their host. For example the blood of the mammalian host is very rich in glucose (5mM), oxygen (95-99%) and proteins (0.6-0.8g/mL) [56]. Trypanosomes possess peroxisome like compartments called glycosomes, which is a specialised compartment for glycolysis and other pathways. Since the glycosomal membrane is impermeable to the entry of ATP, ATP is strictly produced during the cytoplasmic steps in bloodstream forms [57]. During differentiation, the bloodstream form glycosomes are degraded and a new glycosomes are made for the procyclic [58]. The glycosome contents in the two different cell stages vary in many ways, including the enzyme phosphoglycerate kinase (PGK), which carries the last glycolytic step. In the bloodstream form the

major cytosolic PGK is expressed from *PGKC* gene, whereas in the procyclic form *PGKB* expresses the major cytosolic PGK [59]. The bloodstream form trypanosomes depend on Trypanosome alternative oxidase for respiration. It is a cytochrome-independent terminal oxidase of the mitochondrial electron transport chain. Its inhibition is trypanocidal for bloodstream form trypanosomes [60]. In the bloodstream form the enzymes of the TCA cycle are not expressed; however the F_0-F_1 ATPase, preserves the mitochondrial membrane potential (MMP) by ATP hydrolysis, allowing for continuous mitochondrial protein import [61], [62-64] but does not contribute to ATP synthesis.
Not much is known about the stumpy form metabolism. But it has been shown that on inhibition of the major glucose transporter THT1, by sub-lethal concentrations of phloretin, induces mRNA expression of procyclic specific genes like glucose transporter *THT2* and glycolytic enzyme *PGKB* [65]. In the fly midgut there is hardly any glucose, but the environment is rich in amino acids, especially proline (100μM). Proline is the most important substrate for procyclics, but catabolism of threonine and leucine also contribute to the procyclic carbon metabolism, which relies heavily on mitochondrial pathways [66, 67]. In culture, procyclic forms can switch to using glucose for the production of ATP even in presence of proline [68]. We do not have any information about metabolism in epimastigotes or metacyclic trypanosomes.

Gene expression

Trypanosomes are diploid organisms. The genetic information of trypanosomes totals to ≈ 35Mb per haploid cell [69-71]. As mentioned before trypanosomes possess 11 pairs of megabase sized chromosomes, 100 minichromosomes (MC) (30-150Kb) and intermediate-size chromosomes (ICs) (200-900Kb) [33, 69]. The gene expression in trypanosomes are often compared with the other model eukaryotic organisms. Trypanosomes employ three RNA polymerases for transcription (RNA pol I, RNA pol II and RNA pol III), just like other eukaryotes. Like all other eukaryotes RNA pol III transcribes a small cohort of non-coding RNAs, like tRNAs and 7SLRNA [69]. The *T. brucei* RNA pol III is known to include one isoform of RPB5 (specifically the form present in RNA pol II) [72].
RNA polymerase I transcribes rRNA and two types of surface proteins, variant surface glycoproteins (VSGs) in the bloodstream form, and EP or procyclin in the procyclic form [73]. RNA polymerase II (RNA pol II) transcribes the remaining protein coding genes, as in other eukaryotes. In trypanosomes, the genome is organised as directional gene clusters (DGC). Such clusters are also known in *Caenorhabditis elegans*, but DGCs in *C.elegans* comprise of only 15% of the total protein coding genes. In *T. brucei* there are about 388 clusters of two or more genes [33]. Genes within clusters are usually not functionally related. Recently, it was found that *T. brucei* pol II promoters are GT rich and are marked by the occupancy of H2A.Z [74].

Since almost all RNA pol II transcripts are polycistronic, they (as well as the expression site and procyclin mRNAs) have to be *trans* spliced to make a mature mRNA, by adding a 39 nucleotide sequence at 5' end of the transcript. This 39 mer is called the splice leader (*SL*), and is derived from the 140 nucleotide derived splice leader RNA. Therefore gene expression in trypanosomatids depends very highly on

the expression of splice leader RNA gene. It is present in a high copy number: about 200 *SL* RNA genes per genome, and the *SL* RNA has a high turnover rate due to use in splicing (half life ~4 minutes) [10, 25, 75]. Figure 1.5 shows the biochemical reaction of addition of *SL* RNA to a hypothetical exon to form an mRNA.

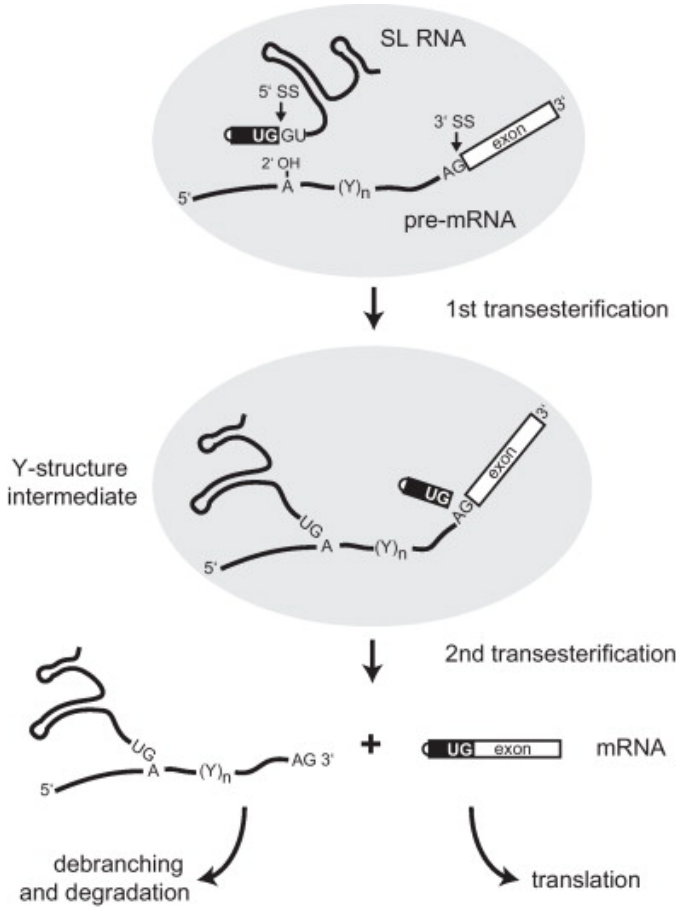


Figure 1.5: Mechanism of trans splicing. In the diagram the SL RNA, along with the pre-mature RNA is depicted. The first *trans*-esterification step involves a nucleophilic attack on the 2' OH site of the branch point Adenosine at the 5' of the splice site. The intermediate is Y-structure consisting of SL mini-exon with SL mini-intron /protein coding exon, with the characteristic 2' 5' phosphodiester linkage. In the second step the 3'OH group of the 5' SL mini-exon attacks the 3' end of the protein coding exon. This reaction leads to the joining of the first 39 nucleotides of the SL mini-exon to attach to the 5' end of the protein coding exon producing an mRNA. The intron is released in the form of a Y-structure and undergoes debranching and degradation (Taken from [10]).

After splicing the mature mRNA is transported by an Mex67/Mtr2 transporter complex, which is conserved in all eukaryotes [76]. Recently, the helicase Hel45 has been identified which is found both in the nucleus and cytoplasm. Knockdown of Mex67 increases the nuclear quantity of Hel45. But its definite role in shuttling mRNA from nucleus to cytoplasm has not yet been identified [77].
Translation in *T. brucei* is similar to other eukaryotes. The eIF4F complex binds the cap. eIF4F consists of the helicase eIF4A and eIF4E, bound to the scaffold eIF4G.

There are six homologs of eIF4E and five homologs of eIF4G in *T.brucei* [78]. Translation is an important step for differential expression of proteins in *T.brucei* [79].

The mRNA degradation machinery

Since genes lack individual promoters the regulation of expression is post-transcriptional. The general modes of post-transcriptional regulation mRNA include control of mRNA processing, translational repression, and mRNA degradation or storage. It has been found, from RNAi and computational studies, that splicing and polyadenylation contribute to homeostasis of mRNAs in trypanosomes [80, 81].
The degradation of a polyadenylated mRNA in yeast and mammalian cells begins with the removal of the poly(A) tail from the 3' end of the mRNA by the poly(A) nucleases or deadenylation complexes. PARN, Pan2-3 and the Ccr4-Not complex are the deadenylation machineries conserved in eukaryotes. The Ccr4-Not complex regulates global gene expression not only by deadenylation, but also via transcriptional regulation [82] and protein ubiquitination [83]. The Ccr4-Not complex of yeast contains 6 subunits, namely Not1p, Not2p, Not3/5p, Not4p and the deadenylases Caf1 and Ccr4. In yeast and humans Ccr4 is the main mRNA deadenylating enzyme. Not1p acts as a scaffold for the other subunits. Not2p interacts with Ada2, a component cofactor of Gcn5p Histone acetyltransferase (HAT), and has a role in transcriptional regulation [84] and Not4p acts as an E3 ubiquitin ligase. Not3/5p supposedly has a role in signal transduction and binds to small rho G proteins. Beside, the above core complex, the yeast deadenylation machinery also interacts with RNA helicase, Dhh1, and has specific proteins like Caf130p. In addition to these roles the Not1p subunit in the nucleus imprints mRNAs of ribosomal proteins in a Not5 dependent manner, which facilitates their translation [67]. In humans, Caf130 is replaced by CNOT10. This protein forms a sub-complex with another subunit called C2ORF29/CNOT11, which is conserved from Drosophila to humans [85]. In addition to this, CAF40 interacts with the N-terminal region of NOT1 by its DUF3819 domain.
A number of factors mediate the interaction between mRNA and the deadenylation complex. In humans and *Drosophila*, both miRNA and Smaug are capable of recruiting the NOT complex to mRNA to promote deadenylation [8]. In humans the scaffold subunit NOT1 interacts with RNA binding protein tristetraprolin and recruits CAF1 to the bound mRNA and hence targets it for deadenylation [86]. Another *in vitro* study to determine RNA binding partner and activity of yeast Ccr4-NOT complex showed that the complex harbours Mmi1 which, when it binds to a recombinant deadenylation complex, can cause deadenylation of a poly (A) tail as fast as 4 minutes [87].
Although many RNA binding proteins have been identified to interact with the deadenylation complex to target bound mRNA deadenylation, its regulatory pathway has been elusive. Recently it has been demonstrated that accumulation of AMP inhibits the deadenylation activity of the CCR4-NOT complex in *Drosophila* [68]. Also in a recent study it was demonstrated that histone acetyl transferases (HATs) p300 and CBP acetylate CAF1, which promotes its deadenylation activity [88]. Besides, a role in total mRNA decay, the mammalian CNOT3 regulates differentiation genes negatively. It is highly expressed in blastocysts, and its knockout is embryonic lethal.

In a conditional knockout study it was found that on depletion of CNOT3 gene, genes involved in developmental process are specifically upregulated [56].
There are three different kinds of deadenylation machineries in trypanosomes: PAN2/3, PARN proteins and the CAF1-NOT complex [89, 90]. The trypanosome CAF1-NOT complex consists of CAF1, NOT1, NOT2, NOT3/5, NOT10, NOT9/CAF40 and DHH1[89]. It lacks homologs of CNOT4 of humans and Caf130 and Ccr4 of yeast. Yeast–two-hybrid assay results suggested that CAF1 interacted with NOT10 and the N-terminal half of NOT1, and NOT10 interacted with NOT3/5 [91, 92]. NOT10 is required for the integrity of the complex, as its depletion in trypanosomes leads to detachment of CAF1 from the NOT complex and lowers the amount of NOT1 [92]. Affinity purifications of TAP-NOT10 revealed the subunits of the complex, namely, NOT1, NOT2, NOT11, CAF1, and the RNA binding proteins PUF2 [92], [93]. A recent study on regulatory networks for gene expression revealed that NOT11 can bind to mRNA [94], [95]. In general we lack information about RNA binding protein partners of the CAF1-NOT complex of trypanosomes, which might target the deadenylation complex to a specific mRNA. This suggested my genome-wide and proteome wide studies to discover novel partners of the CAF1-NOT complex.
Trypanosomes lack homologs of usual Dcp2 decapping enzyme. Instead they use ApaH like-phosphatase as a major decapping enzyme [31]. RNAi against the enzyme arrests growth and increases the total amount of deadenylated mRNAs in the cell. Also, this enzyme has an in vitro decapping capability. It is also established that decapping is followed by 5'-3' exoribonuclease digestion of mRNA. It was also observed that the 5'-3' trypanosome exoribonuclease XRNA localises completely with ApaH-like phosphatase. The trypanosomes have the classical exosome complex, with an exception that the subunit RRP6 is found in both the nucleus and cytoplasm, and is required for exosome stability [96, 97]. The exosome is responsible for degradation of deadenylated mRNA in a 3'-5' direction. Many developmentally regulated mRNAs change their decay pattern from one developmental stage to the other [81, 98, 99].

Post-transcriptional gene regulation

Post-transcriptional regulation of the transcripts by trans-acting factors is extremely important in gene expression regulation in trypanosomes. I here describe some examples or RNA binding protein (RBP) functions.
One RBP that regulates differentiation and development is RBP10. RBP10 is composed of an RRM domain and low complexity regions. It is exclusively expressed in the bloodstream form, and overexpression of RBP10 in the procyclic form of the parasite is lethal [19, 100]. Another RNA binding protein-regulating differentiation of trypanosomes is RBP6. Expression of RBP6 in procyclic trypanosomes induces differentiation to epimastigotes and metacyclic forms [101]. Besides regulating differentiation, RBPs also regulate other cellular processes like metabolism. RBP42 modulates metabolism of *T. brucei*, by binding to mRNAs involved in cellular respiration, carboxylic and amino acid metabolism [102].
RNA binding proteins are also essential for the survival of the parasite under potentially harmful environmental stimuli. One RBP important in stress is the RNA

binding protein ZC3H11. Knockdown of its expression is lethal in the bloodstream form, but not for the procylics. However, procyclic knockdown clones of ZC3H11 are incapable of survival after heat shock [103]. Also the expression of ZC3H11 increases from $2x10^3$ molecules per procyclic cell to $15\text{-}20x10^3$ molecules upon heat shock at 37°C or 41°C [104]. ZC3H11 binds to AU rich elements in the 3' UTR of the chaperone mRNAs, including that of HSP70, and stabilizes them for translation during heat shock [105]. In absence of ZC3H11, *HSP70* mRNAs are localised in the heat shock granules under stress [106]. ZC3H11 interacts with MKT1. MKT1 then interacts with PABP and LSM1 and promotes translation of ZC3H11 bound transcripts [107].

The RNA binding proteins which regulate the fate of mRNAs might also be modified to serve specific functions. For example the RBP DRBD18, has 3 methylation sites. A fully methylated DRBD18 binds target mRNAs and prevents their decay, whereas a hypomethylated DRBD18 promotes decay of the bound targets [108]. Beside post-translational modification of RBPs, the localisation of the RBP itself under certain conditions control the fate of its targets. For example DRBD3, an RBP associated with PABP1, PABP2, RNA helicases, and U1 A snRNP complexes, remains bound to its target mRNAs in distinct subcellular foci upon starvation, and relocalises to nucleus upon treatment with sodium arsenite [109].

Adaptation to stress regulation by *Trypanosoma brucei*

Trypanosomes complete their life cycle by living in their mammalian host (cattle or human) or in their fly vector, the tsetse fly. In these hosts, the parasites experience stresses. Therefore the parasites have evolved mechanisms to combat such conditions for survival. As described before, the RBP ZC3H11 undergoes phosphorylation and stabilisation during heat shock and stabilises chaperone mRNAs during heat shock [104, 105]. Starvation in eukaryotes induces changes in cytosolic RNA-protein granules. P-bodies are present during normal growth, and contain RNA helicases like Dhh1, and decapping enzymes [110, 111]. After stress, larger granules are formed with different protein contents.

P-bodies have not been identified in trypanosomes, but stress granules are formed. In order to comprehend the RNA and protein composition of such granules in trypanosomes, procyclic form cells were starved for 2 hours, and the granules were isolated and the proteins were identified by mass-spectrometry and resident mRNAs by RNA-Seq [112]. A huge repertoire of RBPs was identified in the starvation granules, along with the known markers like PABP1 and DHH1. It was also observed that upon stress most mRNAs are relocalised to the granules; this is thought to prevent their translation. Therefore, these granules serve as storage sites and also regulate stress by repressing translation of almost all mRNAs excepting those of the chaperones.

Trypanosomes do not have the usual machineries for combating stress inflicted by protein denaturation in the ER (unfolded protein response or UPR). Instead, parasites suffering from endoplasmic reticulum stress induces silencing of the splice leader (SL) RNA transcription, thereby inhibiting processing of all mRNAs and causing cell death [113, 114].

One of the recent studies that confirm that trypanosomes possess factors to combat stress other than heat shock was on the potassium channel of *T. brucei* localising in the acidocalicisomes. It was suggested that TbIRK, which is a transmembrane protein selectively transferring potassium ions might have roles in regulating intracellular pH and osmoregulation [29]. Also trypanosomes tend to protect their proteins from oxidative stress by protein-S-trypanothionylation and protein S-glutathionylation [115, 116]. Total proteomics was done to find out which proteins are affected by oxidative stresses. Only 3 proteins were increased by all stresses inflicted upon the bloodstream form parasites - H_2O_2, Diamide and NaOCl. They were heat shock proteins, HSP20, HSC70 (heat shock protein 70 interacting protein) and a protein disulphide isomerase, PDI. Most of the other proteins were specifically enriched by a certain compound, indicating differences in combating stress [117]. Also studies on compounds with trypanocidal activity demonstrate that alkaloids like emetine and homoharringtonine exploit their stress induction ability to interfere in biological pathways in trypanosomes, that change mitochondrial membrane potential (homoharringtonine) or replication inhibition (emetine) [45].
Therefore study of stress and scrutinizing pathways that challenge survival of host or parasite and host-parasite or vector-parasite interactions are interesting for developing drugs or designing methods that prevent transmission of the parasite.

Screens to identify regulators of gene expression in trypanosomes

Since gene expression in trypanosomes is usually regulated by RNA binding proteins, a number of approaches have been used to identify these factors. One study used a "tethering" assay to identify RBPs in a genome wide fashion. A random genomic shotgun DNA library was prepared and was transfected in cell lines expressing positive or negative selectable marker, and bearing a 'Box-B' λ-N recognition element at the 3' UTR of the marker. The library expressed genomic fragments fused to a λ-N peptide at the N-terminus. Thus the λ-N fusion proteins could bind to the selectable marker RNA. The screen identified 300 potential post-transcriptional regulators [118]. In order to validate this screen, a library expressing the complete open reading frame of 300 potential post-transcriptional regulators was constructed. RBPs repressing mRNA expression were identified, and an in vitro mRNA bound proteome screen was carried out to identify true mRNA binding partners. The combination of the above methods helped in identifying 155 mRNA binding proteins, which were either gene expression repressor or activators [94]. In this screen, ZC3H30 was shown to be a repressor of reporter expression and a *bona fide* mRNA binding protein. Also, in the procyclic-form starvation granule study, ZC3H30 was enriched in the granule fraction [112]. All these results provoked me to study this zinc finger RNA binding protein, ZC3H30 and find its function in trypanosomes, both in the bloodstream (BS) and the procyclic form (PC).

Aim of the study

Part1-Objectives

- To discover novel interacting partners of the CAF1-NOT complex of *Trypanosoma brucei*

The subunits of the CAF1-NOT complex have been already demonstrated to interact with each other ([89, 92, 118]). Whether or not the protein subunits interact with RNA binding proteins (RBPs) that guide them to mRNA for deadenylation has not been discovered. Also, no roles of this large complex beside deadenylation are known. Hence, I decided to further investigate its biological role by looking at its protein-protein interaction map.

Part2-Objectives

The aim of the second part of my thesis was to characterise the role of ZC3H30 in trypanosomes. ZC3H30 was already known as a RNA binding protein and a repressor from high throughput study. My steps were:

- Validation of the effect of ZC3H30 when tethered to a reporter mRNA; both on transcript abundance and reporter protein expression and activity.
- To find the role of ZC3H30 in growth and survival of trypanosomes (procyclic form and bloodstream form) under normal conditions in cell culture
- To find the role of ZC3H30 in growth and survival of trypanosomes (procyclic form and bloodstream form) under stress conditions. Cells lacking ZC3H30 had poor survival after stress, so this aspect was followed up.
- Localization of ZC3H30 under normal conditions and heat shock.
- Effect on chaperones and stress markers in absence of ZC3H30.
- To find the protein and RNA-binding partners of ZC3H30.
- Effect on global transcriptome in absence of ZC3H30 expression.
- Validation of interaction between a novel protein and ZC3H30 by co-immunoprecipitation.
- Role of the novel protein in normal growth and under stress in trypanosomes.
- Role of the novel protein in localization of ZC3H30 under heat shock.
- Effect on global translation on inhibition of total protein synthesis in ZC3H30 double knockouts, under normal conditions and after heat shock.

Chapter 2-Materials and Methods

Some of these method descriptions are partially excerpted from previous PhD theses or from papers I wrote.

Part1- The Genome-wide interactors of the CAF1-NOT complex

Vectors and cloning of bait and prey

The full length open reading frame (ORF) of NOT11 was cloned as bait in the pGBKT7 vectors as per the Matchmaker GAL4 Two-Hybrid System3 by Clontech (for the plasmid of NOT10 refer to [92]). The ORFs for NOT2 and CAF40, which were cloned using GATEWAY cloning system, in a special PGBKT7 plasmid modified for GATEWAY cloning. A trypanosomal shotgun genomic DNA library was prepared by cloning the DNA fragments ligated to an adaptor in the prey plasmid pGADT7[119].

Yeast strains and Yeast-two-hybrid mating screen

The bait plasmids were transformed in AH109 yeast strains and the pool of prey plasmids carrying fragmented trypanosomal genomic DNA inserts were transformed in Y187 yeast strains. The AH109 cell lines carrying a specific bait protein was grown for 16 hours in SD media without tryptophan. Following it was diluted to 1OD, and mixed with the cells from prey library in an enriched 2x YPD media. The cells were allowed to mate for 22 hours at 27°C. Due to pheromone production and sub-optimum temperature for growth, mitotic division cannot take place. The mated cells were pulled down by centrifugation at 1000xg for 5 minutes and resuspended in 25 to 30ml of 0.5x YPD media (unenriched media to avoid cell division that might take place while plating the cells). The mated cells were plated on SD agar plates lacking tryptophan, leucine, or both (double drop-out medium or DDO) to calculate mating efficiency, and total number of diploids obtained for each experiment. For screening colonies the mated cells were plated on SD agar high stringency plates lacking tryptophan, leucine, adenine and histidine (Quadruple drop-out medium- QDO). The plates were incubated for about 3 to 5 days at 30°C. The colonies from the high stringency plate were streaked on high stringency SD plates with 40 µg/ml X-α-Gal, and were incubated again at 30°C for 3 to 5 days. Only blue colonies that indicated strong interaction between, bait protein and prey peptide fragment were pooled in high stringency SD media, and the cells were grown and amplified for plasmid isolation. The inhibitor of His3 enzyme, 3-amino triazole (3-AT) was spread on the high stringency SD plates at various concentrations, ranging from 0.5 to 2mM to avoid the growth of clones carrying weakly interacting preys and baits. Only true interactors that interacted strongly even in presence of 3-AT could grow into a colony.

Amplification of prey inserts interacting with bait in screen

For each screen of library against a specific bait, the clones obtained on high-stringency, X-α-Gal, 3-AT, SD plates were pooled in a single flask by washing with high-stringency SD medium. The yeast culture was mixed thoroughly and 1ml of the mixture was re-inoculated in a fresh high-stringency SD medium and grown overnight at 30°C. Plasmid DNA was isolated and a PCR reaction was carried out to amplify only the inserts of the prey vector. The 5' end forward primer includes a linker, which was used to link the sheared genomic DNA inserts and prepare the library, and which has unique sequences that were used as barcodes. A different barcode was used for recognising genes screened from different baits. A universal 3' end primer was used in combination with 5' end forward primers. The number of cycles for PCR was minimized between 22 and 25, so that inserts abundant in library would not be over-represented in the amplicon mixture. The purified amplified and bar-coded DNAs from different experiments were sent for deep sequencing at BioQuant and EMBL, Heidelberg, Germany.

Deep sequencing data analysis (reads and coverage-codes written and run by Dr. Abeer Fadda; RPM calculation and data mining- Chaitali Chakraborty)

The reads were aligned to the trypanosome reference genome Lister 427 strain. The unaligned reads were discarded and only those aligned reads that were predicted to be translated were kept. Translation of the sequences were considered for both sense and anti-sense strands. The final reads obtained were used for data analysis. First those genes, which had reads less than 10, were discarded. This list was compared with the unique gene list, to obtain a list of interacting proteins and their respective genes. From this list RPMs were calculated for every gene. A list of genes with RPMs lower than 100, were discarded. A list of abundant genes in the library was prepared by comparing the unique gene list of interactions. All those genes, which appeared not only in the screens for the subunits of the CAF1-NOT complex, but also in other the yeast-two-hybrid interaction list, were discarded. The remaining genes on the list were compared with data previously obtained from different studies to find a role of the interaction between the bait (the subunits of the CAF1-NOT complex) and the prey (interactors obtained in the mating assay).

Bioinformatic tools for data visualisation

The tool Venn Diagram by UGent was used to make a venn plot to visualise common genes in Biological processes and putative interactome list of the screened subunits. I used the GO tool from Gene Ontology consortium, to look for classes of proteins, and their molecular function, which were abundant in our interactome list. We looked for protein class enrichment by comparing our list with the genome of the unique list of genes from the *T.brucei* Lister 927, and used Fisher's exact test to check for significance, using the online tool Social Science Statistics.
(The online URL for Venn Diagram from UGent is-
/http://bioinformatics.psb.ugent.be/webtools/Venn/)

Part2: Trypanosomes lacking the RNA Binding Protein ZC3H30 are sensitive to stress.

Cell culture techniques

Trypanosome cell culture

All experiments were done using Lister strain 427 monomorphic procyclic or bloodstream form (BS) parasite expressing the tet-repressor [120]. Procyclic forms (PC) were cultured in MEM-Pros Medium at 27°C at densities lower than $8x10^6$ cells/ml. Bloodstream form parasites were cultured in HMI-9 medium at densities lower than $1.5x10^6$ cells/ml. Stable cell lines were created, expressing tagged endogenous copy or tagged ectopic copy in frame with the open reading frame or carrying stem-loop RNAi constructs were created. For details on the construct, please see 'Oligos' and 'Plasmids' (section right after Materials and Methods).

Components of media

MEM-Pros (500ml): The powdered MEM-PROS medium was mixed with MEM-Pros vitamins, non-essential amino acids in a solution with 100mg Phenol Red (pH 7.4). The media was sterilised by filtration and stored at 4°C.

To 450ml of MEM-Pros medium the following components were added before use:

- Heat Inactivated FBS (55°C, 30 minutes); 10% v/v (50ml)
- Haemin 7.5mg/l (1.5ml from stock)
- Penicillin/ Streptomycin 50U/ml [5ml of stock (5000U/ml)- preparted by Ute Leibfried]

HMI-9 medium (500ml): The media contained 17.66g/l Iscove's modified Dulbecco's medium, 3.024g/l of $NaHCO_3$, 136mg/l Hypoxanthine, 110mg/l Sodium Pyruvate, 39mg/l Thymidine and 28mg/l bathocuprono disulfonic acid disodium salt (pH 6.3). The media was sterilised by filtration and stored at 4°C.

To 450ml of HMI-9 medium the following components were added before use:

- Heat Inactivated FBS (55°C, 30 minutes); 10% v/v (50ml)
- L-Cysteine (1.5mM)
- β-mercaptoethanol (0.14%)
- Penicillin/ Streptomycin 50U/ml [5ml of stock (5000U/ml)- preparted by Ute Leibfried]

Transfection of Bloodstream or Procyclic trypanosomes

For each transfection 2-2.5 $x10^6$ cells were used. The cells were centrifuged at 850xg for 8 minutes and the cell pellet was washed twice with the transfection buffer. Bloodstream form parasites were transfected with Amaxa buffer (90mM NaH_2PO_4, 5mM KCl, 0.15mM $CaCl_2$, 50mM HEPES, pH 7.3). The procyclics were transfected with ZPFM buffer (132mM NaCl, 8mM KCl, 8mM NaH_2PO_4, 1.5mM KH_2PO_4, 1.5mM $MgOAc.4H_2O$, 90μM $CaOAc_2$, pH 7).

The washed cell pellet was resuspended in transfection buffer (100μl for BS and 400μl for PC) and mixed well with 10μg of linearized vector construct in a 2mm gap electroporation cuvette. Bloodstream form cells were transfected using Amaxa Nucleofactor (Lonza Cologne AG, Germany). The procyclic parasites were transfected by the BTX electroporation machine (Harvard apparatus) at reistance R2 and 1.5KV. The electroporated cells were transferred to 25 ml of pre-warmed media and left overnight (16h) to recover. The following day selection antibiotics were added and the culture was plated on a 24 well plate. For bloodstream form parasites, positive clones were selected after 5-7 days and for procyclics, positive clones were selected after 10 or 14 days.

Antibiotics	**Bloodstream form**	**Procyclic form**
Phleomycin	0.2μg/ml	0.2μg/ml
Blasticidin	5μg/ml	10μg/ml
Puromycin	0.2μg/ml	1μg/ml
Neomycin-G418	5μg/ml	15μg/ml
Hygromycin	15μg/ml	50μg/ml

DNA techniques

Cloning and plasmid construction

The lists of plasmids, plasmid maps and oligonucleotides are given at the end of 'Materials and Methods' section. Each construct was designed virtually on Seq Builder. Primers were ordered from Biomers.net. Genomic DNA was isolated by ammonium acetate-isopropanol precipitation. For cloning a DNA fragment from genomic DNA, NEB Q5 DNA polymerase was used. The PCR product was purified using Machery–Nagel PCR purification kit, and was digested with appropriate restriction enzymes from NEB or Thermo Fischer's Fast Digest. The PCR product was ligated to the vector of interest digested by same restriction enzymes as the PCR product and dephosphorylated (Alkaline Phosphatase NEB) using NEB T4 ligase. The ligated product was transformed in DH5α chemical competent cells by heat shock and then plated on LB-agar plates the selection antibiotic against which the vector harbours the resistance marker. Positive clones were isolated the following day and confirmed first by colony PCR using NEB Taq polymerase, then by plasmid isolation (Machery-Nagel Plasmid isolation kit) and restriction digestion. Only one correct plasmid was sent for sequencing at GATC for verification.

Genomic DNA extraction

For isolation of genomic DNA, $2x10^8$ cells either BS or PC were centrifuged at 1125xg for 10 minutes. The supernatant was discarded and the cell pellet was washed once with 1x PBS (phosphate buffer saline) and resuspended in 0.5ml EB

Lysis buffer (10mM Tris-HCl pH 8.0, 10mM NaCl, 10mM EDTA, 0.5% SDS) with 1.2μl of RNase A (stock-10mg/ml). The suspension was incubated at 37°C for 1 hour followed by addition of 200μl of ice cold ammonium acetate and then gently vortexed. The suspension was then centrifuged at 6000xg for 5 minutes at 4°C. The clear lysate was carefully transferred in another tube and to it 0.7 sample volumes of Isopropanol was added, followed by vigorous mixing. The mixture was then centrifuged at 16,000xg for 15 minutes. The supernatant was discarded and the DNA pellet was washed with 1ml of 70% and 100% Ethanol sequentially. After the final wash the pellet was dried shortly and then resuspended in TE at 37°C. The quality of the purified DNA and the concentration was measured using NanoDrop 2000/2000c by Thermo Scientific.

Southern Blotting

In order to verify the genomic arrangement and deletion of endogenous *ZC3H30* gene, a Southern blot was done. Genomic DNA was isolated from the double knock-outs and was digested by *Apa* I and *Sal* I enzymes which would digest the DNA from the 5' UTR of the ZC3H30 locus into the open reading frame of the gene, resulting in 1Kb fragment. To ensure that DNA was well transferred, a probe was designed for *β-tubulin* gene. 10μg of digested genomic DNA from the wild type and the double knockouts was loaded on a 0.8% agarose gel and electrophoresed at 120V for 4 hours. Before blotting the gel was treated with 0.25M HCl for 15. This treatment hydrolysed long DNA. Then it was treated with 0.4M NaOH and 1.5M NaCl to denature and depurinate DNA fragments. Finally it was washed twice for 15 minutes with neutralization buffer (1.5M NaCl, 0.5M Tris-HCl pH 7.5). Then the southern blot was assembled like Northern blot and the process of pre-hybridisation and hybridisation of probes were alike. The probe against *β-tubulin* was used as a loading control.

Oligo ID	Name	Sequence
cz6181	ZC3H30miniORF(HindIII)Fw	AAT**AAGCTT**CCATGCCTCCCGCTGACAGG
cz6182	ZC3H30miniORF(ApaI)Rv	AAT**GGGCCC**TGTTGACGGAGCATCGGAAGAA
cz2581	tubulin probe FW	CCTTTGGCACAACGTCACCACGG
cz2724	tubulin probe RV	TGACTCGCCGCAACCTCGAT

Table: Oligonucleotides used to prepare PCR amplicons for radioactive probe preparation for Southern blot.

Protein Techniques

Cell fractionation

A subcellular fractionation method was applied to designate the localization of V5-ZC3H30 in bloodstream trypanosomes as previously published [121]. About $2x10^8$ bloodstream form cells were harvested (850xg, 10minutes). The cells were washed once in 1ml of trypanosome homogenisation buffer (THB: 25mM Tris-HCl pH 7.8, 1mM EDTA, 10% Sucrose, 5μg/ml Leupeptin). In another eppendorf tube an equal volume of Silicon Carbide (SiC: about the same volume as the pellet of trypanosomes), were taken, and washed with 1ml of THB. The SiC-THB mixture was spun (16,000xg, 5 minutes), and the buffer was removed. The washed trypanosomes were resuspended in 0.5ml THB and mixed thouroughly. The homogenate then was added to the abrasive SiC and mixed thoroughly. The homogenate was centrifuged briefly (6000xg, 5 minutes) and the supernatant was discarded. The SiC-trypanosome mixture was ground on ice with an ice-cold pestle for 5 minutes. To it 0.5 ml of THB was added and centrifuged (16000xg, 15 minutes). From the supernatant, 50μl aliquot was collected for Western Blot analysis from the soluble cytosolic fraction, and to the pellet (nuclear and organellar debris) 0.5ml of THB buffer was added. The pellet was homogenised and the sample was passed through a 21G syringe for 20 times. It was then centrifuged (16,000xg, 5 minutes); a 50μl aliquot was collected as Western Blot sample of the organellar soluble fraction (NSup) fraction, and to the pellet 50μl of Laemli buffer was added and labelled as nuclear pellet (NPel) fraction. The samples were analysed by Western blotting method. The antibody against Trypanothione reductase (α-TR) was used as a cytoplasmic control and antibody against exoribonuclease D (α-XRND) was used for a nuclear marker.

Co-immunoprecipitation

About 10^8 cells expressing C-terminally YFP-tagged Tb927.8.3820 and/or N-terminally V5-tagged ZC3H30 were centrifuged at 2300 rpm for 10 minutes. The cells were washed in cold 1x PBS (phosphate-buffer saline), followed by lysis in hypotonic solution (10mM NaCl, 10mM Tris-HCl pH-7.5, 10μg/ml Leupeptin, and 0.1% Igepal). The lysed cells were centrifuged at 16000xg for 20 minutes and the salt concentration was adjusted to 150mM NaCl. Then the cleared lysate was incubated with 30μl of either GFP-binding protein conjugated or anti-V5 antibody conjugated agarose beads for 1.5 hours at 4°C. The unbound fraction was discarded and the beads were washed 4 times in wash buffer (150mM NaCl, 10mM Tris-HCl pH 7.8, 0.1% Igepal, and 10μg/ml Leupeptin). Western blot samples were collected during the procedure from total lysate (Input-I, Unbound-Un: clear lysate post-incubation with beads; Beads-B: protein bound to beads post incubation and washing).

Western Blotting and protein detection

SDS-polyacrylamide gels (standard recipe) ranging from 10-15% were prepared depending on the size of the protrein to be analysed. The protein samples were loaded on the gels and they were run at 80V till bromophenol blue from the loading dye entered the resolving gel, and then switched to 124V with the BIORAD system. Standard 1xSDS running buffer was used (25mM Tris, 192 mM Glycine and 0.1%

SDS) [104, 105]. Antibodies used were against V5-tag (AbD sero Tec, 1:1000), GFP-tag (Roche, 1:1000), myc-tag (Santa Cruz, 1:1000), aldolase (rabbit, 1:50,000), SCD6 (rabbit, 1:5000), Dhh1 (rabbit, 1:10000), trypanothione reductase (TR, rabbit, 1:1000), HSP70 (chicken, 1:1000), ZC3H11 (rat, 1:1000), peroxiredoxin (TxNPx, rabbit, 1:1000), ribosomal protein S9 (rat, 1:1000), HSP60 (mouse, 1:4000), Protein A (Sigma, 1:50,000), PAP (Sigma, 1:5000) and CAT (Sigma, 1:5000).

CAT Reporter Assay

Cell lines expressing constitutively CAT mRNA with or without 5x BoxB sequence (from λ-phage) at the 3'-end were transfected with λ-ZC3H30-myc constructs and stable lines were selected. Expression was induced with tetracycline (200ng/ml). The expression of the λ-ZC3H30-myc construct was validated on a Western blot using anti-myc (α-myc) antibody (1:1000, mouse IgG, Santa Cruz). About $2.5x10^7$ bloodstream form trypanosomes (BS), were harvested (850xg, 8minutes), and washed with 1ml of 1x PBS once. For each sample, the cells were resuspended in 200μl of assay buffer (100mM Tris-HCl, pH 7.8), and was lysed by freeze-thawing in liquid Nitrogen and at 37°C respectively at least thrice. The lysed samples were centrifuged to remove cell debris (13000rpm, 3minutes) and the lysate was transferred in a separate eppendorf tube. The protein concentration for each sample was determined by Bradford Colorimetric assay, using different concentrations of bovine serum albumin (BSA) to make a standard curve. BIO-RAD SmartSpecTM3000 recorded the OD for each sample at 595nm. The protein concentration for each sample was determined using the BSA standard curve. To determine CAT enzymatic activity, equal amount of protein from each sample (~ 1μg of total protein) was taken. Each sample was diluted using CAT assay buffer to 50μl. To the diluted sample about 200μl of CAT assay buffer, 2μl of Chloramphenicol (40mg/ml) and 10μl of ^{14}C-butyryl CoA and 4ml of scintillation cocktail (Ultima Gold^{TM} F, Perkin Elmer) were added. Beckman LSC6000IC, scintillation counter, using program #7, recorded the amount of butyryl ^{14}C chloramphenicol. The butyryl ^{14}C chloramphenicol is soluble in the scintillation cocktail, whereas the butyryl ^{14}CoA and CAT enzyme remain on top in the aqueous phase.

Polysome fractionation

About $4x10^8$ procyclic trypanosomes (unstressed or treated with 25μg/ml of Hygromycin B) were treated with cyclohexamide (100μg/ml) for 5 minutes and harvested at room temperature (880xg, 10 minutes). Lysis was performed as previously described (Minia et al., 2016- ZC3H11 paper), and centrifuged to obtain a clear lysate (16,000xg, 10 minutes). The clear lysate was layered on sucrose gradient (15-50%; 4ml) and centrifuged for 2 hours at 4°C in Beckman SW60 rotor. The fractionation of gradients and their absorbance profiles were monitored using Teledyne Isco Foxy Jr system. Aliquots were taken from each fraction for Western blot analysis.

Methionine Pulse assay

About $5x10^6$ procyclic form trypanosomes were harvested by centrifugation (850xg, 8min, 27°C), and washed in 1ml of 1xPBS at room temperature. The pellet was resuspended in labelling medium (modified DMEM from Gibco, lacking L-methionine) and incubated at 27°C for 1 hour. Then 20μCi of S^{35}-L-Methionine was added to the medium and incubated for 20 minutes at 27°C .The cells were then pelleted (850xg, 5 minutes), washed twice in 1ml of 1x PBS, and resuspended in 15μl of Laemli buffer. The samples were loaded on 10% SDS-PAGE, and analysed by autoradiography.

Tandem Affinity Purification and Mass-Spectrometry analysis

This is a two-step purification process-

First purification:

Proteins interacting with ZC3H30 was identified by tandem affinity purification (TAP) and was done using TAP-ZC3H30 fusion cell line. The tag contains Protein A peptide which binds to IgG, followed by a protease cleavage site (specifically Tobaco Etch Virus cleavage site) and a calmodulin binding peptide at the C-terminus of the peptide [122]. About 10^{10} procyclic form trypanosomes expressing N-terminal TAP tagged ZC3H30 in the endogenous locus were harvested at 1125xg for 10 minutes. Three replicates were done with or without 200mg/ml RNase A. The cell pellet was washed once with 1x PBS and then resuspended in 4 ml of lysis buffer (TAP-lysis buffer: 10mM Tris-HCl pH7.8, 10mM NaCl, 0.1% Triton X-100, 5mg/ml Leupeptin). The suspension was mixed thoroughly by pipetting up and down and then the cells were lysed by passing through the 21G syringe for 25-30 times. The supernatant obtained was salt adjusted to 150mM NaCl. An aliquot (input-I) for Western blot was collected which was equivalent to 10^7 cells; to it equal volume of 2x Laemli buffer was added. The cell lysate was centrifuged for 10 minutes at 16,000xg at 4°C. In the meantime 100ml of IgG-Sepharose beads were washed with 1ml of 1x PBS and then equilibrated with TAP-wash buffer (10mM Tris-HCl pH7.8, 150mM NaCl, 0.1% Triton X-100, 5mg/ml Leupeptin) by rigorous washing for 3 times. To the equilibrated beads the cell lysate was added. The beads together with the lysate were incubated for 2 hours at 4°C. After incubation a certain volume equivalent to 10^7 cells was collected as the unbound fraction. The beads were washed thoroughly 4 times and then incubated with Tobaco etch virus (TEV) protease in 1ml TEV cleavage buffer (TAP wash buffer adjusted to 0.5mM EDTA and 1mM DTT) for 1 hour at room temperature. The eluate was recovered by gravity elution.

Second purification:

In the meantime, 100ml of Calmodulin affinity beads were washed and equilibrated in Calmodulin binding buffer (TAP-wash buffer, 10mM β-Mercaptoethanol, 1mM Magnesium acetate, 1mM Imidazole, 2mM $CaCl_2$). To the eluate from the first purification, 3ml of 1M $CaCl_2$ was added and it was added to the washed Calmodulin

beads and incubated for 2h at 4°C. After incubation the beads were washed for 4 times and the eluate was recovered by Calmodulin elution buffer (TAP wash buffer, 1mM Magnesium acetate, 1mM Imidazole, 2mM EGTA). The eluate was concentrated by standard TCA precipitation and run for 1cm on a 10% SDS-PAGE resolving gel and stained with soluble Coomassie. The protein containing gel area was then sliced in 5 pieces, followed by mass spectrometry analysis.

RNA methods

RNA isolation and agarose gel electrophoresis and Northern blotting

Total RNA was extracted using peqGOLDTrifast (Peqlab). Typically 10μg of RNA (from total cell lysate or from fractions RNA immunoprecipitation) was mixed 2x RNA loading dye (1.6% MOPS-morpholino propane sulfonic acid, 7% formaldehyde, 65% formamide, 50μg/ml ethidium bromide, 0.025% bromophenol blue). The mixed samples were denatured for 15 minutes at 65°C and resolved on a 1.5% agarose /2.5% formaldehyde /1% MOPS gel, and blotted overnight with 10xSSC (saline sodium citrate buffer) on nylon membranes (Hybond N+ from GE Healthcare). After blotting the blots were UV crosslinked at 400mJ/cm^2 (Stratagene UV Stratalinker 2400) and stained for 10 minutes with Methylene blue solution (0.5M Sodium actetate, pH 5.2 and 0.04% Methylene blue). The blots were pre-hybrideised in hybridisation solution (5x SSC, 0.5% SDS, 5x Denhardt's solution and 100μg/ml Salmon sperm DNA) for 1 hour at 65°C with *CAT*, *HSP70* (Tb927.11.11330), β-*tubulin* (Tb927.1.2370). Post-hybridisation the blots were washed twice with Wash Solution 1 (2xSSC, 0.1% SDS) at room temperature for 10 minutes and once with Wash solution 2 (1xSSC, 0.1% SDS) at 65°C for 15 minutes.
To detect total RNA the membranes were pre-hybridised for 2 hours at 42°C in pre-hybridisation solution for splice-leader oligonucleotide (6x SSC, 0.5% SDS, 5x Denhardt's solution, 0.05% Sodium pyrophosphate and 100μg/ml Salmon sperm DNA) and then hybridised for 3-4 hours with a radiolabelled oligonucleotide anti-sense to mini-exon (labelled using [γ-^{32}P]-ATP, and T4 polynucleotide kinase, NEB) in the hybridisation solution for oligos (6x SSC, 0.5% SDS, 1x Denhardt's solution, 0.05% Sodium pyrophosphate and 100μg/ml Salmon sperm DNA). Then the membranes were washed thrice at room temperature for 15 minutes and once at 42°C for 10 minutes with SL wash solution (6xSSC, 0.05% Sodium pyrophospahate). The *rRNA*s from the Methylene blue staining were used as loading control. MultiGauge Software was used for quantification of signals and analysis.

Oligo ID	Name	Sequence
cz2827	HSP70-Fw	ggggatccATGACATACGAAGGCGCC
cz3257	HSP70-RV	ACGACTCCGCAACCTCCT
cz2697	CAT probe FW	TACACCGTTTTCCATGAGCA

cz2698	CAT probe RV	CCTGCCACTCATCGCAGTA
cz2581	tubulin probe FW	CCTTTGGCACAACGTCACCACGG
cz2724	tubulin probe RV	TGACTCGCCGCAACCTCGAT
cz4490	spliced leader oligo (antisense)	CAATATAGTACAGAAACTGTTCTAATAATAGCGTTAGTT

Table 2: List of oligos used for PCR amplification for DNA probes used for Northern blot hybridisation.

RNA-Immunoprecipitation

10^9 procyclic cells which had an in situ N-terminal TAP-tagged ZC3H30 and the second copy of the ZC3H30 open reading frame was replaced by a blasticidin resistance casette were harvested by spinning at 1125xg for 10 minutes. After centrifugation the media was discarded and the cell pellet was resuspended in MEM-Pros medium without FCS. The culture was spread on a 27 cm diameter Petri dish and was UV-cross linked (400mJ/cm^2). The cross linked culture was centrifuged again at 1125xg for 10 minutes and the media was discarded. The cell pellet was resuspended in 1ml of 1x PBS and washed and centrifuged to remove remaining media. Finally the pellet was resuspended in 1ml of tandem affinity purification lysis bufffer (TAP-lysis buffer: 10mM Tris-HCl pH7.8, 10mM NaCl, 0.1% Triton X-100, 5mg/ml Leupeptin, 283mM VRC, 40U/ml RNasin). The cells were lysed by passing through a 21G syringe for 50 times, followed by centrifugation at 16,000xg for 10 minutes. Lysis was confirmed microscopically. After centrifugation, the salt concentration of the clear lysate was adjusted to 150mM NaCl. In the meantime, IgG conjugated Sepharose beads were first washed twice in 1x PBS, and then with TAP-wash buffer (10mM Tris-HCl pH7.8, 150mM NaCl, 0.1% Triton X-100, 5mg/ml Leupeptin, 283mM VRC). The clear lysate was resuspended with the washed beads and incubated for 2 hours at 4°C. After incubation, the unbound fraction was discarded and the eluate was recovered by protease cleavage by Tobace etch virus (TEV) protease in TEV cleavage buffer (TAP wash buffer adjusted to 0.5mM EDTA and 1mM DTT) for 1 hour at room temperature. Samples from total lysate, unbound fraction and eluate were collected from each step for protein and RNA analysis. RNA bound to cross-linked protein was recovered by 20µg Proteinase K treatment at 42°C for 15 minutes. RNA isolation was done using Trifast (peqlab). For RNA-Seq, the RNA preparations from the unbound (Un) fractions and the eluate (E) fractions were rRNA depleted. The preparations were incubated with oligonucleotides complementary to rRNAs and then they were treated with RNaseH. The building of the library and sequencing was done as per standard Illumina protocols. The alignments to the latest available T.brucei TREU927 genome sequence was using Bowtie, allowing for up to 20 alignments and one mismatch. All reads aligning to open reading frames were counted using custom scripts. Differences in the eluate and unbound fraction were determined using DESeq2.

Total RNA Transcriptome

Total RNA from $5x10^7$ procyclic ZC3H30 double knockouts and knockouts carrying and ectopic inducible copy of ZC3H30 were isolated using Trifast (peqlab). 5μg of total RNA was subjected to rRNA depletion using oligonucleotides complementary to rRNAs, followed by RNaseH treatment. The integrity of mRNAs was checked by Northern blotting using a probe that detects the β-tubulin mRNA. The samples were subjected to deep sequencing, and the reads were aligned to the latest available *T. brucei* TREU927 genome sequence using Bowtie, allowing for up to 20 sequence matches. To extract the reads for individual reading frames, we used a modified version of unique open reading frame list from Siegel et al. Differences in mRNA abundance at different conditions were assessed using DESeq2.

Stress techniques

Stress conditions

For stress experiments pertaining to heat shock, oxidative, ethanol, starvation and translation stress, trypanosomes were treated as follows:

- Heat shock- Log phase BS cells were heat shocked at 41°C or 42°C in a pre-warmed water bath for 1 hour. This was followed by centrifugation at 850xg for 8 minutes. The supernatant was discarded and the cell pellet was resuspended in 10ml of HMI-9. The heat-shocked cells were then allowed to recover, grow and proliferate at 37°C. Similarly PCs were mildly heat shocked at 37°C or severely stressed at 39°C or 41°C. To the 37°C heat shocked culture, MEM-Pros medium pre-warmed at 27°C was added and the culture was recovered by letting it grow at 27°C. The severely heat shocked cultures were centrifuged at 850xg for 8 minutes. The supernatant was discarded and the cell pellet was resuspended in 10ml of MEM-Pros medium. The culture was allowed to recover at 27°C.
- For oxidative stress with sodium arsenite, either 10μM or 20μM (final concentration) sodium arsenite was added to log phase procyclic culture at 27°C. After stressing the cells for 1 hour the cultures were centrifuged at 850xg for 8 minutes; the supernatant was discarded and the cell pellet was resuspended in 10ml of MEM-Pros medium and allowed to recover at 27°C. Although the pellet was not washed, to minimise extra stress, the remaining arsenite in the recovery medium was not more than 0.01-0.02μM (final concentration).
- For stress with ethanol, either 1% or 2% (final concentration) ethanol was added to log phase procyclic cultures at 27°C for 1 hour. The cultures were then centrifuged at 850xg for 8 minutes, the supernatant was discarded and the cell pellet was resuspended in 10ml of MEM-Pros medium and allowed to recover at 27°C. Since the pellet was not washed, the remaining Ethanol in the recovery medium was about 0.001-0.002% (final concentration).

- For translational stress, 50μg/ml (final concentration) of Hygromycin was added to log phase procyclic culture at 27°C. Then, the growth and proliferation of trypanosomes were monitored for a period of 72-96 hours.
- For starvation: Log phase procyclic trypanosomes were centrifuged at 850xg for 8 minutes; the supernatant was discarded and the cell pellet was resuspended in 1x PBS at a final density of 10^6/ml for 5 hours. Following incubation the cultures were centrifuged at 850xg for 8 minutes; the supernatant was discarded and the cell pellet was resuspended in 10ml of MEM-Pros medium and allowed to recover at 27°C.

Heat Shock granule preparation and purification

Granules from healthy proliferating or heat-shocked (41°C) PC trypanosomes were prepared as previously published [106, 112]. About $5x10^8$ cells were harvested by centrifugation at 2500rpm for 10 minutes. The pellet was washed once with 1ml of 1x PBS and lysed in 200μl ice-cold buffer A (20mM Tris-HCl pH 7.6; 2mM $MgCl_2$;0.25M Sucrose; 1mM DTT; 10% Glycerol and 1% Triton X-100), supplemented with 5μg/ml of Leupeptin (protease inhibitor). Lysis was confirmed microscopically and the lysate was clarified (20,000xg, 10 minutes). About 50μl of the supernatant (SN1) was taken for western blot, and the rest was discarded. The remaining supernatant was removed by centrifugation (20,000xg, 3 minutes). To the pellet 200μl ice-cold buffer A was added and it was lysed by passing through a 21G syringe for 30-40 times, and then it was vortexed and centrifuged (20,000xg, 5 minutes). A sample for western blot was collected from the supernatant (SN2) and the pellet was resuspended in 200μl ice-cold buffer A. The whole procedure was repeated to obtain the supernatant (SN3), and 50μl for Western Blot. The pellet was resuspended in 200μl ice-cold buffer A and 14.2μl (283mM) of 4M NaCl was added to disrupt the microtubules. The samples were once more passed through 21G syringe for 30-40 times, followed by 30 minutes of incubation on ice with vortexing every 5 minutes. The samples were centrifuged (20,000xg, 10 minutes) and the last soluble fraction, supernatant SN4 was obtained. An aliquot for the Western Blot was taken and the rest of the supernatant was discarded. The pellet was washed in buffer A without re-suspension (20,000xg, 10 minutes), and finally resuspended in 50μl Laemli buffer for Western Blot analysis.

Cytoskeleton-free total protein extraction

$2.5x10^7$ PC or BS form trypanosomes were harvested at 850xg for 8 minutes, and washed once with 1x PBS. The cell pellet was then re-suspended in extraction buffer [1% (vol/vol) IGEPAL in 0.1M PIPES (pH 6.9), 2mM EGTA, 1mM $MgSO_4$, 0.1mM EDTA, supplemented with 10μg/ml Leupeptin and 1x PhosSTOP™ (Roche)]. The cells were lysed by pipetting up and down and centrifuged at 3400xg for 10 minutes at 4°C. The supernatant was carefully transferred into another eppendorf and mixed with 1x volume of 2x Laemli buffer.

Cytotoxicity assay

The half maximal effective concentration of Hygromycin B on procyclic trypanosomes (Wild type-Wt and ZC3H30 double knockouts- DKOs) was measured using a fluorescent probe Resazurin. It is a non-fluorescent; blue coloured component which when metabolically processed by living cells is converted to a pink coloured fluorescent compound called resorufin. Resazurin (Sigma-Aldrich) was used at the final concentration of 44μl.
About 200μl of procyclic form trypanosomes (Wt or DKO) were inoculated on a 96 well plate at a density of $2x10^5$ cells/ml. An appropriate range of drug dosage was selected based on previously published half maximal effective concentration (EC50) [123, 124]. The cells were allowed to proliferate and grow for 72 hours. After 4 hours of incubation with Resazurin, cellular fluorescence was measured using FLUOstar Omerga plate reader (MBG Labtech). Growth inhibition was analysed using Variable Slope model.

Immunofluorescence

For intracellular localisation of N-terminally TAP-tagged ZC3H30 in BS parasite in presence and absence of heat shock (41°C) trypanosomes were centrifuged (850xg, 8 minutes). The pellet was washed in 1ml of 1x PBS at room temperature, followed by resuspension in 2.4% paraformaldehyde in 1x PBS for 18 minutes. The cells were re-sedimented (2 minutes, 6000xg) and washed once in 1ml 1xPBS and resuspended in 100μl 1x PBS and allowed to settle down on a poly-L-Lysine coated glass slide overnight. The cells were then permealised with ice-cold 0.2%- Tritonx-100 in 1x PBS for 20 minutes and blocked with 0.5% Gelatine in 1x PBS for 20 minutes. The cells were incubated with anti-Protein A antibody (1:10,000, anti-mouse IgG, Sigma) for 1 hour. The slides were washed thrice with 1x PBS and stained with anti-mouse IgG conjugated with Cy5 (1:1000) for 1hour followed by washing in 1x PBS. The nucleus and kinetoplast were stained with DAPI (100ng/ml) and the slides were analyses using an Olympus CellR microscope.

Oligos

Part1

Oligo ID	Name	Sequence
cz5185	GW-NOT2-F	5'aaagcaggctccatgaacaatagcaaccagggaac3'
cz5186	GW-NOT2-R	5'gtacaagaaagctgggttcagtcccctcccaccaca3'
cz5229	CAF40GWF	5'aaagcaggctccatgcaccaaacccaagcgtatt3'
cz5230	CAF40GWR	5'gtacaagaaagctgggttcacgcactggcatcacg3'
cz5231	attL1	5'ccccgatgagcaatgctttttttataatgccaactttgtacaaaaaagcaggctccatg3'
cz5232	attL2	5'gggggataagcaatgctttcttataatgccaactttgtacaagaaagctgggt3'

Part2

Oligo ID	Name	Sequence
cz5503	Z305'UTRF(NotI)	5'aatgcggccgcacgagaggaagcaagagggaagg3'
cz5504	Z305'UTRR(XbaI)	5'aattctagagttaaagcggcaccaatctg3'
cz5505	Z30CDSF(XhoI)	5'aatctcgagatgcctcccgctgacaggaac3'
cz5506	Z30CDSR(ApaI)	5'aatgggcccctacgctttttgccgacg3'
cz5673	TetheringZC3H30F	5'aatgggcccatgcctcccgctgacaggaac3'
cz5674	TetheringZC3H30R	5'aatgttaaccgctttttgccgacgtttc3'
cz5810	tethzc3h30-1F	5'-aat gggcccatgcctcccgctgacaggaac3'
cz5811	tethzc3h30-1R	5'aatgttaacatgaatatttttgcaatcttcca3'
cz5812	tethzc3h30-2F	5'aatgggcccatttgccgcgagtatgac3'
cz6137	ups5'UTR-F	5'ATGAACAAGTTCAATGAATGTAA3'
cz6138	inZC3H30CDS Rv	5'CTGTTGACGGAGCATCGGAAGAA3'
cz6139	BLAint rv	5'TGAGGGTGGATTCTTCTTGA3'
cz6140	puro int-rv	5'TGGCGAGGCGCACCGTGGGCT3'
cz6145	bla int fw	5'GTGAATTGCTGCCCTCTGGT3'
cz6146	puro int fw	5'TGCCCGAAGGACCGCGCACCT3'
cz6147	RvZC3H30out3'utr	5'ACAAACGTAAAGGCACATAC3'
cz5815	ZC3H30 in sidecdsF	aatGGGCCCCCCGCGTCACAGCGTTCG
cz6098	ZC3H30KO5'UTRF(XhoI)	AATCTCGAGACACCCCTGAACTACTGTAGACG
cz6099	ZC3H30KO5'UTRR(HindIII)	AATAAGCTTGTTAAAGCGGCACCAATCTG
cz6100	ZC3H30KO3'UTRF(EcoRI)	AATGAATTCGCGGTGGTCCACAGCAATG
cz6101	ZC3H30KO3'UTRF(BamHI)	AATGGATCCCCCGCATACCATCACTG
cz6179	5'UTRZC3hH30-TAP(SacI-NotI)-Fw	AAT**GAGCTCGCGGCCGC**ACACCCCTGAACTACTGTAGACG
cz6180	5'UTRZC3hH30-TAP(NdeI)-Rv	AAT**CATATG**GTTAAAGCGGCACCAATCTG
cz6181	ZC3H30miniORF(HindIII)Fw	AAT**AAGCTT**CCATGCCTCCCGCTGACAGG
cz6182	ZC3H30miniORF(ApaI)Rv	AAT**GGGCCC**TGTTGACGGAGCATCGGAAGAA
cz6611	Tb927.8.3820 RNAi(1,2)-BglII-HindIII-F	AATAGATCTAAGCTTACATCACCTTTCGTGGGAAG
cz6612	Tb927.8.3820 RNAi)(3,4)-SalI-EcoRI-R	AATGAATTCGTCGACGACGAGAATGGGATCGTGTT

Plasmids

Part1

Name	Description
pHD2673	pBD-NOT11(gift by Dr. E.Erben
pHD2674	pBD-NOT2-Gateway cloned ORF
pHD2677	pBD-CAF40-Gateway clonecd ORF

Part2

Name	Description
pHD 2613	N-V5-ZC3H30-bla-5ÙTR(NotI, XbaI)ORF(XhoI,ApaI)
pHD2667	N-lambda-ZC3H30 ORF-myc (ApaI, HpaI)
pHD2703	N-lambda-ZC3H30(5'-1-849bp-3')ORF-myc (ApaI, HpaI)
pHD2704	N-lambda-ZC3H30(5'-850-1692bp-3')ORF-myc (ApaI, HpaI)
pHD2826	Knockout construct-5'UTR(XhoI, HindIII)-Blasticidin resistance-3'UTR (EcoRI, BamHI)
pHD2827	Knockout construct-5'UTR(XhoI, HindIII)-Puromycin resistance-3'UTR (EcoRI, BamHI)
pHD2842	N-terminal TAP-ZC3H30-5'UTR (NotI, NdeI)-Puromycin- mini ORF (HindIII, ApaI)
pHD2950	Tb927.8.3820RNAi- RNAi(HindII,SalI)-Spacer (EcoRI, BglII)-iANR
SK184	in situ C-terminal YFP tagged hypothetical protein (8.3820)- a kind gift from Sussane Kramer

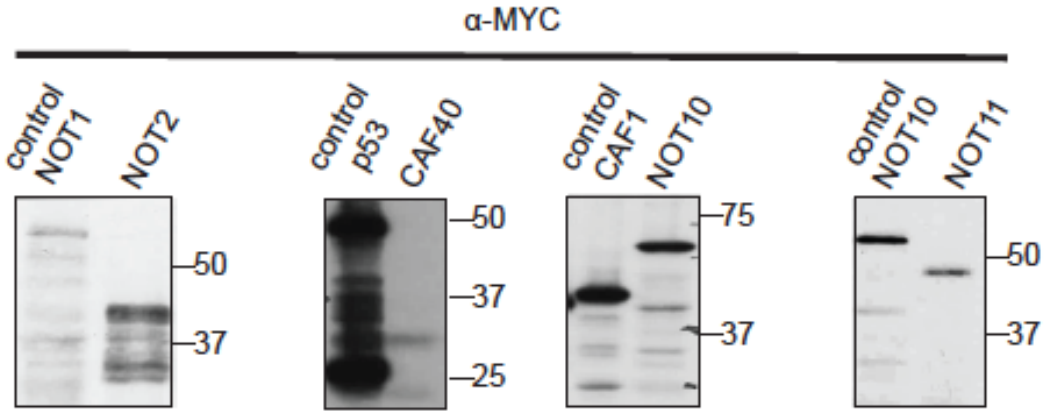
Chapter 3- Results

Part1- Interactors of the CAF1-NOT complex

Yeast transformation with bait constructs, mating efficiency, diploid generation and number of clones obtained

The ORFs of NOT2, CAF40, NOT10 and NOT11 were cloned in pGBKT7 (Clontech), as a GAL4BD-myc-bait. The cloned constructs were transformed in yeast strain AH109. The expression of each bait protein was analysed by Western blotting (Figure 3.1.1 A), using anti-myc antibody. Haploid cells carrying either the bait or the prey were mixed and allowed to mate randomly in an enriched YPD medium, as per the 'Two-Hybrid Library Screening Using Yeast mating' protocol offered by Clontech's Matchmaker GOLD Yeast Two-Hybrid system manual. About 1 million diploids were screened for each mating experiment, so that each clone in the trypanosomal genomic DNA yeast library could be represented at least twice. For each screen, a total of 20 million haploid prey cells or more were mated with about equal number of haploid bait cells. Screens that yielded a mating efficiency lower than 2% were discarded. The positive clones interacting with the baits were selected on high stringency SD plates, with 3-amino triazole (3-AT). Figure 3.1.1 B shows the details of the number of diploid (mated yeast clones expressing both bait and prey protein), and their mating efficiency and number of clones and all of the prey protein encoding genes per bait protein screened.

A



B

SUBUNITS	TOTAL DIPLOIDS (million)	MATING EFFICIENCY (%)	LIBRARY COVERAGE	TOTAL CLONES (number of blue colonies)	PUTATIVE INTERACTORS
CAF40	3.5	6	5	179	37
NOT10	2	5	4	137	311
NOT2	3	3	5	496	47
NOT11	1	6	2	720	99

Figure 3.1.1. Expression of subunits of the trypanosomal CAF1-NOT complex in yeast and screening for their genome-wide interactors. (A). Western blot to show the expression of bait proteins in yeast. The expression of each bait protein was checked using anti-myc antibody (α-MYC). **(B)**. Details of yeast two hybrid screen for each bait subunit. Total number of diploids indicate summation of the number of diploids obtained for at least 2 screens. Mating efficiency is the average mating efficiency for 2-3 screens. The genome wide (**Y**east-**T**wo-**H**ybrid) YTH library of trypanosome fragments consists of 435,000 clones. Library coverage indicates the average number of times the whole library has been represented per screen. The total number of individual blue colonies obtained on quadruple drop-out medium (QDO) is represented by TOTAL CLONES for 2-3 screens. PUTATIVE INTERACTORS are the number of genes identified by deep sequencing for each bait protein.

Class and function enrichment and role in biological processes of the identified genes from screen

The list of genes that have a RPM > 100, number of reads ≥ 10, and were not identified using multiple other baits used for screening the genome wide library is provided in the published article [93]. I obtained a total of 47, 311, 99 and 37 interacting partners for NOT2, NOT10, NOT11 and CAF40. I first looked for genes, which were identified in DeepSeq data for all the subunits of the CAF1-NOT complex. I used Venn Diagram, a tool from UGent, where I uploaded gene lists for putative interacting partners for each bait protein. Figure 3.1.2 C shows total number of common genes identified in the screen of CAF40, NOT2, NOT10 and NOT11. Considering all interactors, I used the on-line Gene ontology tool to find out what GO terms were enriched with respect to protein class and molecular function (Figure 3.1.2-A, -B). Since a large number of trypanosomal genes have not been functionally annotated, the tool could use only 243 genes for functional analysis. The class of proteins largely represented among the putative interactors was Nucleic acid binding (19%), followed by hydrolases (18%) and transferases (17%). Although transmembrane proteins, receptors, chaperones, are amongst the most abundant proteins found in a cell, their percentage in summation was not higher than 5%. Among the molecular functions, the most enriched ones were catalytic activity (63%) and binding activity (26%). I also wanted to see if the genes identified in the screens had any role in specific biological process. Hence, I compared the gene list from my screen with proteomics data for RNA-binding proteins (RBPs) identified on polysome pull downs, tethering and RBPome screen to identify potential gene expression regulators, proteomics data for differentiation, starvation granule proteomic data [94,

95, 112, 125, 126]. Figure 3.1.2 D depicts a Venn digram (diagram made using UGent Venn diagram tool) shows the number of genes obtained in the biological processes of RNA biology, stress and differentiation, with or without overlaps. I found 13 genes associated with all three biological processes. Interesting candidates from this list are two novel hypothetical proteins with gene IDs Tb927.6.5010 (partner of NOT10 and NOT11) and Tb927.10.15360 (partner of NOT10) and ZC3H22 (Tb927.7.2680), a protein having zinc-finger domain. Interestingly, the hypothetical protein expressed by Tb927.6.5010 is also known to bind to RNA. The putative interacting partner of NOT10 and NOT11, ZC3H22, represses reporter mRNA expression and is also upregulated upon stress and differentiation. The interacting partner of NOT2, zinc finger protein ZC3H30 also binds to RNA [94], and is enriched in starvation stress granules (see below). Also, I found a hypothetical protein, Tb927.9.7690, which is supposedly a positive regulator of mRNA, binds to mRNA, and the quantity of protein is relatively higher in stumpy forms.

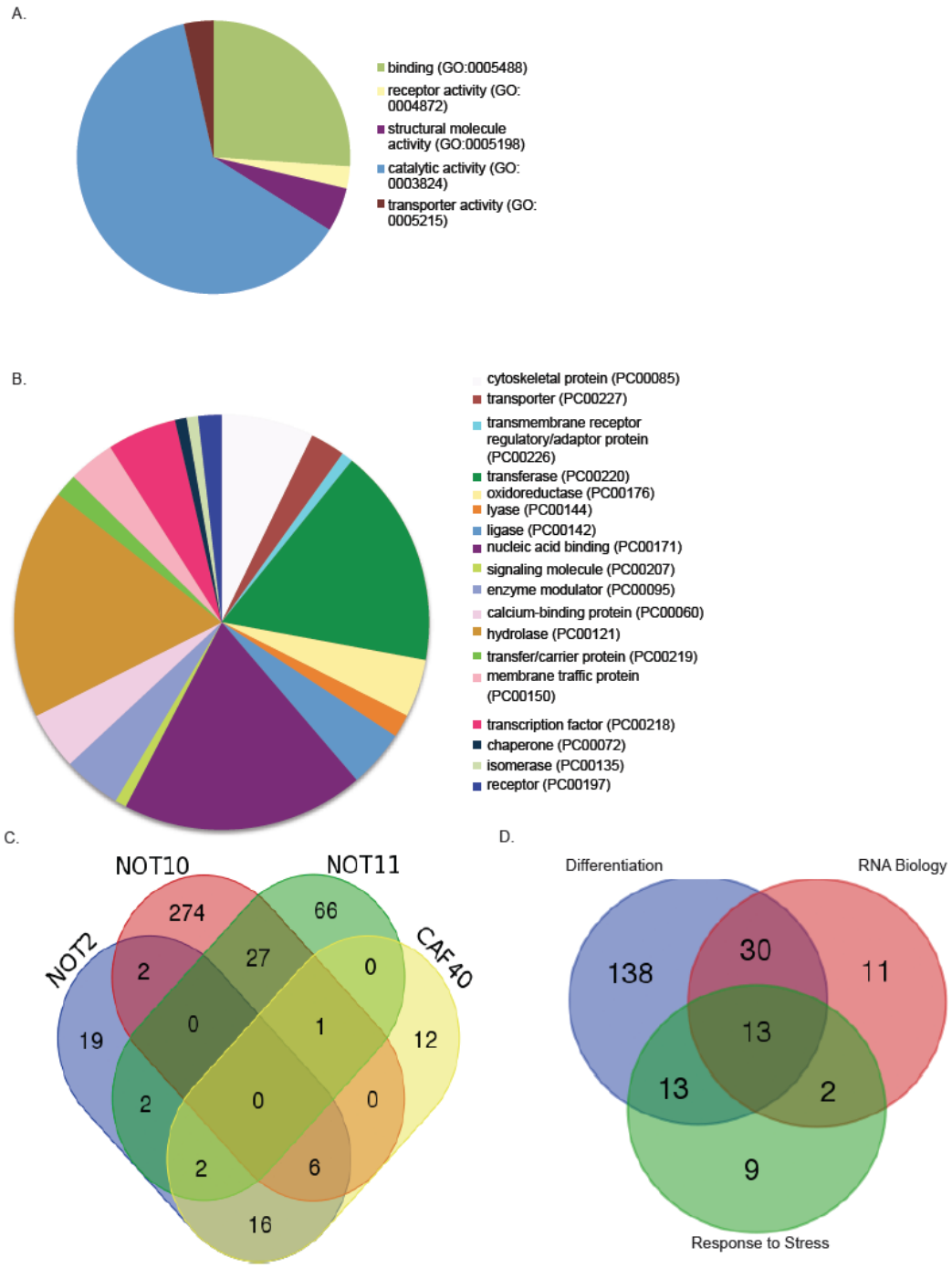


Figure 3.1.2. Gene classes and functions enriched in the Yeast-two-hybrid screen, and Venn diagram representation of common genes and biological processes they might be involved in. (A). and **(B)**. Pie charts showing GO classification of the putative interacting partners obtained from the screen with respect to the molecular function (A) and classes of protein (B). **(C)**. Common genes identified in the individual screens of CAF40, NOT2, NOT10 and NOT11. **(D)**. Genes identified in the Yeast-two-hybrid screen with possible biological function.

Interactors identified in the Yeast-two-hybrid screen

(Analysis done by Prof. Christine Clayton)

More stringent criteria to find interactors were applied to exclude amplification artefacts. Open reading frames for which there were at least two different in-frame sequences represented by at least 20 reads were selected. This gave 6,158, 15 and 3 interaction partners for NOT2, NOT10, NOT11 and CAF40 respectively. The possible specific interactions for NOT2, NOT11 and CAF40 are shown in Table 3.1. In order to find out whether these interactors were specific or random the list of interactors was compared to the yeast-two-hybrid list of interactors for 4E-IP (translational repressor), Tb927.7.2780 (a hypothetical protein affecting the stability of bound mRNA positively), and centromere associated protein CFB1. It appears that many of the interactors pulled down by the yeast-two-hybrid screen are common to all screens (eg., POMP37 and Tb927.8.7960). The localisation of the proteins was determined using available online data from http://tryptag.org [127]. Since the number of putative partners of the subunit NOT10 were too numerous, I did not follow them up. In fact, the only interaction that can confidently be labeled as specific is the interaction between NOT10 and NOT11. Also, there was no enrichment in protein function after the data analyses. I therefore decided to focus on the RNA binding protein ZC3H30 in the second part of my thesis.

Tb927	Function	NOT2	NOT10	NOT11	CAF40	4EIP	7.2780	CFB1	Loc
7.7210	POMP37	2	12	0	0	0	y	0	?
8.7960	unknown	2	0	0	2	y	0	0	
11.16280	Ribosomal protein L2 L8	2	0	0	0	y	0	0	nuc, cyt
11.13310	spectrin repeat protein	2	0	0	0	y	0	0	nuc
8.1960	**NOT11**	0	10	2	0	0	y	0	
7.2070	DNAJ	0	2	6	0	0	0	0	
9.10370	TAX-1, flagellar protein	0	4	5	0	0	0	0	cyt
11.7040	pterin enzyme	0	2	4	0	0	0	0	
5.2440	WW domain	0	19	2	0	0	0	0	
6.860	unknown	0	14	4	0	0	0	0	
1.3290	unknown	0	7	2	0	0	0	y	
9.1560	DNAj domain	0	0	2	0	0	0	0	
10.13360	Mitochondrial EF-Tu	0	0	2	0	0	0	0	mit
10.4930	protein phosphatase	0	0	2	0	0	0	y	cyt
3.1880	unknown	0	0	2	0	0	0	y	cyt

Table 3.1: Most specific interaction partners of NOT2, NOT11 and CAF40 [93].
Key: Tb927- Gene ID; Function- possible annotated biological function; NOT2/10/11/CAF40/4E-IP/7.2780/CFB1- protein subunits used for Yeast-two-hybrid screen with shotgun yeast genomic DNA library; loc-localisation of the protein, nuc (nuclear) or cyt (cytoplasmic); numbers- number of different locations from which reads (>20) were obtained in the screen. "y" - positive in the screen with 4EIP, 7.2780 or CFB1.

Part 2 - The role of ZC3H30 in combating stress in *Trypanosoma brucei*

Sequence analysis of ZC3H30

The RNA binding protein Zinc finger 30 (ZC3H30) is located on chromosome 10. Its gene identification number is Tb927.10.1540. The gene encodes an mRNA of length greater than 5 Kb which has a very long UTR of 3.269Kb. The gene encodes a 564 amino acid long protein, which has 2 zinc finger domains in the N-terminal half and has 3 low-complexity regions (LCRs). The Zinc fingers are of CCCH type (C-$X_{6–14}$-C-$X_{4–5}$-C-X_3-H). The CCCH zinc finger motif is known to play important roles in RNA processing as RNA binding proteins (RBP) [128, 129]. The protein sequence is also spanned with LCR, which have been known to promote protein-protein interactions are found in proteins involved in almost all kinds of biological processes including, stress, nucleic acid binding translation etc [130]. Figure 3.2.1 shows the location of the LCRs and the zinc finger domains on the primary sequence of ZC3H30. The Zinc finger domains were manually searched on the primary sequence and the low complexity regions were found by running the protein fasta sequence in SMART [131].
Below is the protein sequence of ZC3H30

N terminal-
MPPADRNDIEEERVSRWCVTDEERPIPPDSYLSPNGALTVATNAVSPPPY
CSVQRVVHPEGPVHHFVDSDSPGCGYDDKFGAGDDEVKGGELYMWV
DPHSRKLRVPLSLMVPTHATANKGALPSLCLSFLEGHCRHEWCRQAHV
PHVLPMLRQQALNAPTCCRLHKDPHSTTELTDRFKFIRLVGNEGSYSSR
TNGEQSDLIPAERVALTVGLQLIIAQSAQAAKEKEEQSQNLGETNQGAA
NGAKVERKDVLDVPAKDVCRLHISRLCRYVEDCKNIHICREYDLQLPPP
PNLVCLLNGITPSMTMINIGERSYSSTMLSLGDVTDEVFNIICDQQRRSV
MNNTPAARMTPASQPYQHANQVVSCANPGVDAAINSASPLSLGDANLY
TAPRGGTSPASCGMGDATPFAPASQRSLGQNGSPNLGAAAAGSSNVTR
VLRIFDVRPKSAGEGGRERGSTHRGHNTNGNGSAGASPFLNGRGSESAS
RSRNGASVSYNNYTCPNSKNGSGNGCAAGAGVPPCGINGSDRNACSGS
HRRGSRHLTGQSNNSGRGSAKRRQKA- C terminal

Zinc fingers are marked in red-CCCH- C-$X_{6–14}$-C-$X_{4–5}$-C-X_3-H X-any amino acid
2 CCCH domains- 1:128-145; 2:265-283
Low complexity regions are gray total 3 LCRs- 1:217-236; 2:290-302; 3:507-526

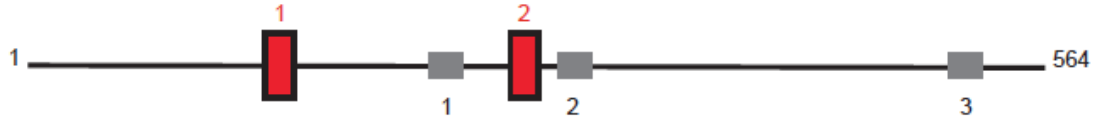


Figure 3.2.1: A schematic diagram of the protein sequence oft he protein ZC3H30. The entire primary sequence is represented by the black bar; 1 represents the first amino acid and the N-terminus and 564 represents the last amino acid and the C-terminus of the protein

ZC3H30. The red bricks are the two Zinc finger domains of CCCH type and the gray boxes are the Low Complexity Regions (LCRs).

Full length ZC3H30 can repress CAT reporter mRNA expression

The repressive activity of the complete ZC3H30 protein was found in a mini-library tethering screen [94]. In addition, the screen using tethered random genomic DNA fragments showed that the segments starting at amino acid 276, 279 or 290 were repressive [95]. Since screen results can be misleading, I first used the tethering assay to confirm the effect of full-length ZC3H30 protein on bound mRNAs. Cell lines expressing a constitutive chloramphenicol acetyl transferase (*CAT*) reporter gene were used. In one cell line the *CAT* mRNA contained 5 copies of λ-boxB hairpin sequences between the resistance reporter and Actin 3' UTR (ACT). The other cell line expressed identical *CAT* mRNA without the boxB (B) sequence at its 3' end. A construct expressing λN-transcriptional anti-termination protein [132] fused N-terminally to ZC3H30, which was C-terminally myc-tagged was cloned. Then this construct was transfected in the both the above cell lines. The λN protein is capable of binding the BoxB sequence [132], and depending on the effect of the fused protein on mRNA, the reporter expression will be increased, decreased or unaffected. Cell lines lacking the λN-ZC3H30-myc construct were used as controls.

Figure 3.2.2 A depicts a schematic representation of λN-ZC3H30-2xmyc construct and B represents the reporter mRNA constructs with respect to tethering assay. Figure 3.2.2 C shows the effect of ZC3H30 on *CAT* mRNA and protein in reporter fused with boxB and also the cell line with *CAT* reporter without the boxB. For Northern blot (NB) and Western blot (WB) respectively, bloodstream form trypanosomes were harvested and RNA and protein were extracted from cell lines co-expressing *CAT-5xB* or only *CAT* mRNA with λN-ZC3H30-2xmyc protein. In the Northern blot (NB), the cell line expressing *CAT-5xB-ACT* co-expressing λN-ZC3H30-myc protein has very low levels of the reporter mRNA, compared to the cell line which does not express the λN-ZC3H30-2xmyc protein. The *CAT* mRNA and protein levels in the cell lines without BoxB remain almost the same in presence or absence of λN-ZC3H30-2xmyc protein (Figure 3.2.2B-WB and NB). Although it's a tetracycline (tet) inducible cell line the expression is extremely leaky (Figure 3.2.2B-WB, α-myc). What is clear from the blots (NB and WB) is that cell lines expressing the construct *CAT-5xB*, have considerable lower CAT mRNA and protein in presence of λN-ZC3H30-2xmyc protein. The CAT enzyme activity was measured by CAT reporter assay (scintillation measurement): the CAT-5xB cell-lines co-expressing λN-ZC3H30-myc protein showed much lower CAT activity (<10% of control). Figure3.2.2 D shows the relative CAT activity against CAT mRNA abundance for 2 individual experiments. This indicates that when ZC3H30 is tethered to the reporter mRNA, it is able to repress the expression by mRNA degradation and/or translational repression.

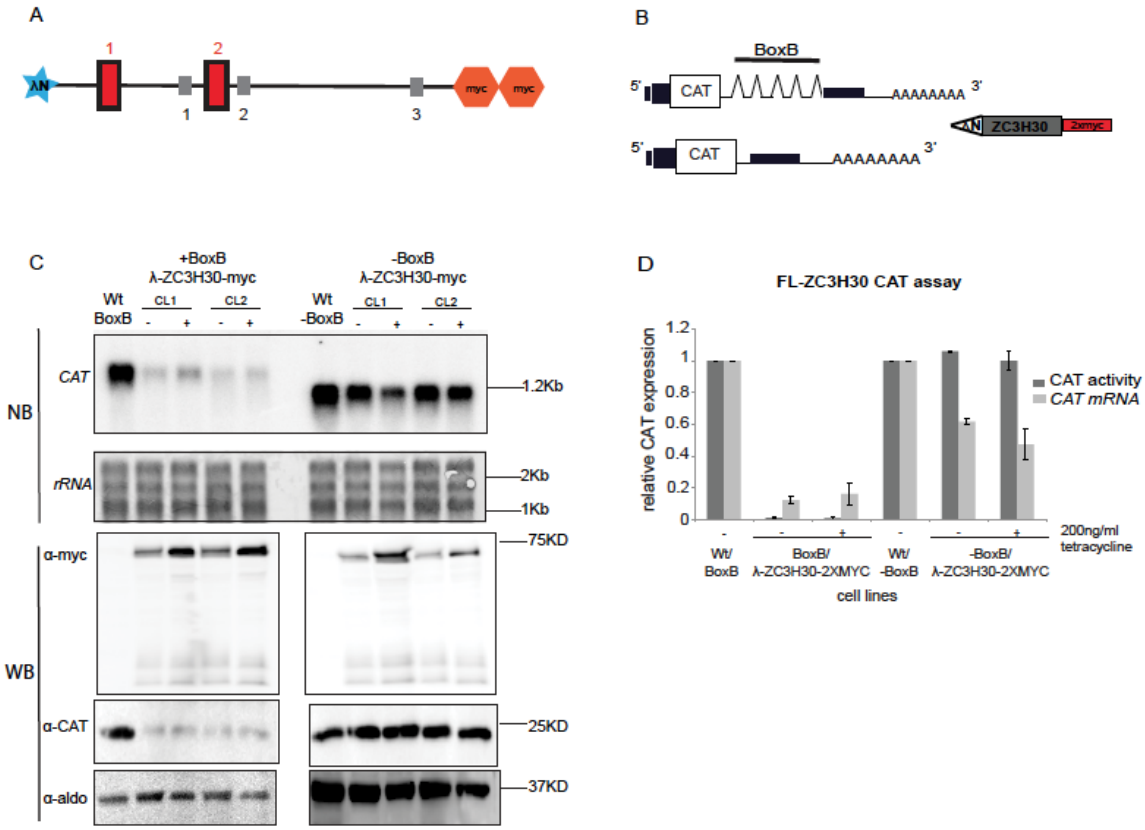


Figure 3.2.2: ZC3H30 can destabilise and degrade tethered reporter mRNA in its full-length form. (A). A schematic representation of the λN-ZC3H30-2xmyc constructs used in tethering assay. The N-terminal of ZC3H30 is fused to λN peptide, and the C-terminal is fused with two myc tag peptides. The primary structure of ZC3H30 is spanned with two Zinc fingers (red bricks) at the N-terminal, and it has two central and one C-terminal low complexity regions (gray boxes - LCRs). **(B)**. A schematic representation of the tethering constructs to investigate the fate of a reporter mRNA, when ZC3H30 is tethered to it. **(C)**. Effect of full-length λN-ZC3H30-2xmyc when tethered to CAT reporter. BS cells co-expressing CAT-5xBoxB and λN-ZC3H30-2xmyc or CAT and λN-ZC3H30-2xmyc were cultured to isolate total RNA and extract total protein. Effect on *CAT* mRNA was analysed on a Northern Blot (NB). Quantity of *CAT* mRNA (+*BoxB*- 1.5Kb or -*BoxB* -1.2Kb) was normalised to the lowest *rRNA* band. Expression was checked on a Western blot (WB). The blot shows the expression of ZC3H30 is leaky. Expression was checked by anti-myc for ZC3H30 (α-myc) antibody; anti-aldolase (α-aldo) was used as loading control; expression of CAT was checked using anti-CAT (α-CAT) antibody. **(D)**. Relative CAT activity with respect to relative abundance of *CAT* mRNA in cells expressing λN-ZC3H30-2xmyc construct and constitutive expression of CAT mRNA with (BoxB) or without BoxB (-BoxB) in the 3' UTR. The expression of λN-ZC3H30-2xmyc was induced using 200ng/ml of tetracycline. The graph represents an average of two experiments on the effect of induced (+) or uninduced (-) λN-ZC3H30-2xmyc on reporter CAT mRNA. Relative CAT expression for cell lines expressing CAT reporter mRNA (Control cell lines: *CAT-5xBoxB*/*CAT*) is 1 and CAT expression of cell lines co-expressing *λN-ZC3H30-2xmyc* is normalised to the controls.

Both N-and C-terminal half of ZC3H30 can decrease reporter mRNA expression upon tethering

Next, I decided to find out which domain or fragment in ZC3H30 is responsible for inflicting repression on the reporter mRNA. Since ZC3H30 was known to bind and repress reporter mRNA expression, I believed that either the zinc fingers or the LCRs interact with mRNA degradation or repression machinery. Hence I decided to clone deletion mutants. The N-terminal mutant had both the zinc-finger domains, and two LCR, and the C-terminal domain had one LCR (Figure 3.2.3 A –bottom diagrams, left-N-terminal half-F1 and right- C-terminal half-F2). Both the cell lines expressing λN-F1-2xmyc or λN-F2-2xmyc could decrease *CAT-5xBoxB* reporter mRNA, but only the expression of λN-F2-2xmyc could reduce CAT protein (Figure 3.2.3 B and C). The control cell lines used here were the ones like Figure 3.2.2 C. In the control cells, CAT-5xBoxB cell line without any deletion mutant construct and cell lines which expressed the CAT construct without the boxB showed no anomaly in the *CAT* mRNA or protein abundance. Also, the cell line expressing *CAT* without boxB, carrying the deletion mutants showed no effect either on *CAT* mRNA or on CAT protein.

This indicates that either deletion mutant can interact with the trypanosomal mRNA degradation machinery but only the C-terminal fragment might have an additional effect on mRNA translation. It might be so, that the LCR in the C-terminal half interact with proteins, which inhibit translation. In order to have a clear idea as to which domain is responsible for repression by destabilisation of mRNA, and which one represses mRNA by translation inhibition only, one has to clone each Zinc finger domain and each LCR separately and test the activity of CAT protein by tethering assay. Also, another rationale that explains the repressive activity of the C-terminal protein is if the C-terminal half is capable to dimerise with the endogenous protein. Then the dimer will repress the reporter gene expression both by translation inhibition and mRNA destabilisation. RBPs are known to bind mRNAs and we do not know if the mRNA binding character of ZC3H30 is specific or ambiguous. If endogenous ZC3H30 binds to the actin 3'UTR then it would also be capable of repressing the CAT mRNA by degradation and translational repression. But it seems unlikely to happen; otherwise the fragment F1 would also be capable of translational repression.

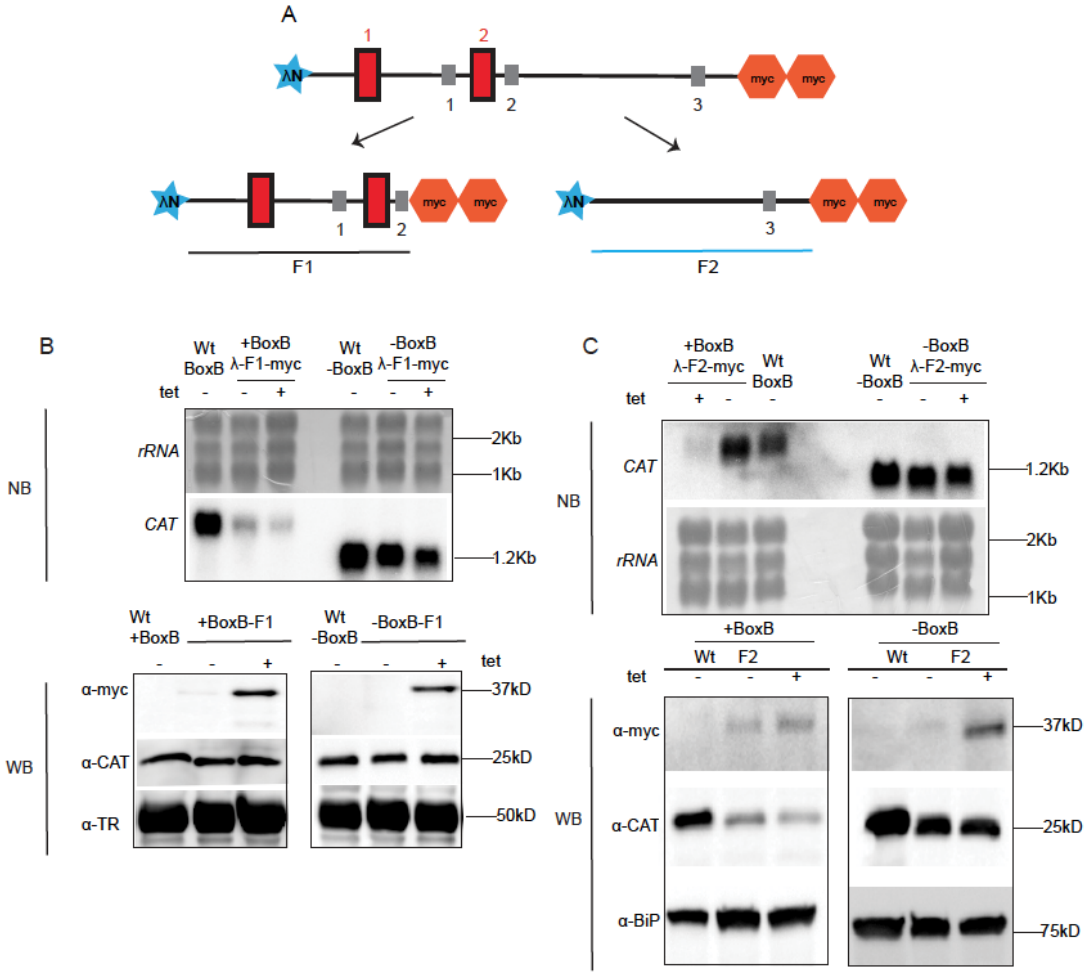


Figure 3.2.3: Both N- and C-terminal halves of ZC3H30 can lower CAT mRNA abundance, but only C-terminal half of ZC3H30 can reduce abundance of CAT protein. (A). Schematic representation of the deletion fragments, F1 and F2 with respect to the full-length protein. The thin black and blue bars represent the part of the protein the deletion fragments, F1 and F2 respectively. The thick black bar represents primary sequence of ZC3H30 and the red bricks and gray boxes represent zinc fingers and LCRs. (**B).** and **(C).** Effect of λN-F1/F2-2xmyc when tethered to CAT reporter. The effect of the deletion mutants on CAT reporter mRNA and protein were studied as in 3.2.2C. (This experiment needs to be replicated before final conclusion).

Localisation of ZC3H30

RNA binding proteins involved in post-transcriptional regulation via decay or translation are expected to be localised in the cytoplasm, while those implicated in regulation of mRNA export or splicing are expected to be at least partially in the nucleus. The localisation of zinc finger RNA binding protein, ZC3H30 was determined by using TAP-ZC3H30 bloodstream (BS) form cells, using cell fractionation and differential centrifugation. Cells expressing ZC3H30 with an N-terminal tandem affinity purification (TAP) tag were harvested by centrifugation and the membrane was disrupted using the abrasive SiC in isotonic buffer. The cytoplasmic soluble fraction was separated from the organellar fraction by

centrifugation. The organelles and nucleus remain intact, leaving the soluble proteins in the cytosolic fraction. The pellet obtained was then lysed in presence of detergent and the suspension was re-centrifuged. The supernatant contained soluble organellar and nuclear proteins, and the pellet contained proteins assoiciated with larger structures. Figure 3.2.4A, is a Western Blot showing the localisation of TAP-ZC3H30. It was found in the cytoplasmic fraction. The cytoplasmic marker trypanothione reductase (TR) and exoribonuclease D (XRND), a nuclear marker, were used as controls.

I also analysed the localisation of TAP-ZC3H30 in BS trypanosomes by immunofluorescence. Figure 3.2.4 B shows localisation of TAP-ZC3H30 at 37°C and upon heat shock 41°C. TAP-ZC3H30 was found to be in the cytoplasm. The staining of TAP-ZC3H30 was more intense upon heat shock.

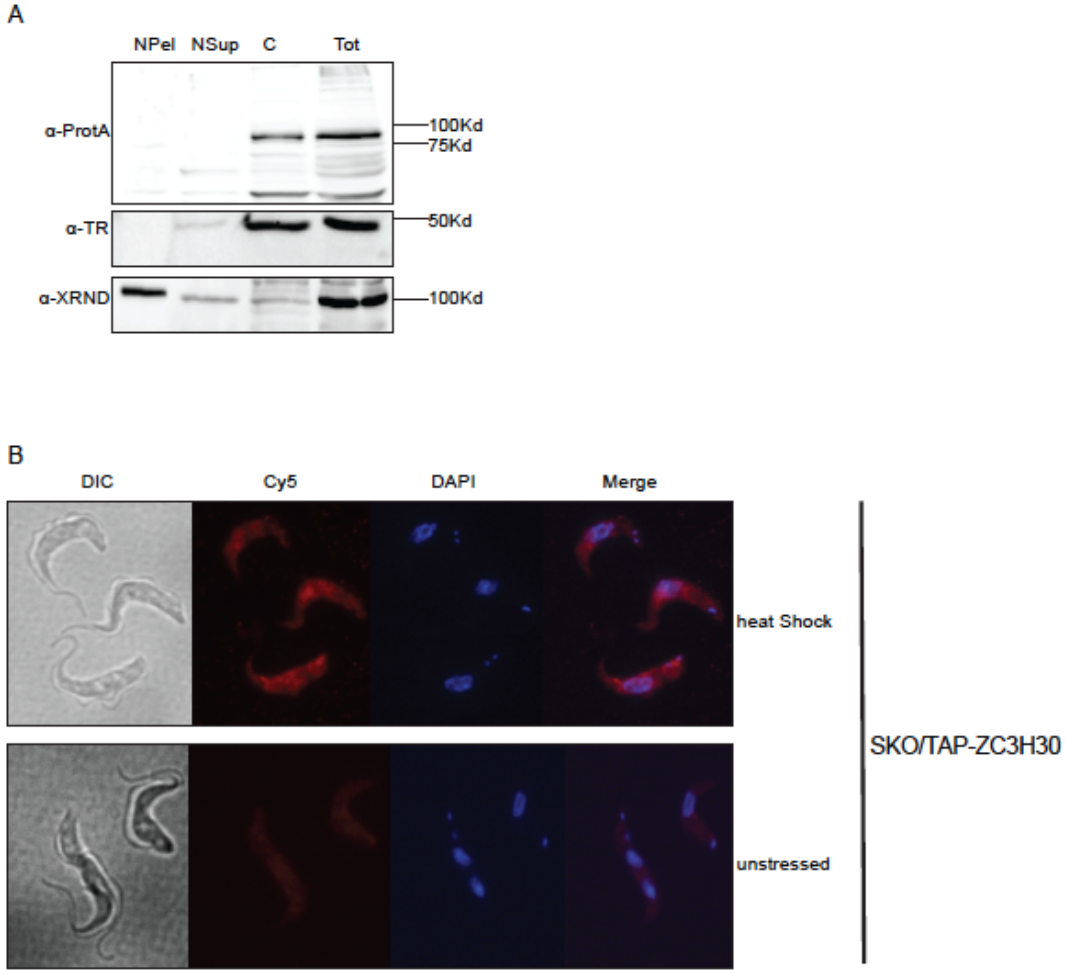
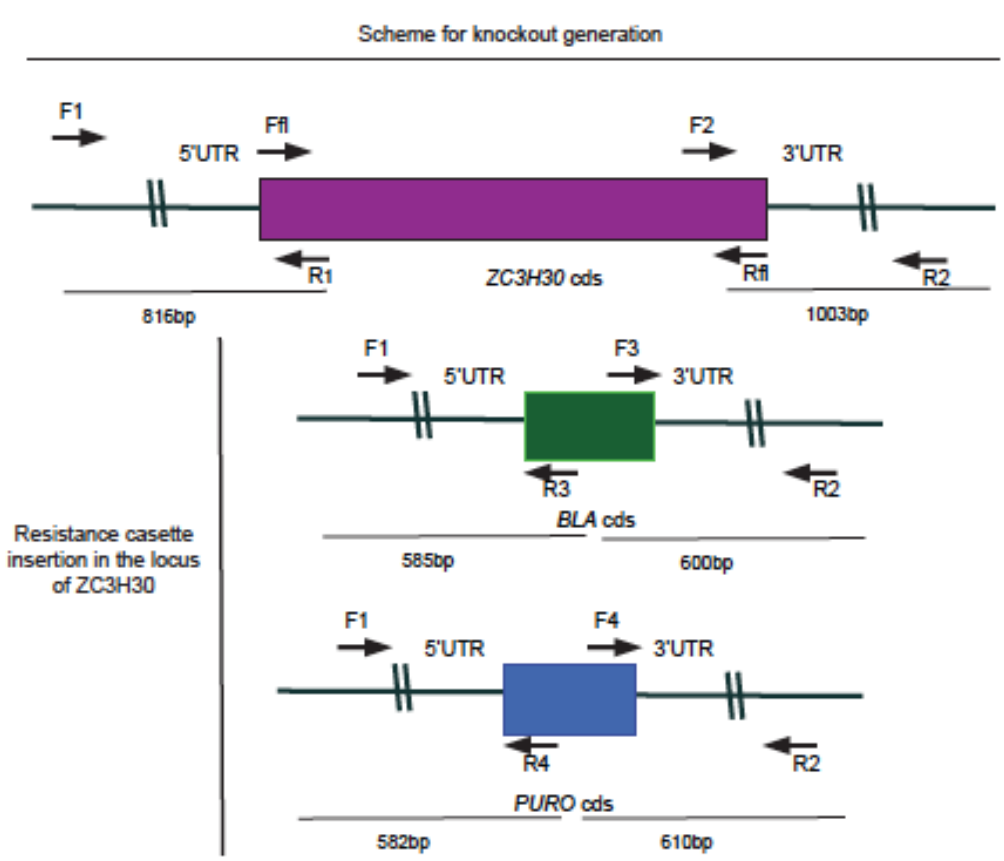


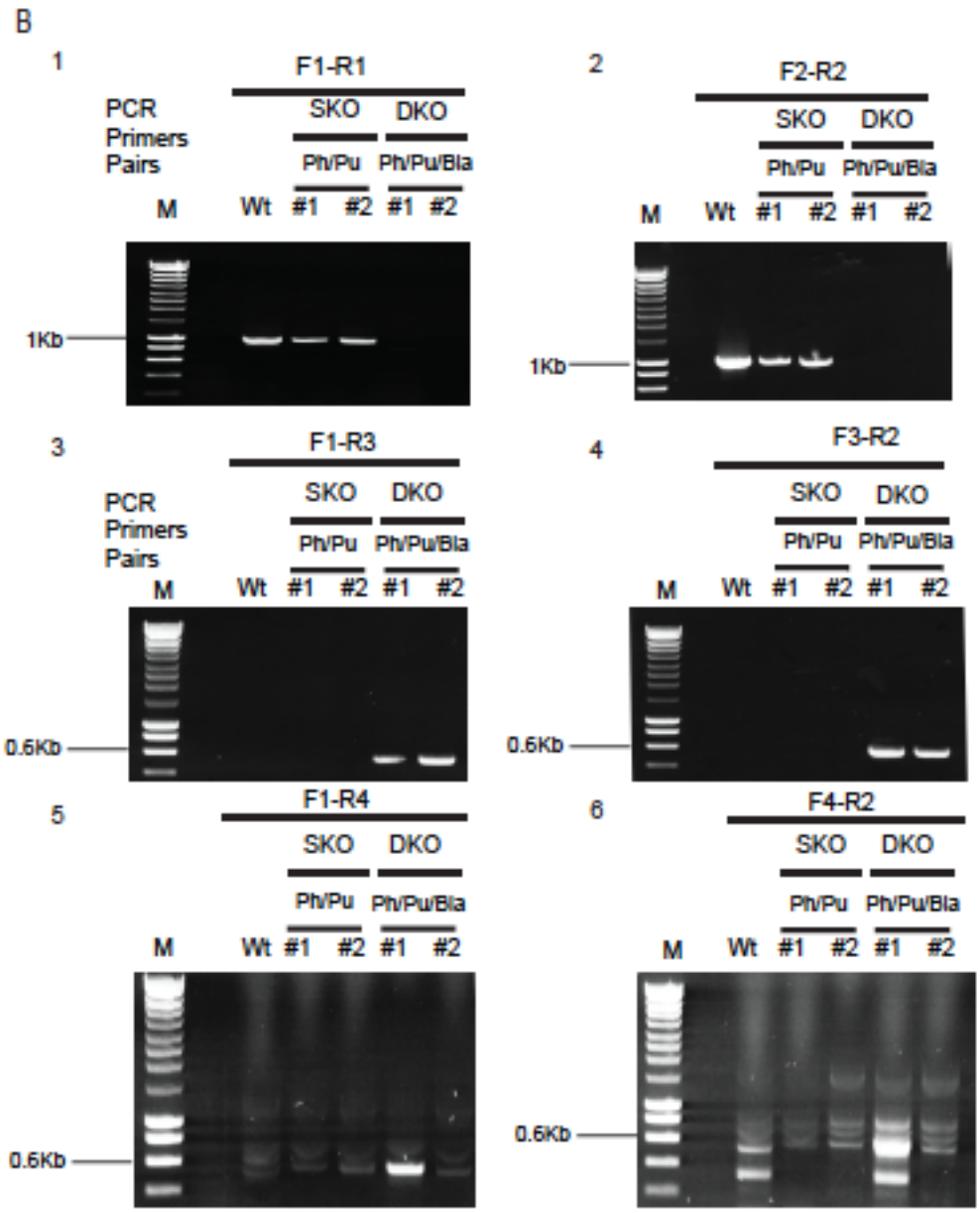
Figure 3.2.4: ZC3H30 is a cytoplasmic protein. (A). SiC cell fractionation assay. Bloodstream form cells expressing an *in situ* N-terminally TAP- tagged ZC3H30 were fractionated by differential centrifugation. Equal cell-equivalents of sample from the total lysate (Tot), cytoplasmic (C) and organellar fraction (detergent-insoluble organellar and cytoskeletal pellet -NPel; detergent-soluble - NSup) were analysed by Western blotting. The tagged protein was identified by anti-Protein A antibody (ProtA), and it localised in the same fraction as the cytoplasmic marker trypanothione reductase (TR), and not with XRND, which, is a nuclear marker. **(B)**. Immunofluorescence of bloodstream-form cells expressing TAP-ZC3H30 (labelled with Cy5) with or without heat shock at 41°C. DAPI labels the nucleus and kinetoplasts.

Generation of ZC3H30 double knockouts

Next the question was, whether the expression of ZC3H30 gene is essential for trypanosomes under normal growth conditions. Knock-out cell lines were made as previously published ([133, 134]) both in procyclic and bloodstream form cells. Blasticidin (*BLA*) and Puromycin (*PURO*) resistance cassettes replaced the endogenous alleles of *ZC3H30*. The schematic diagram for *ZC3H30* knockout generation in bloodstream and procyclic form trypanosomes is given in Figure 3.2.5 Knockouts were confirmed by genomic DNA isolation and PCR analyses. To check if the endogenous *ZC3H30* gene has been removed I did a PCR with the primer pairs F1-R1, and F2-R2. The primers F1 and R2 anneal to the genomic DNA sequences outside the cloned 5' and 3' UTR respectively. The primers R1 and F2 anneal to sequences within the coding DNA sequence (CDS) of *ZC3H30*. PCR followed by agarose gel analyses shows that *ZC3H30* gene was removed from the endogenous locus both in the bloodstream and procyclic form double knockouts (DKO) and not in the single knockouts (SKO) (Figure 3.2.5 B-1,2, C-5,6). Also a PCR was done in the procyclics to ensure that the gene was not relocated elsewhere in the genome (Figure 3.2.5 C7) by using primers, which anneal at the beginning (Ffl) and the end (Rfl) of the *ZC3H30* open reading frame (ORF). Here, only the DNA from the wild type (Wt) cells gave a ≈1.7Kbp PCR product (*ZC3H30* CDS-1695bp). PCR products were obtained for DKOs (Figure 3.2.5 B3-6 C1-4) with the primer pairs F1-R3, F3-R2, F1-R4 and F4-R2. R3 and F3 are primers annealing to the genomic DNA of *BLA* resistance cassette and R4 and F4 anneal to the genomic DNA of *PURO* resistance cassette. A PCR using primers annealing to a different region in genomic DNA was carried out to ensure that the genomic DNA isolated from the procyclic form cells was not degraded (Figure 3.2.5 C8).

A


Scheme for knockout generation
F1
5'UTR
Ffl
F2
3'UTR
R1
ZC3H30 cds
Rfl
R2
816bp
1003bp
Resistance casette insertion in the locus of ZC3H30
F1
5'UTR
F3
3'UTR
R3
R2
BLA cds
585bp
600bp
F1
5'UTR
F4
3'UTR
R4
R2
PURO cds
582bp
610bp


B
1
F1-R1
PCR
Primers
Pairs
SKO
DKO
Ph/Pu
Ph/Pu/Bla
M
Wt
#1
#2
#1
#2
1Kb
2
F2-R2
SKO
DKO
Ph/Pu
Ph/Pu/Bla
M
Wt
#1
#2
#1
#2
1Kb
3
F1-R3
PCR
Primers
Pairs
SKO
DKO
Ph/Pu
Ph/Pu/Bla
M
Wt
#1
#2
#1
#2
0.6Kb
4
F3-R2
SKO
DKO
Ph/Pu
Ph/Pu/Bla
M
Wt
#1
#2
#1
#2
0.6Kb
5
F1-R4
SKO
DKO
Ph/Pu
Ph/Pu/Bla
M
Wt
#1
#2
#1
#2
0.6Kb
6
F4-R2
SKO
DKO
Ph/Pu
Ph/Pu/Bla
M
Wt
#1
#2
#1
#2
0.6Kb

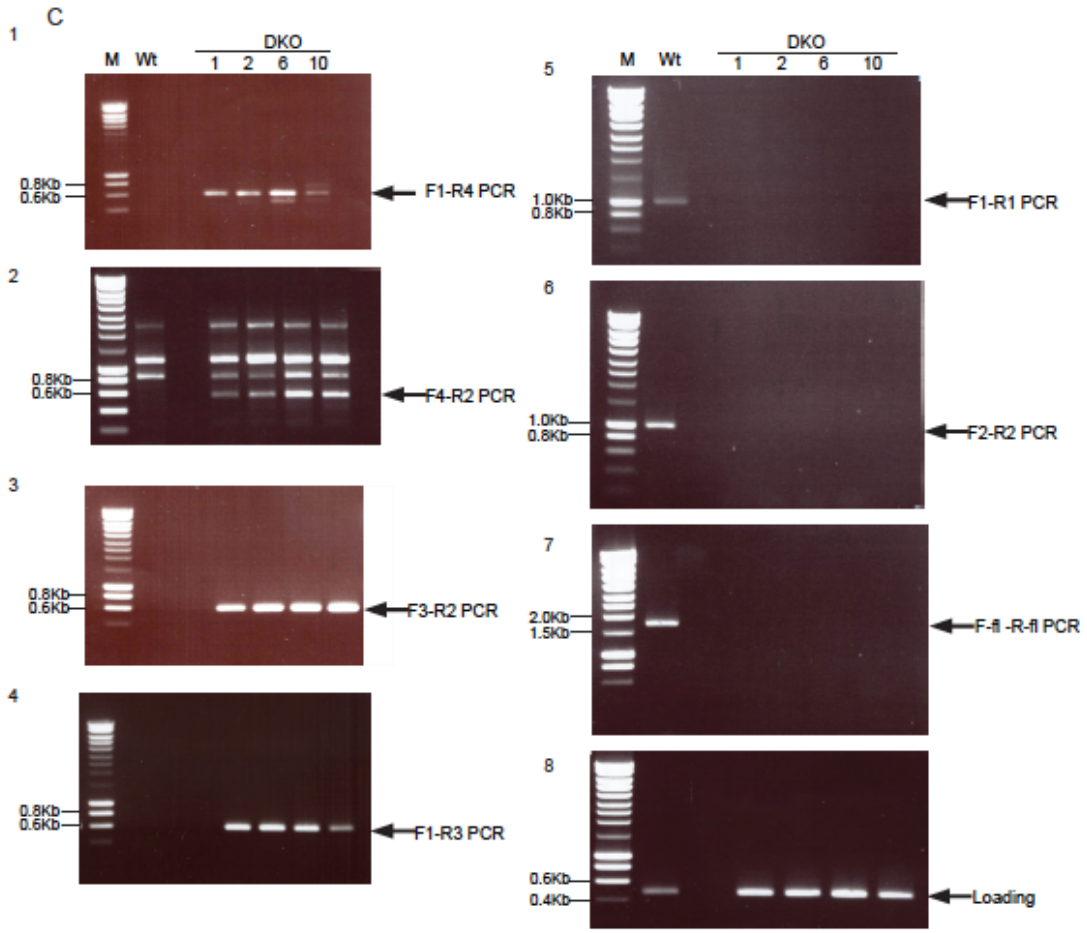


Figure 3.2.5: Demonstration of elimination of *ZC3H30* open reading frame (ORF) in procyclic (PC) and bloodstream form (BS) trypanosomes by PCR analyses. (A). A schematic representation of ZC3H30 alleles, and the resistance cassette used to replace them by homologous recombination, the primers' location (small arrows) and the resulting PCR product, to confirm *ZC3H30* knockouts in bloodstream form parasites (B) and Procyclic trypanosomes (C). **(B)** and **(C)**. Agarose gel analyses to show the generation of *ZC3H30* double knockouts by PCR in BS and PC. The clones for ZC3H30 SKO and DKO in the BS are marked as #1 and #2. The clones in the PC are marked as 1, 2, 3 and 4.

Southern blot to confirm ZC3H30 gene deletion in procyclic cells

A Southern blot was performed to confirm gene deletion in PC. Genomic DNA was isolated from double knockouts and wild type cells. The genomic DNA was digested with *Apa*I and *Sal*I. 10μg of digested genomic DNA was loaded on 0.8% agarose gel. The gel was run for 4 hours and the DNA in the gel was fragmented, denatured and neutralized before being set to transfer on a nylon membrane. Figure 3.2.6 A shows the schematic diagram of genomic DNA digestion at ZC3H30 locus, and binding site of ZC3H30 gene specific probe. The PCR probe used for radiolabelling was a 300bp fragment amplified from the open reading frame of *ZC3H30* gene. As expected upon hybridisation with the probe, a 1Kb band was observed only in wildtype procyclic cells but not the knockout (ZC3H30$^{-/-}$) clones (Figure 3.2.6 B). Also, no other band was obtained for ZC3H30 probe on a southern Blot, confirming the PCR results.

Thus the gene was truly deleted and was not relocalised elsewhere. A probe against *β-Tubulin* gene (Tb927.1.2370) was made and hybridised to ensure the equal transfer of genomic DNA.

The successful generation of the knockouts indicated that the expression of ZC3H30 is not required for growth and survival under normal growth conditions either in the bloodstream or in the procyclic form.

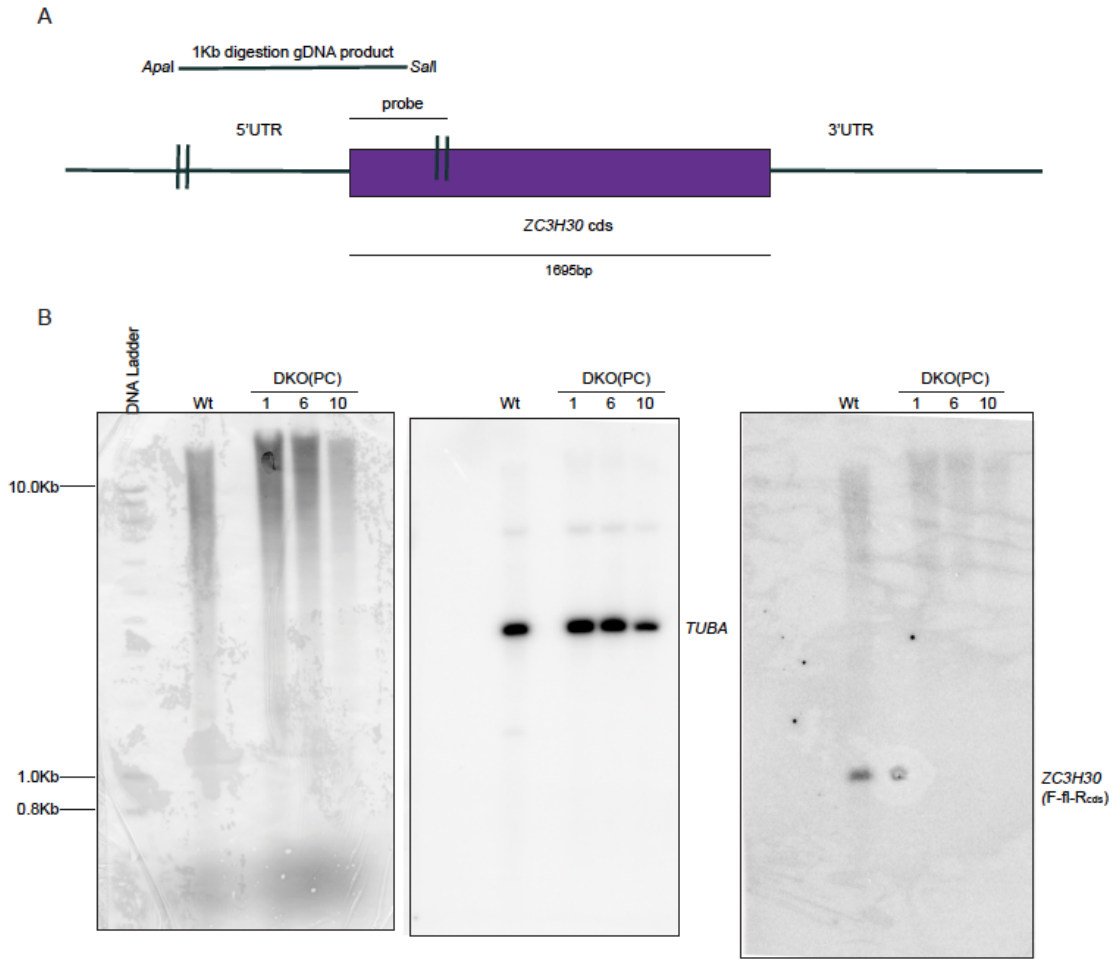


Figure 3.2.6: Southern blot to confirm ZC3H30 double knockout in procyclic trypanosomes. (A). Schematic diagram showing the design of DNA digestion to demonstrate endogenous elimination of *ZC3H30*. The double bars on the locus (outside the 5' UTR) and on the open reading frame (ORF) of ZC3H30 represent restriction digestion sites. The black bar on top represents the length of the digested product (1Kb). The thin black bar with the name probe (300bp) represents the location on the ORF, from where it was PCR amplified, and its binding site on the *Apa*I and *Sal*I digested DNA. The double black bars cutting the genomic DNA upstream of the ZC3H30 5'UTR is the *Apa*I restriction site and the other double black bars cutting right into the ORF of ZC3H30 is the digestion site of *Sal*I. The thin black bar at the bottom represents the length of ZC3H30 ORF (1695bp). **(B)**. Methylene Blue staining of Southern blot demonstrating quantity of DNA transfers; endogenous *ZC3H30* confirmed by PCR probe; equal transfer of all genomic DNA confirmed by using probe amplified from tubulin locus (*TUBA*). For each experiment 10μg of genomic DNA was used and digested with *Apa* I and *Sal* I

Generation of ZC3H30 double knockouts expressing an ectopic ZC3H30 copy

To obtain procyclic trypanosomes lacking endogenous *ZC3H30* alleles but expressing an ectopic copy of the gene, I transfected clone 6 (see Figure 3.2.6 B) with a λN-ZC3H30-2xmyc construct. I wanted the *ZC3H30* DKOs for differentiation and stress experiments. I wanted to ensure that any effect on growth or a specific gene expression I see is actually due to *ZC3H30* gene deletion. Therefore, if I put back the gene and if the phenotype reversed back to the same as the wild type cells, I could confidently conclude that the observed phenotype is due to the involvement of *ZC3H30* gene and not because of some random mutation, which took place in the trypanosomes during transfection or selection. The clones obtained were cultured, Genomic DNA was isolated and the generation of the double knockout with an ectopic copy of the gene was confirmed by PCR analysis (Figure 3.2.7 A). When the PCR reaction was carried out using primers annealing only to the coding DNA sequence (cds) of the gene, a band between 1.5 and 2Kb was obtained, for all reactions, whereas with a PCR reaction with primers annealing outside the 5' or 3' UTR of ZC3H30 and inside the cds of the gene, only the wild type cells yielded the PCR band of expected size. I also checked whether the expression of the ectopic copy was tetracycline inducible or not, by Western blotting. Expression of ZC3H30-myc in the bloodstream form was inducible (Figure 3.2.7 B), whereas the expression in the procyclics (Figure 3.2.7 C) was extremely leaky.

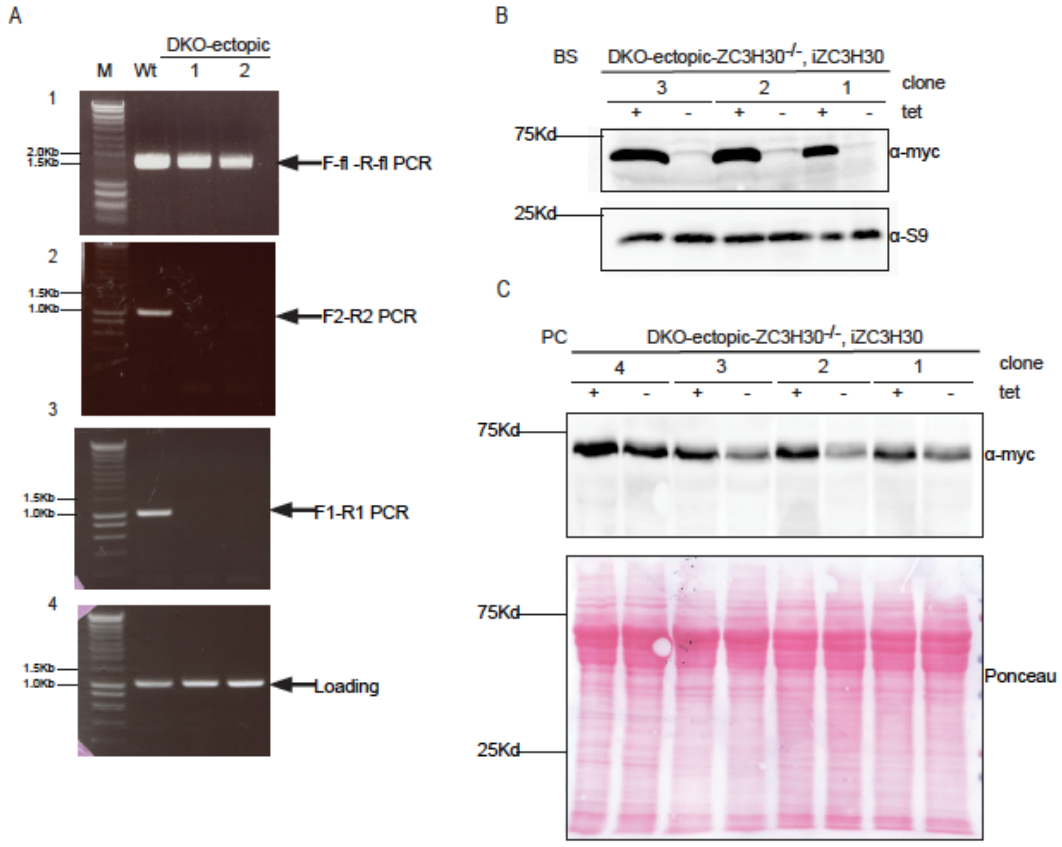


Figure 3.2.7: Confirmation of ZC3H30 double knockout cell lines expressing ectopic ZC3H30 by PCR and Western blotting. (A). PCR analyses demonstrating removal of endogenous ZC3H30 and integration of ZC3H30 in a different location. PCR reactions to test for ORF (1), endogenous ZC3H30 gene (2) and (3) were done as in figure 3.2.5 C. To test that the quality of genomic DNA used for each experiment were intact and nearly same

amount of template DNA was used, a PCR reaction using primers annealing to a different locus was done (4). **(B)**. and **(C).** Western blots to demonstrate tetracycline inducible expression of ectopic ZC3H30 in BS (D) and PC (C) cells. 200ng/ml of tetracycline (tet) was used to induce λN-ZC3H30-2xmyc expression, and the expression was confirmed using anti-myc (α-myc) antibody.

The absence of ZC3H30 does not affect growth of bloodstream form and procyclic form trypanosomes

To find out whether DKOs of ZC3H30 demonstrated any growth and proliferation defect, bloodstream form (BS) and procyclic form (PC) parasites were grown at 37°C and 27°C (normal growth temperatures of BS and PC parasites respectively). As a control, wild type cells (Wt), and double knockout parasites expressing an ectopic copy of λN-ZC3H30-2xmyc (DKO-ectopic) were used. The expression of the ectopic copy was induced by 200ng/ml of tetracycline, and the growth and proliferation of the trypanosomes were monitored for 4 days in BSs and 3 days for PCs. Figure 3.2.8 A and B shows the cumulative growth of BS and PC trypanosomes. There was no significant difference in cell proliferation between Wt, DKO and DKO-ectopic(s) in BS and PC.

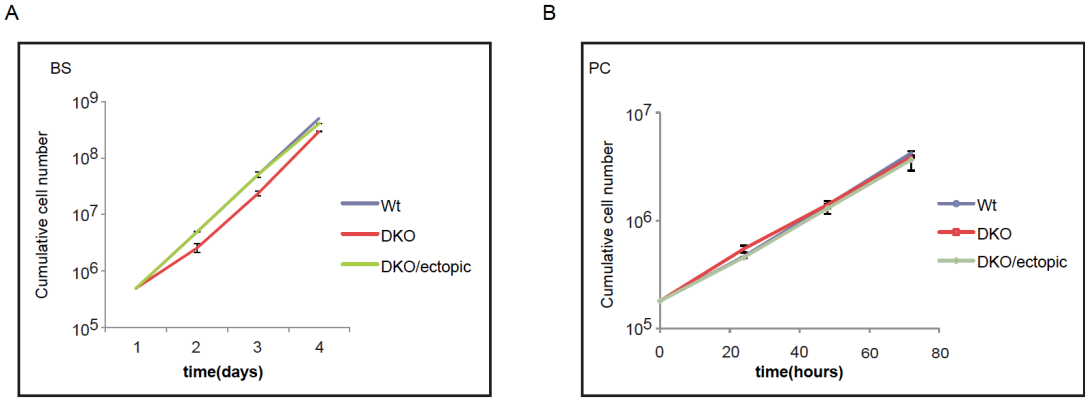


Figure 3.2.8: (A). and (B). Growth curve of cells (BS or PC) in presence of endogenously expressed ZC3H30 (Wt), or lacking the endogenous expression of ZC3H30 with (DKO/ectopic) or without (DKO) carrying an inducibly expressed ectopic copy of *ZC3H30*.

ZC3H30 bloodstream form double knockouts can produce EP procyclin on induction with cis-aconitate

At this point I was interested in knowing which cellular process ZC3H30 might be involved in. I asked myself if ZC3H30 plays a role in differentiation and development of trypanosomes. To test the hypothesis, I checked if ZC3H30 double knockouts were capable of producing EP procyclin. BS ZC3H30 DKOs and Wt cells were treated with 6μM cis-aconitate. After treatment with 6μM cis-aconitate for 17 hours, samples were collected for Western blot and 10^6 cells were pelleted and resuspended in MEM-Pros medium and transferred to 27°C. The growth and survival of the cis-aconitate treated cells were monitored for 72hours. I saw that both Wt and ZC3H30 DKOs could produce EP procyclin on treatment with cis-aconitate (Figure

3.2.9A). The wild type cells and the knockouts died at the same rate in procyclic MEM-Pros medium at 27°C (Figure 3.2.9B). The death of BS 427 cells is inevitable, since they are monomorphic cells and cannot be committed to differentiate to another cell type. So far, there was no evidence that ZC3H30 participates in differentiation and development in BS trypanosomes. This was a preliminary data on differentiation. BS monomorphic cell lines are incapable of differentiation and they perish at 27°C. Therefore, to ensure whether ZC3H30 has a role in differentiation, it would be important to generate double knockout cell lines in pleomorphic trypanosome BS cells (EATRO 1125). Then, it should be induced for differentiation to procyclic form cells by cis-aconitate (CCA). Following induction, growth and differentiation of Wt, DKO cells, DKO-ectopic cells and expression of differentiation specific proteins should be monitored.

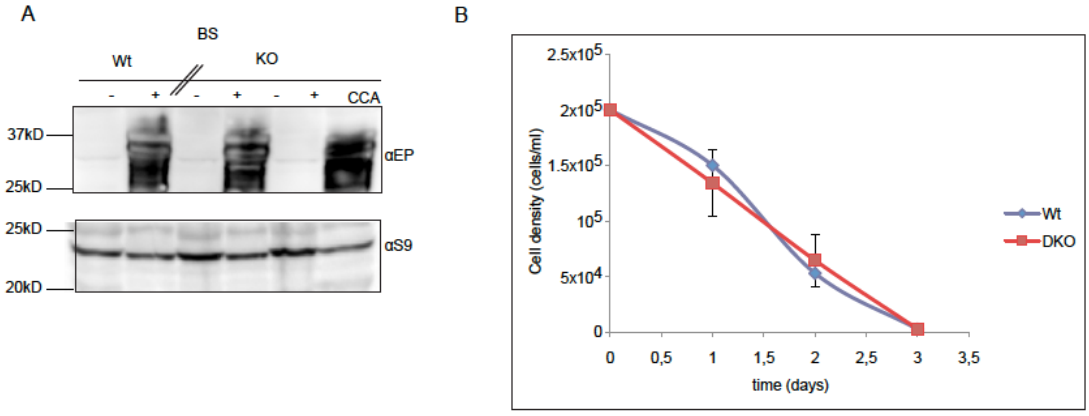


Figure 3.2.9: ZC3H30 double knockouts are capable of making EP procyclin protein upon induction with cis-aconitate. 2x10^6 log phase Wt, KO were treated with 6μM cis-aconitate (final concentration in HMI-9 medium). Post- treatment, the cells were centrifuged and was resuspended in 10ml of MEM-Pros medium. The resuspended culture was shifted to 27°C for 3 days, and the growth of the cultures were monitored. **(A).** Western blot showing that Wt and KOs express EP after 17 hours post-treatment with cis-aconitate. Anitibody against endogenous EP –Procyclin (α-EP) was used for detection and antibody against ribosomal protein S9 was used as loading control. **(B).** Growth/death of Wt and KOs post cis-aconitate treatment, and transfer to 27°C.

Trypanosomes lacking ZC3H30 expression are sensitive to stress

According to the recently published data [112] ZC3H30 is enriched in stress granules upon starvation with 1x PBS for 2 hours. I therefore decided to see if ZC3H30 participated in recovery from stress. First I decided to investigate whether ZC3H30 has a role in heat shock in bloodstream form trypanosomes. For each stress recovery experiment, both BS and PC experiments, the Wt, DKOs and DKO-ectopics were first grown for 2 to 4 days without drugs. The ectopic copy of *ZC3H30* in DKO-ectopic were induced with 200ng/ml of tetracycline (tet) for 2 days before using the cultures for stress. Also, immediately after stress the expression of the ectopic copy was induced (200ng/ml of tet).

I took 10^6 BS form cells and heated the culture in a water bath for 1 hour at 41°C. Then the cultures were centrifuged; the media was discarded and the cell pellet was resuspended in 10 ml of HMI-9 medium without drugs. Growth of the heat-shocked

cultures was monitored for 48 hours. Wt and DKO-ectopics recovered completely within the next 24 hours (Figure 3.2.10A), but the DKOs took 48 hours to recover and proliferate like the wild type cells. I also heat shocked the BS ZC3H30 double knockouts at a higher temperature (42°C). After the shock the cells were harvested like before and their recovery was monitored. Figure 3.2.10B shows the growth tendency of Wt and double knockouts and the double knockouts expressing ectopic inducible copy of ZC3H30 at 42°C. Cells expressing endogenous ZC3H30 (Wt) or ectopic copy again recovered faster than the knockouts (DKOs).

Next, I wanted to know if the lethal phenotype was exclusive for heat shock treatment. Therefore, I stressed the log phase growing cell cultures of wild type, double knockouts and double knockouts expressing ectopic ZC3H30 with 0.1mM dithiothreitol (DTT), a strong endoplasmic-reticulum-associated-degradation (ERAD) stress agent. After 1 hour of DTT treatment, the cells were harvested (850xg, 8minutes) and resuspended in 10ml of HMI-9 medium and their growth was monitored for 72 hours. The Wt, and DKO-ectopic cell lines recovered faster (10 hours) than the DKOs (Figure 3.2.10C). The DKO cell lines achieved similar growth rates after 24 hours, but their number never reached the same as the Wt or the DKO-ectopics. Possibly this is due to more cell death in ZC3H30 DKOs. In my hands 0.1mM DTT was the highest concentration, which could allow recovery of growth of BS cultures after 1 hour of stress. Any concentration above it was lethal for the cells.

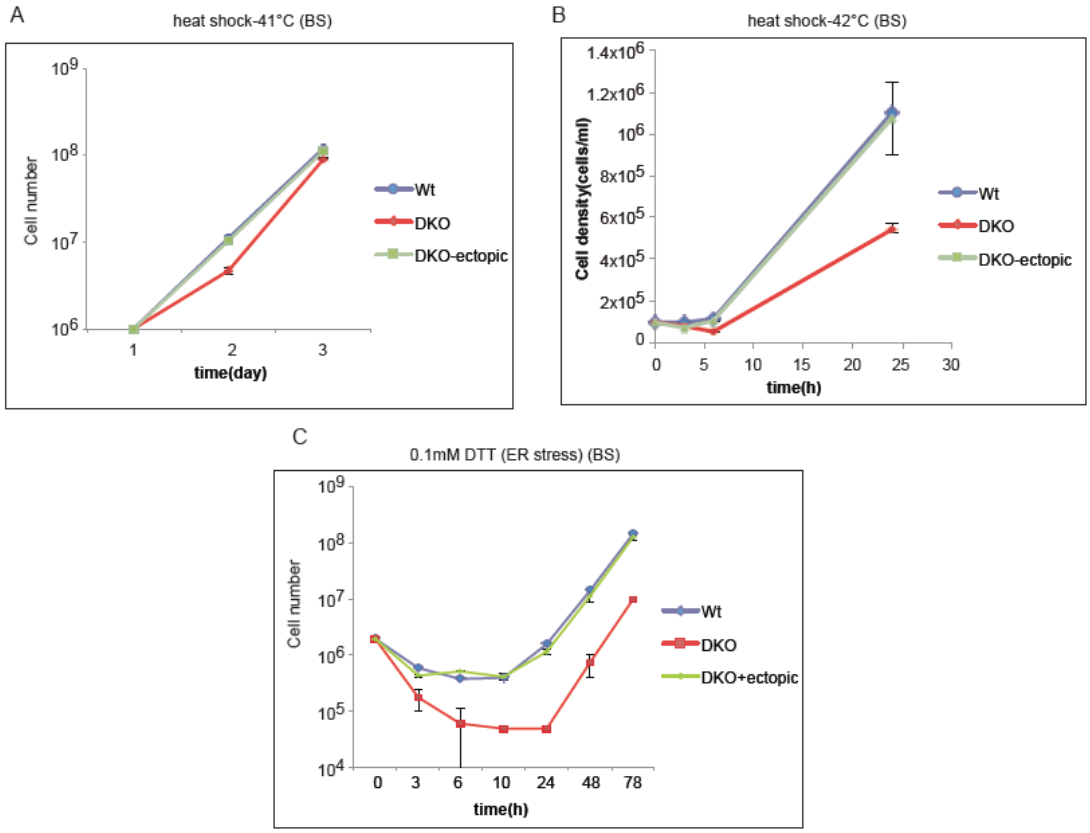


Figure 3.2.10: Bloodstream form *ZC3H30* double knockouts are sensitive to heat shock and ER stress. (A). and **(B).** Growth of BS cells after heat shock. BS (10^6 cells) Wt, DKO and DKO-ectopic were heat-shocked in a water bath at 41°C or 42°C for 1 hour, followed by the recovery and growth was monitored for 48 hours. **(C).** Growth of BS cells after ER stress for 1 hour with 0.1mM DTT. BS ($2x10^6$ cells) trypanosomes Wt, DKO and DKO-ectopic were

treated with 0.1mM DTT for 1 hour; later the cultures were centrifuged and the media with DTT was discarded. The cell pellet was resuspended in 10 ml of HMI-9 media without drugs. Recovery and growth was monitored for 48 hours. Each graph represents two independent experiments.

Next, I wanted to find out whether ZC3H30 plays a role in stress responses in PC trypanosomes. Hence, I shifted logarithmically growing Wt, DKO and DKO-ectopic cells from the optimum growth temperature of 27°C to 37°C for 1 hour. Then the cultures were shifted back to normal 27°C and the recovery and growth was monitored for 48 hours. While the Wt and the DKO-ectopics recovered in less than 20 hours, some DKOs did not survive the heat shock and as a result the culture density did not reach the same number as that of the Wt or the DKO-ectopic (Figure 3.2.11A). I also tested whether procyclic form ZC3H30 double knockouts are as susceptible to heat shock at 41°C as the BS knockouts. I stressed PCs (Wt, DKOs and DKO-ectopic) at 41°C. Later the cells were sedimented, and the pellets were resuspended in MEM-Pros media without drugs. I observed that Wt and DKO-ectopic cells recovered and gained full growth after 48 hours, whereas the ZC3H30 DKOs failed to survive and perished during the recovery period after heat shock (Figure 3.2.11B).

Formation of reactive oxygen (ROS) or nitrogen species (RNS) during oxidative stress results in oxidation of membrane lipids and possibly macromolecules. Sodium arsenite is thought to induce oxidative stress. Ethanol also induces oxidative stress in some cells [135, 136], although it is not known to do so in trypanosomes. I used two different concentrations of Sodium arsenite (10μM and 20μM) to treat Wt, DKO, and DKO-ectopic cultures for 1 hour. Then the cells were transferred to medium without drugs and the recovery and growth of the parasites were monitored over 48 hours. A mild stress of 10μM sodium arsenite was capable of inhibiting the growth and proliferation of the DKOs (Figure 3.2.11D) while 20μM lead to severe growth defect and they failed to recover completely within 48 hours (Figure 3.2.11C). In contrast, the Wt and DKO-ectopics recovered completely after both concentrations of Sodium arsenite. After treatment with 1% and 2% Ethanol for 1 hour, DKOs were susceptible to growth inhibition (Figure 3.2.11E and Figure 3.2.11F), but the wild type and double-knockouts expressing ectopic copy of ZC3H30 were unaffected.

Nutritional limitation or short term starvation [137] in trypanosomes leads to inhibition of cell proliferation and induction of autophagy. I now asked whether the expression of ZC3H30 is essential in short-term starvation stress. Wt, DKO and DKO-ectopic cell cultures were harvested and washed with PBS, then resuspended in 1x PBS for 5 hours (10^6 cells/ml-density in 1x PBS). Afterwards, growth in normal medium was monitored for 48 hours. Unlike the Wt and DKO-ectopics, the DKOs did not recover completely after the short-term starvation (Figure 3.2.11 G).

These results led me to infer that the expression of ZC3H30 might play an important in the survival of cells that have undergone stress.

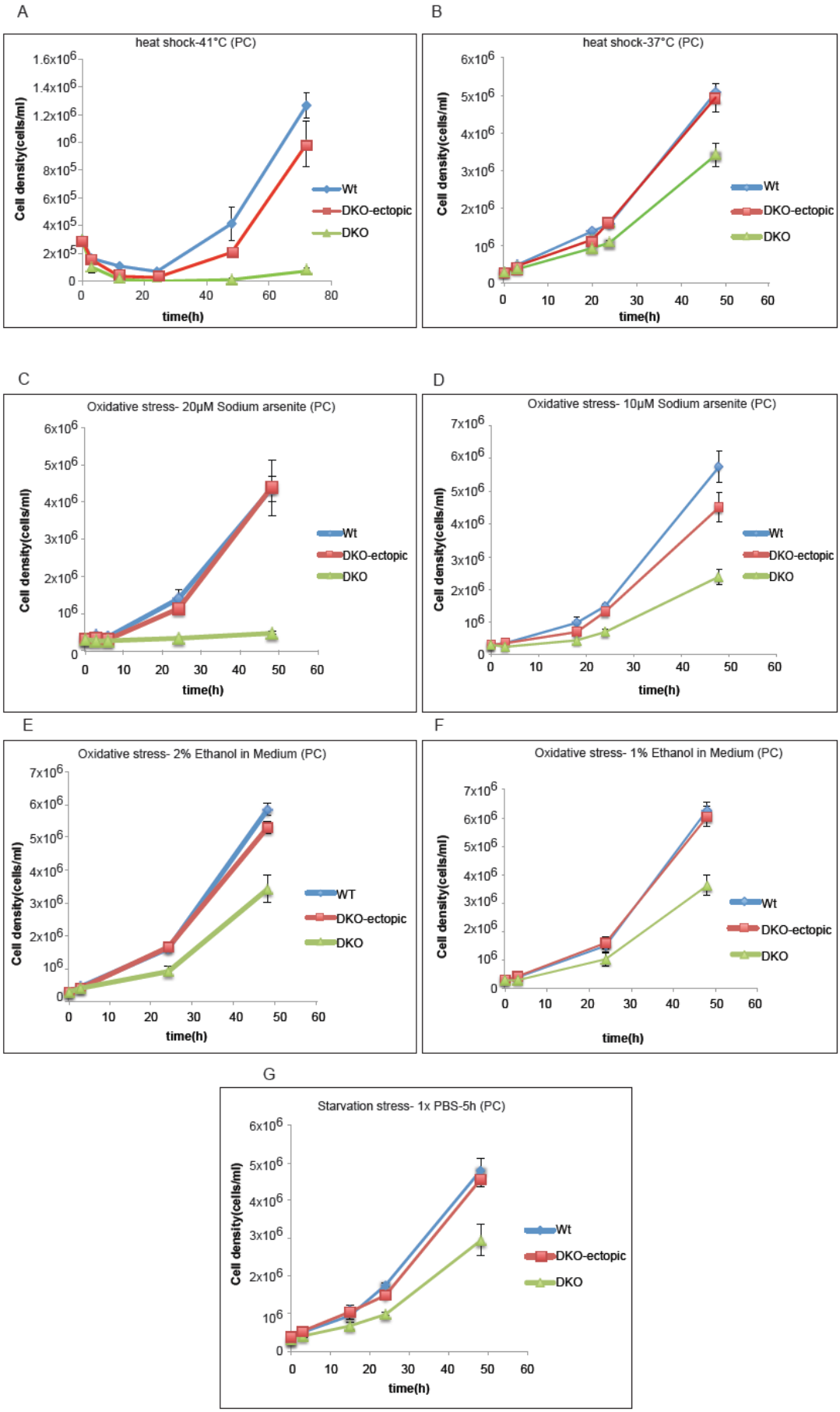


Figure 3.2.11: Procyclic form ZC3H30 double knockouts are also sensitive to stress.

(A). Growth of PC heat shocked cells at 37°C. PC Wt, DKO, and DKO-ectopic were shifted to 37°C incubator for 1hour, and shifted back to normal growth temperature, 27°C; their recovery and growth was monitored for 48 hours. **(B).** Growth tendency of PC heat shocked cells at 41°C. Log phase growing cultures of Wt, DKO, DKO-ectopic were heat shocked in a water bath at 41°C for 1 hour. The growth and recovery was monitored for 48 hours. **(C)** and **(D).** Growth tendency of PC cultures stressed with oxidative stress reagent Sodium arsenite. Wt, DKO and DKO-ectopic were treated with 10μM or 20μM of Sodium arsenite (final concentration in MEM-Pros media), for 1 hour. Then the cells were recovered like B. **(E)** and **(F).** Growth tendency of PC cultures stressed with Ethanol. To the log-phase cell cultures of Wt, DKO and DKO-ectopic, 1% and 2% Ethanol was added for 1 hour. Later the cells were recovered like in B and the growth of recovered cells was monitored for 48 hours. **(G).** The log-phase procyclic cultures (Wt, DKO and DKO-ectopic) were centrifuged and resuspended in 3ml of 1x PBS (final density- 10^6cells/ml 1xPBS) for 5 hours. After starvation, the cells were recovered again as in B. Each graph represents two independent experiments.

Procyclic ZC3H30 double knockouts are more resistant to translational inhibitor Hygromycin than the wild type cells

I investigated the effect of lack of ZC3H30 expression in procyclic form cells upon induction of translational stress. I saw that Wt cells died within 72 hours post-treatment with translation inhibitor Hygromycin (final concentration-50 μg/ml), whereas the DKOs appeared to be more resistant (Figure 3.2.12 A). In order to find out the trypanocidal dose of Hygromycin for the ZC3H30 double knockouts I did a cytotoxicity assay. The IC_{50} of Hygromycin for PC trypanosomes lacking ZC3H30 expression is about 10μg/ml (Figure 3.2.12 B), whereas the Wt cells have an IC_{50} of around 1μg/ml. (I could not use the double knockouts with the ectopic copy because the selection marker for the construct is the hygromycin resistance gene.)

Next I wanted to know, what happens to global translation in presence of hygromycin to Wt, DKO, and DKO-ectopic PCs. Therefore, I treated them with different concentrations of Hygromycin for 18 hours, followed by a S^{35}-Methionine pulse assay. I saw that the knockouts translated more than the wild type growing under normal growth conditions. With increase in Hygromycin concentration, translation decreased in Wt as well as DKOs but at 20μg/ml of Hygromycin the rate of radiolabelled Methionine incorporation was higher in the DKOs than the Wt (Figure 3.2.12 C). At very high concentrations of Hygromycin (100 and 200μg/ml), there was no translation.

This data indicated that ZC3H30 might be involved in negatively regulating translation of PC, and in absence of its expression, there is damage to the balance in basal translation and hence the cells are sensitive to stress. Another possible explanation is that the clones for ZC3H30 DKOs are mutants, which in general translate more and hence can withstand the stress inflicted by Hygromycin.

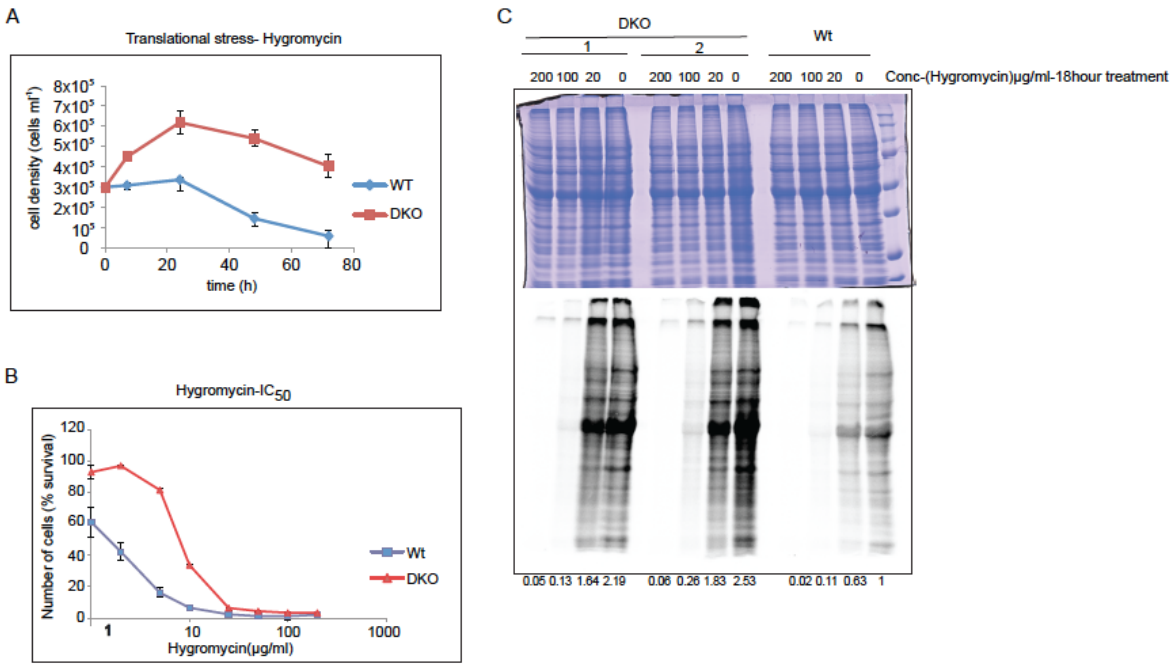


Figure 3.2.12: PC ZC3H30 DKOs translate more in presence of Hygromycin. (A). Growth curve showing that Procyclic form ZC3H30 double knockout cells are more resistant to hygromycin. Log phase procyclic cell cultures of Wt, DKO and DKO-ectopic were grown in MEM-Pros medium containing 50µg/ml of Hygromycin. **(B).** Hygromycin IC_{50} of Wt and DKO. Cell cultures were treated with different concentrations of Hygromycin form 0 to 200µg/ml. The trypanocidal assay and IC_{50} calculation was carried out as described in 'Cytotoxicity assay'. **(C).** SDS-PAGE gel showing effect of S^{35}-Methionine labelling on treating Wt and DKOs with different concentration of hygromycin. Cells growing exponentially were treated with 0, 20, 100 and 200µg/ml of translation inhibitor Hygromycin for 18 hours, followed by metabolic labelling with S^{35}-Methionine, washing and resuspending in Laemli buffer. The numbers below the gel implies a ratio of radioactive signal to the intensity of bands obtained after Coomassie staining.

Translation of Knockouts and knockout expressing ectopic copy upon heat shock

It was strange to infer that ZC3H30 regulated basal level translation of total mRNAs. It was irregular that a protein not essential for regular growth of the parasite regulated an essential task like regulation of homeostatic translation. Therefore I decided to look at translation of wild type, ZC3H30 double knockout and ZC3H30 double knockout expressing ectopic copy of the gene at heat shock. Figure 3.2.13 shows the results of S^{35}-Methionine incorporation in absence and presence of heat shock. $5x10^6$ million cells were either heat shocked or not at 37°C. I observed is that the clones selected as knockouts, were translating at a higher rate than the wild types. I hypothesized that this unusual phenotype is not due to the absence of ZC3H30, since ectopically expressing ZC3H30 clones were also translating significantly much more than the wild types. The intriguing thing was that right after heat shock the expression of chaperones HSP60 and HSP70 were up in wild type and knockout cells expressing ectopic copy whereas ZC3H30 double knockouts failed to elevate the chaperone protein production compared to the control protein

production (Figure 3.2.13 B). Their signals were compared and normalised to other control proteins and then further normalised to the Commassie stained bands of the same.

This is a preliminary data and a lower temperature for heat shock has been used (37°C). To ensure that the ZC3H30 DKOs are truly incapable of shutting global translation during stresses like heat shock, this experiment needs to be repeated at 39°C and 41°C, since these two are most used temperatures for heat shock studies [105, 106, 138]. Also what needs to be further checked is if this phenotype is exclusive to a particular life stage of the parasite. Hence this experiment of metabolic radiolabelling of proteins during heat shock needs to be done in BS cells as well. Figure 3.2.13 C, shows that the ectopic ZC3H30 was well expressed.

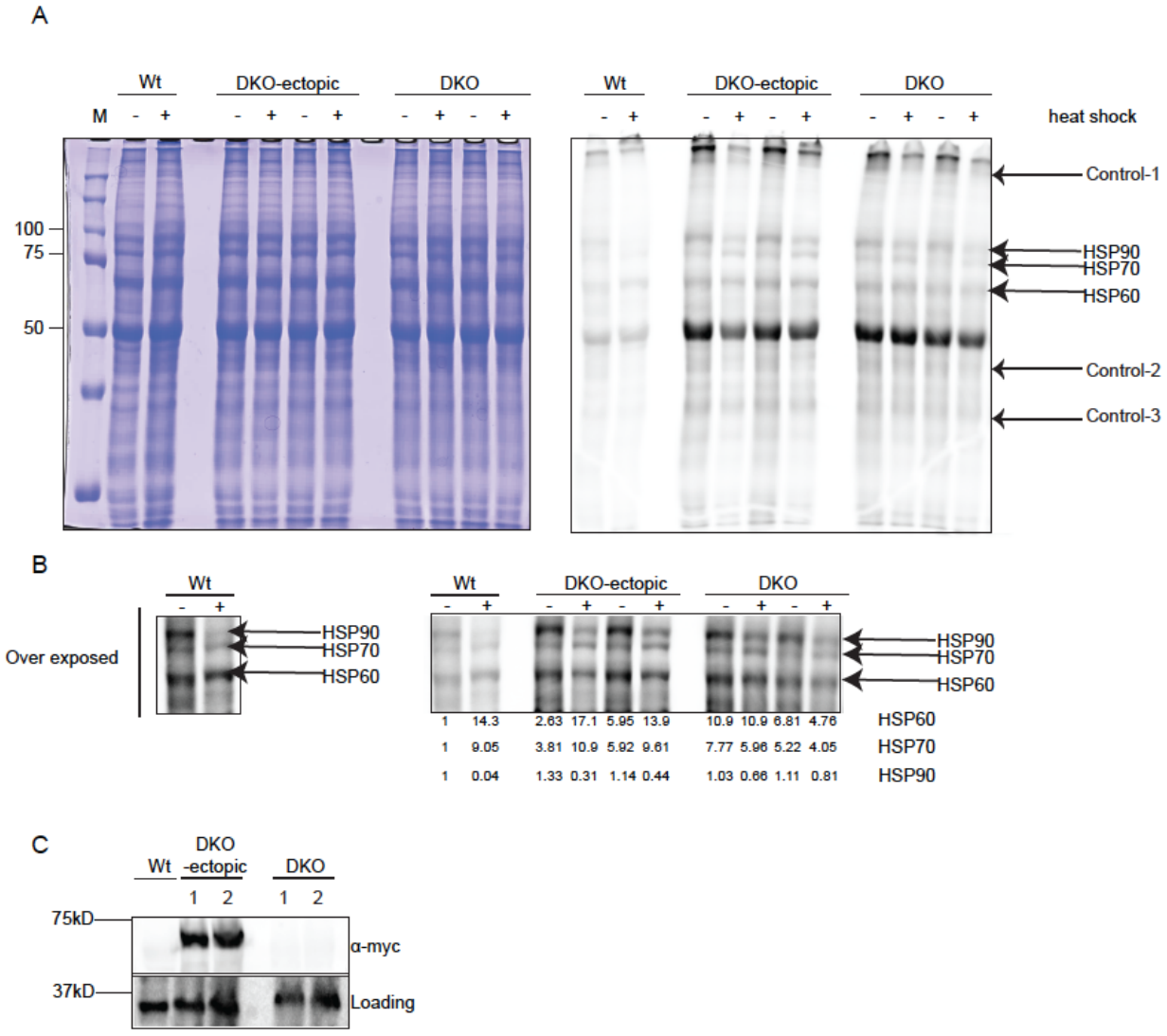


Figure 3.2.13. Translation during heat shock. Procyclic cultures of Wt, DKO and DKO-ectopic were subjected to heat shock. **(A)**. Metabolic labeling assay. Incorporation of S^{35}-Methionine was checked in presence and absence of ZC3H30 with or without heat shock at 37°C. The cells were washed to remove unincorporated radioactive methionine and each sample was loaded on a 10% SDS-PAGE, and was stained with Coomassie R250. On the left is the picture of the gel after washing excess Coomassie stain, and on the right is the autoradiograph, depicting signal from each sample in presence and absence of heat shock.

(B). Over-exposed signals from the autoradiograph to show the signals upon heat shock from heat shock proteins (HSP). The numbers below the right picture are the normalized signal from HSP60, -70 and -90. Quantification of signals from other parts of the autoradiograph was obtained, and then it was further normalised to the Coomassie stained bands. **(C).** Western blot depicting expression of ZC3H30. Expression of ZC3H30 was checked by probing the membrane with anti-myc (α-myc) antibody.

Stress markers in ZC3H30 knockout cell lines

Next I wanted to know whether the lack of *ZC3H30* expression affects chaperones. Since the DKOs were sensitive to stress, I wondered whether the chaperone expression was disturbed in absence of ZC3H30. BS Wt, DKO and DKO-ectopics were heat shocked at 41°C, and the expression of chaperones were analysed by Western and Northern blotting. There was no difference in the induction of HSP60 protein (Figure 3.2.14 B), however, at the normal temperature, the ZC3H30 DKOs had higher (~x2) levels of the major cytoplasmic *HSP70* (Tb927.11.11330) mRNA than the Wt and the DKO-ectopics (Figure3.2.14 A). The same was observed in PC ZC3H30 DKOs (Figure3.2.14 F). Upon heat shock at 41°C (BS and PC) and 42°C (BS) the *HSP70* mRNA level in the DKOs did not increase, unlike the Wt and DKO-ectopics. But when I checked the levels of HSP70 protein in PC, there was no difference in the Wt, DKO and DKO-ectopics at 27°C or 41°C. Also on probing the membrane with the antibody for HSP70, a lot of unspecific bands were noticed. Possibly the effect on HSP70 (Tb927.11.11330) gene expression is only on the transcript and does not affect translation, or I do not see any difference because of unspecific detection due to the antibody. This result is uncertain and was not followed up because when we later did RNA Seq analysis in PCs, the difference was not seen.

ZC3H11 binds the HSP70 mRNA 3' UTR and recruits stabilizing factors [105, 107]. Expression of ZC3H11 is elevated upon heat shock or oxidative stress [104]. At 27°C (PC) and 37°C (BS), there was no difference in the protein levels of ZC3H11 between the Wt, DKO and DKO-ectopics. Figure 3.2.14 C and D are Western blots showing induction of ZC3H11 in BS and PC. This implies that lack of ZC3H30 expression does not influence the expression of ZC3H11.

I also extracted total protein from 0.1mM DTT treated cells. The expression of ERAD marker BiP and stress induced RBP, SCD6 were checked. Figure 3.2.14 shows a Western blot depicting that 0.1mM DTT treatment does not affect either marker.

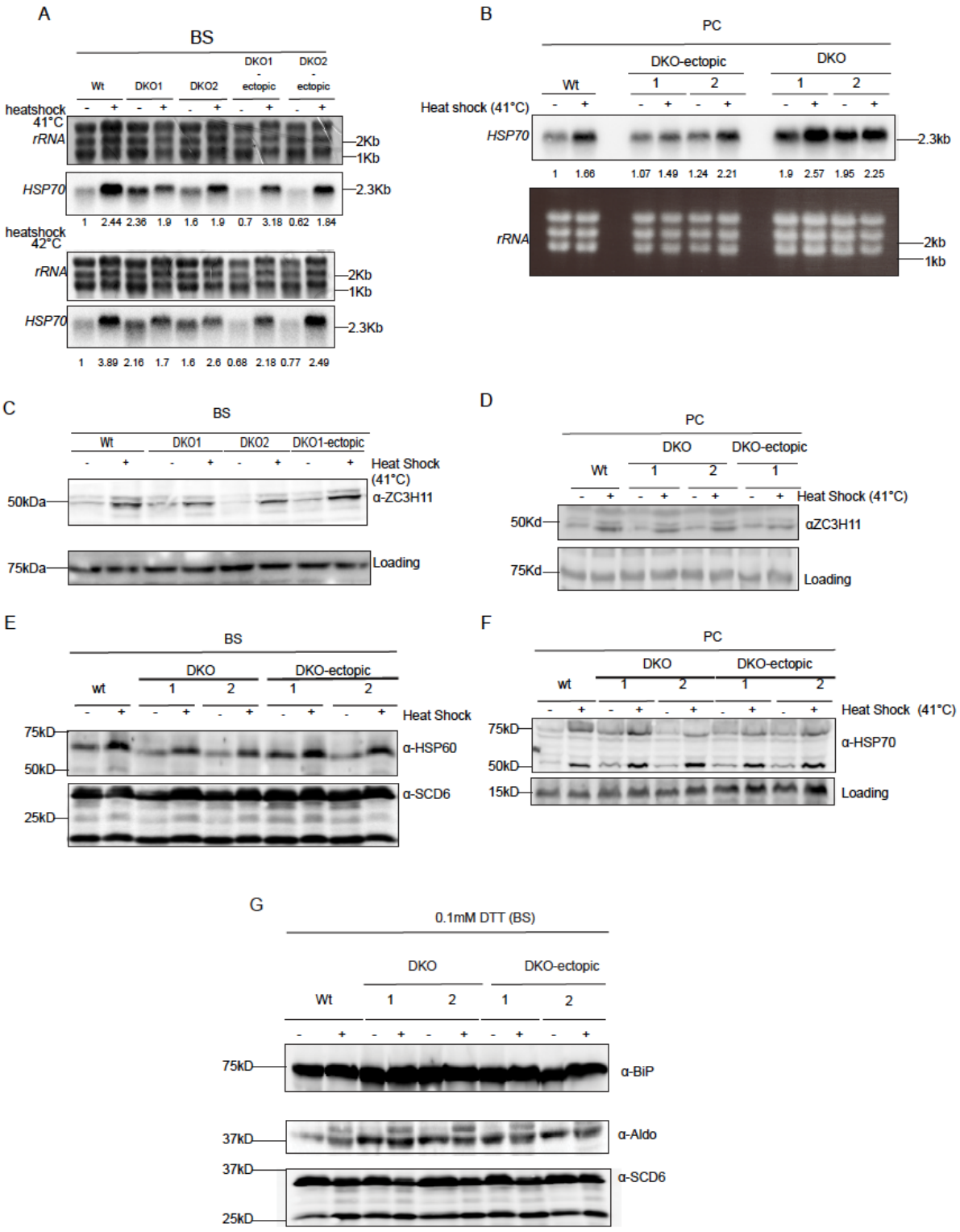


Figure 3.2.14: Chaperone mRNA/proteins are not affected due to lack of expression of ZC3H30. **(A)**. and **(B).** Amount of *HSP70* mRNA (major cytoplasmic) in DKOs. BS (A) Wt, DKO and DKO-ectopic cells were heat shocked at 41°C or 42°C. Total RNA from non-heat shocked and heat shocked cells were extracted and the *HSP70* mRNA level was analysed on a Northern blot. Also PC (B) Wt, DKO and DKO-ectopic cells were heat shocked at 41°C. Total RNA from non-heat shocked and heat shocked cells were extracted and the *HSP70* mRNA level was analysed on a Northern blot. **(C).** and **(D).** ZC3H11 levels in ZC3H30 BS and PC DKOs. Cytoskeleton free total protein was extracted from BS (C) and PC (D) from non-

heat shocked and heat shocked (41°C) cells (Wt, DKO and DKO-ectopic) and processed for Western blotting. ZC3H11 was identified using anti-ZC3H11 antibody (α-ZC3H11). Induction of ZC3H11 upon heat shock was confirmed by increase in intensity of protein band for ZC3H11. **(E).** Induction of HSP60 upon heat shock in BS ZC3H30 DKOs. Log phase cultures were harvested and loaded on a poly acrylamide gel to check for HSP60 induction upon heat shock at 41°C. HSP60 protein was identified using anti-HSP60 antibody (α-HSP60). A stress induced RBP SCD6 was probed (α-SCD6) for loading control. **(F).** Induction of HSP70 upon heat shock in PC ZC3H30 DKOs. Log phase cultures were harvested and loaded on a poly acrylamide gel to check for HSP70 induction upon heat shock at 41°C. HSP70 protein was identified using anti-HSP70 antibody (α-HSP70). A non-specific band was used as control. **(G).** BiP (ER HSP70) and SCD6 (stress induced RBP) induction in BS ZC3H30 DKOs upon ER stress with 0.1mM DTT. Log phase cultures were harvested and loaded on a poly acrylamide gel to check for BiP and SCD6 induction upon heat shock at 0.1mM DTT. BiP and SCD6 proteins were identified using anti-BiP and anti-SCD6 antibody respectively (α-BiP, α-SCD6). Metabolic enzyme Aldolase was probed with anti-Aldolase antibody (α-Aldo) as loading control.

Expression of V5-ZC3H30 protein is increased upon stress

ZC3H11 is upregulated and phosphorylated upon heat shock [104, 105]. To find out whether stressing trypanosomes had any impact on the protein abundance of ZC3H30, I tagged ZC3H30 N-terminally with a V5-tag in frame with the open reading frame. Since, the 3' untranslated region (UTR) of the gene is conserved by this procedure, I expected that the protein expression should not be altered, unless tagging affects protein stability. PC cells were heat shocked at 37°C and 41°C, or stressed with 10μM Sodium arsenite for 1 hour. The cells expressing V5-ZC3H30 were also treated with drugs inhibiting translation - Hygromycin and Puromycin. (The cell line expressing V5-ZC3H30 did not carry any genes that can provide resistance against these two drugs.) I had previously established that PC ZC3H30 DKOs could withstand translational stress by Hygromycin better than Wt cells. Therefore, I wanted to test if Hygromycin had any impact on ZC3H30 protein levels. Also, I wanted to test if other translational inhibitors like Puromycin had any effect on the protein level of V5-ZC3H30. For each inhibitory drug two different concentrations were selected. The higher concentration is the one used routinely in cell culture for selection of constructs carrying the resistance cassettes (≈ x50 IC_{50}). Both the concentrations are capable of clearing non-resistant procyclic form cells from media within 3 days. Translational stress was applied by adding the drugs to log phase cultures for 24 hours. Cells from stressed cultures and cultures where we added translation-inhibiting drugs were harvested and analysed on a Western blot. Upon heat shock or sodium arsenite treatment, the abundance of V5-ZC3H30 increased (Figure 3.2.15A). Interestingly, the addition of 25μg/ml and 50μg/ml of Hygromycin B could also elevate V5-ZC3H30 protein level, but 0.5μg/ml and 1μg/ml of Puromycin could not do the same. The higher protein abundance in ZC3H30 indicates that its expression might be important in stress. Also, the clear higher abundance in hygromycin stressed cells against the puromycin stressed cells, indicate, that the protein might participate in a translation regulation pathway inhibited by Hygromycin. Figure 3.2.15B shows the induction of ZC3H11 upon stress as a positive control; it was increased only after the 41°C treatment.

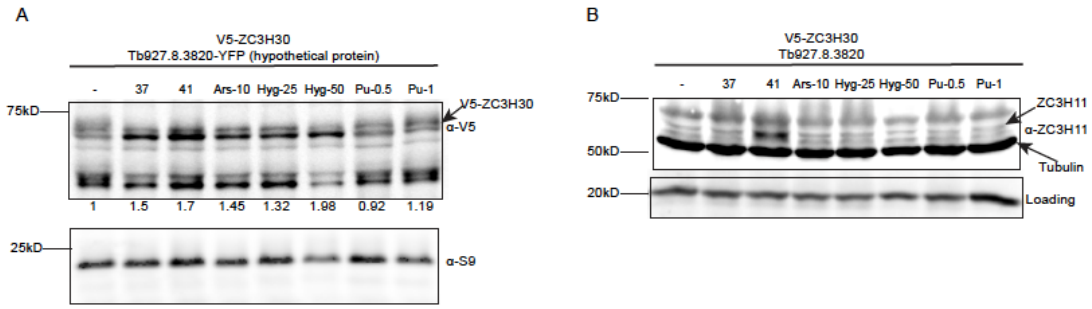


Figure 3.2.15: ZC3H30 protein level is up upon stress. (A). Western blot to show response of V5-ZC3H30 and YFP-Tb927.8.3820 (hypothetical protein) upon stress. Procyclic cell line expressing in situ N-terminal V5-tagged ZC3H30 and C-terminal YFP tagged hypothetical protein (Tb927.8.3820) were grown and maintained at log phase. The cells were either harvested under normal conditions (27°C), or after 1hr of stress infliction in the form of heat shock (37°C-37, 41°C-41), or oxidative stress with 10μM Sodium arsenite (Ars-10), or translational stress with 25 and 50μg/ml Hygromycin (Hyg-25, Hyg-50) and 0.5 and 1μg/ml Puromycin (Pu-0.5, Pu-1). The cells were washed and resuspended in Laemli buffer and then analyzed by Western blotting. The band lower to 75kD when probed with anti-V5 antibody (α-V5) is the ZC3H30 intact ZC3H30 band, and the band above to 75kD when the blot was probed with anti-GFP antibody (α-GFP) is the band for the hypothetical protein tagged C-terminally with a GFP tag. **(B).** Western blot to show up regulation of ZC3H11 upon stress. ZC3H11 was identified by probing with anti-ZC3H11 antibody (α-ZC3H11). A non-specific band at 20 kD was used as loading control.

ZC3H30 localizes in heat shock granules after heat shock

In a recently published article [112], it was demonstrated that ZC3H30 protein was localised in the granule fraction when PC form trypanosomes were starved in 1x PBS for 2 hours. I wanted to find out whether the migration of ZC3H30 to the granules is a general stress phenotype. Hence I decided to heat shock the cells carrying an *in situ* N-terminal V5-tagged ZC3H30 and separate the soluble and the granule fraction and find the fraction where V5-ZC3H30 migrates upon stress. Therefore, I harvested log phase PC culture with or without heat shocked at 41°C for 1 hour, and isolated stress granules as previously described [106]. Samples were analysed by Western blotting. Upon heat shock at 41°C for 1hour, V5-ZC3H30 moved to the granule (G) fraction (Figure 3.2.16) along with SCD6 and DHH1 (markers known to localise in stress granules). As expected, cytoplasmic markers trypanothione reductase (TR) and ribosomal protein S9 did not move to the granules upon heat shock. From this experiment I concluded that ZC3H30 is in granules after both heat shock and short-term starvation. The Western blot below is one representative result of three independent experiments to check migration of V5-ZC3H30 to the heat shock granules upon heat shock at 41°C. I wanted to check by immunofluorescence whether ZC3H30 localises to stress granules upon heat shock, so I used the V5-ZC3H30 PC cells. Since the signal these images (using anti-V5 antibody) was very poor (no difference between the signal and the background), I decided not to follow up.

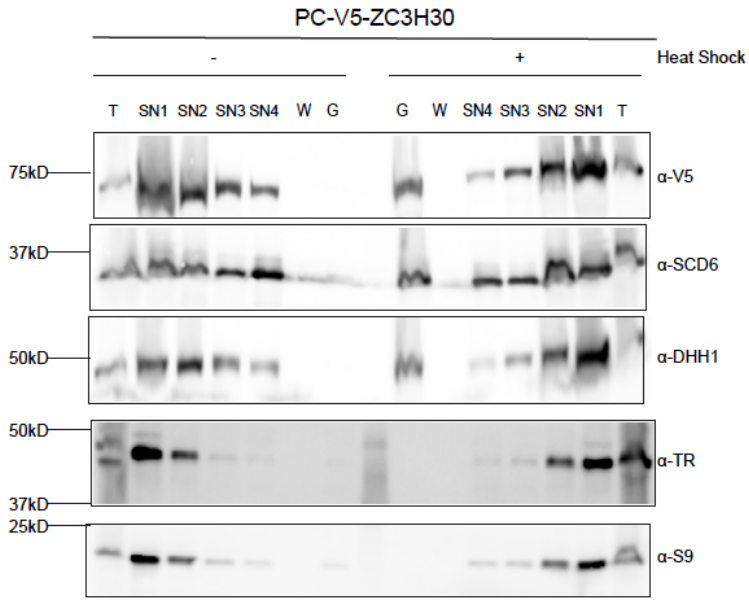


Figure 3.2.16: V5-ZC3H30 localizes to heat shock granules upon heat shock at 41°C. Localization of V5-ZC3H30 upon heat shock was studied by Western blot analysis of protein samples taken during heat shock stress granule isolation as described previously [112]. About equal amount of PC expressing in situ tagged V5-ZC3H30 either heat shocked (41°C) or not (27°C) were harvested for heat shock stress granule enrichment. Antibodies against SCD6 (α-SCD6) and DHH1 (α-DHH1) serve as positive controls for stress granule formation, and antibodies against trypanothione reductase (α-TR) and ribosomal protein S9 (α-S9) serve as negative control, as they are cytoplasmic and do not relocate to granules upon stress. Antibody against the V5 (α-V5) tag was used to detect V5-ZC3H30. The lanes represent the fraction during sequential isolation of the granules; T-total (total fraction), SN1-soluble fraction 1, SN2 and 3- soluble fraction 2 and 3 (isolated after shear force application with 21G needle), SN4- fraction obtained upon solubilisation of cytoskeleton (serial increase of salt concentration in lysis buffer), W- wash (fraction obtained on washing the granules with lysis buffer without detergent and resuspension in buffer), G- granule fraction obtained after wash.

Generation of N-terminal in situ TAP-ZC3H30 cell line and testing the functionality of the protein

To understand the role of ZC3H30 in stress I decided to look for its protein and RNA binding partners. ZC3H30 protein was TAP tagged endogenously at the N-terminus, and the other copy was knocked out, as designed in the scheme of Figure 3.2.17A in PC trypanosomes. The deletion of the other copy and the presence of the endogenous TAP tagged copy were confirmed by PCR (Figure 3.2.17B). In order to ensure the functionality of my TAP-ZC3H30 protein, I stressed the cells by heat shock. The PC SKO expressing TAP-ZC3H30 (TAP) recovered as well as Wt indicating that the protein was functional (Figure 3.2.17C). Hence I proceeded to affinity purification and investigation of bound RNA and proteins with the cell line.

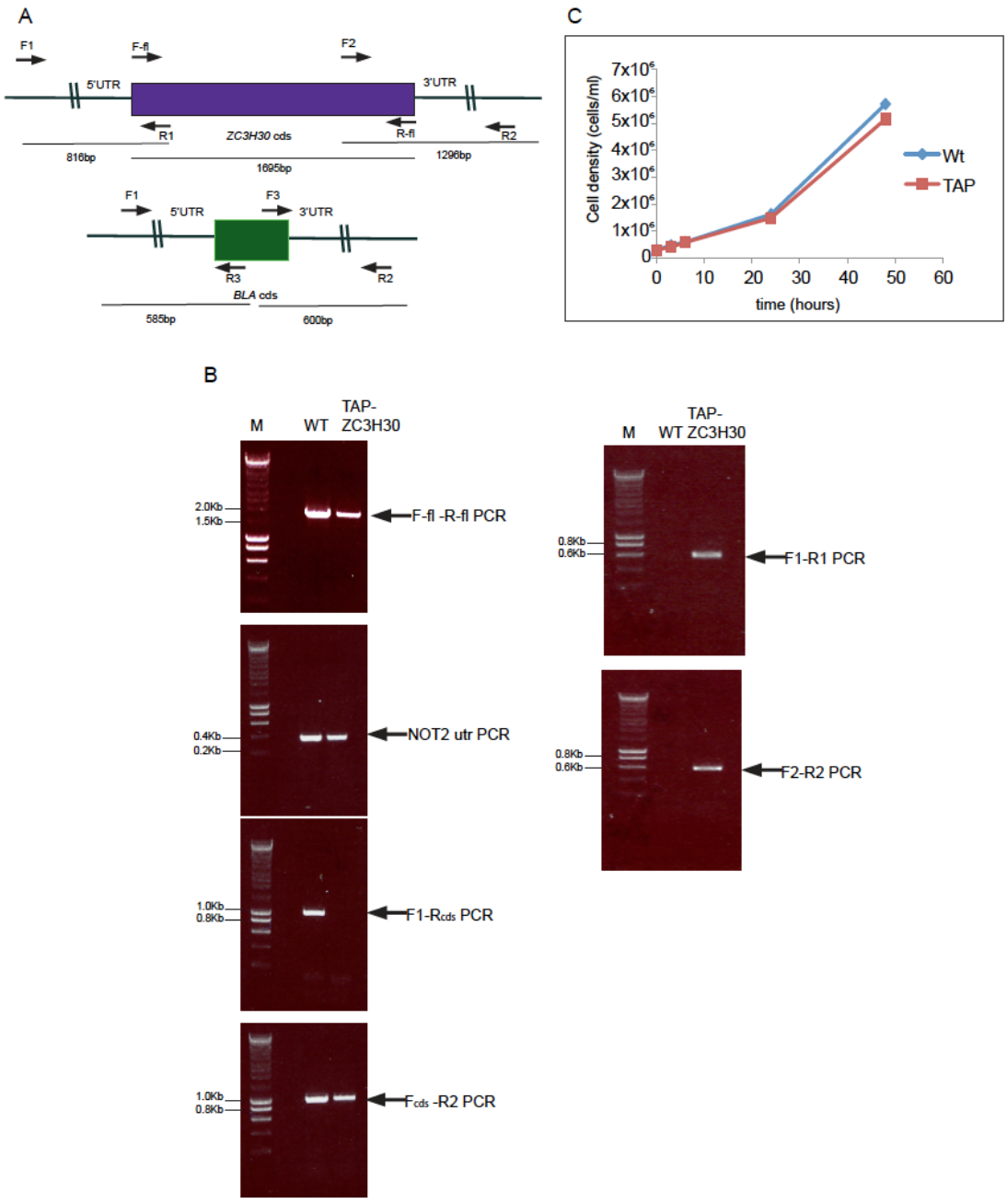


Figure 3.2.17: Generation of TAP-ZC3H30 single-knockout in PC. **(A).** Schematic diagram of *ZC3H30* gene to confirm its presence in the right locus and for the replacement of *ZC3H30* gene with a blasticidin resitance (*BLA*) cassette in the locus. Oligos used for PCR analysis is represented by arrows. The length of PCR products are represented by underlined bars, and their length are given below in base pairs (bp). **(B).** PCR analysis of genomic DNA isolated from wild type (WT) and single-knockout (SKO) N-terminal TAP-tagged cell line (TAP-ZC3H30) to ensure the replacement of one copy of ZC3H30 gene with *BLA* resistance cassette using the genomic DNA of Wt and SKO/TAP-ZC3H30 as template. **(C).** SKO/TAP-ZC3H30 cell line behaves like the Wt upon heat shock. Wt and SKO/TAP-ZC3H30 (TAP) cells were heat shocked at 37°C for 1 hour; recovery of the cells were monitored at 27°C for 48 hours post heat shock.

The protein-binding partners of ZC3H30

To understand the role of ZC3H30 in stress better I did a tandem affinity purification to identify proteins that interact with ZC3H30 both in presence and absence of RNase A. I expected proteins interacting with ZC3H30 will shed light on its function in trypanosomes. Figure 3.2.18A is one representative blot to show the efficiency of the pull down after the first purification. For this analysis, 10^7 million cells were loaded in each well from each fraction on a SDS-PAGE followed by Western blotting. Figure 3.2.18B shows a graphical representation (using Cytoscape [139]) of the proteins identified in the affinity purification with at least 10% coverage and identification of minimum 2 peptides from 2 replicates. A huge number of abundant proteins (cytoskeletal-Actin A, metabolic enzymes- pyruvate phosphate dikinase, glutamate dehydrogenase, ribosomal proteins) were pulled down by TAP-ZC3H30. These proteins are pulled down by nearly identified in almost all purification. I think they interacted with ZC3H30 unspecifically and hence I decided not to validate their interaction. I had observed that ZC3H30 gene deletion led to increase in total translation and my mass-spectrometry data showed that it pulled down a number of ribosomal proteins. Yet, I decided not to follow in validating these interactions because ribosomal proteins are extremely abundant and verification by co-immunoprecipitation would not be conclusive, since it would have pulled down a lot of non-specific proteins. There was only one protein (Tb927.8.3820- hypothetical protein), which was unique to TAP-ZC3H30 purification. This was known to be over-represented in the granule fraction of procyclic form starvation granules [112], along with ZC3H30. The hypothetical protein encoded by Tb927.8.3820 has a molecular weight of protein of 34.81 kDa and its structure is spanned with 5 low complexity regions, as predicted by the TritrypDB. Figure 3.2.18 C represents the primary structure with the identified low complexity regions (LCRs). The LCRs were also identified using SMART online tool to find annotated domains [131].

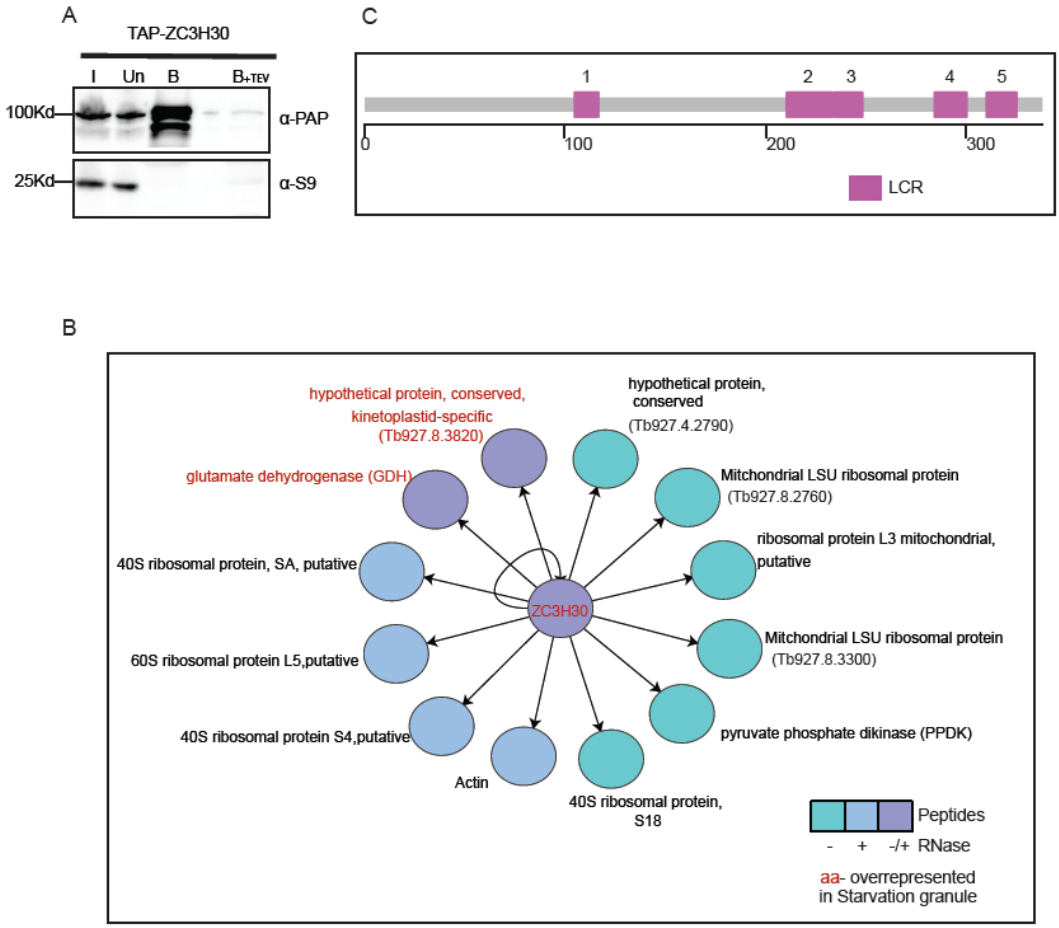


Figure 3.2.18: Affinity purification to identify protein interacting partners of ZC3H30. 10^{10} PC SKO/TAP-ZC3H30 trypanosomes were harvested and affinity purified in two steps as described in 'Materials and Methods- Tandem Affinity purification followed by Mass spectrometry' section. **(A).** A representative western blot to represent the efficiency of the first purification. Anti-PAP antibody (α-PAP probing Protein A of the TAP tag) was used to identify TAP-ZC3H30 on a western blot. The blot was also probed against ribosomal protein S9 (α-S9) as a loading control. The lanes indicate the Input (I), unbound (Un) of the first purification with IgG-sepharose beads, bound (B) fraction to the IgG-Sepharose beads, bound fraction after cleavage of the TAP tag with TEV protease (B_{+TEV}). **(B).** Protein binding partners of TAP-ZC3H30. A schematic diagram showing proteins pulled down by TAP-ZC3H30 on affinity purification. The proteins listed had at least 10% coverage and appeared in at least 2 replicates with a minimum number of 2 peptides. The proteins marked with cyan colored balls were found in the samples where the lysis buffer without RNase A, the ones in light blue from the samples where the initial lysis buffer contained RNase A, and the ones in violet were found in both conditions. **(C).** A schematic representation of the hypothetical protein identified as a specific binding partner of ZC3H30 from TAP-MS analysis. The grey bar represents the primary structure, and the pink boxes represent low complexity regions (LCRs). The black line below represents scale of the primary sequence. The graphical representation was obtained by running the primary sequence in SMART sequence analysis [131].

Interactions of ZC3H30 and the hypothetical protein encoded by Tb927.8.3820

In order to verify the interaction with Tb927.8.3820 I decided to do co-immunoprecipitation. A C-terminally YFP tagged construct for *Tb927.8.3820* was obtained as a generous gift from the laboratory of Dr. S. Kramer, University of Würzburg. I used this construct to replace one endogenous copy of the Tb927.8.3820 gene in procyclics. Next I transfected the cell line with my construct to replace an endogenous copy of *ZC3H30* by a N-terminal V5-ZC3H30. I pulled down V5-ZC3H30 using anti V5 antibody conjugated agarose beads and looked for C-terminally tagged hypothetical protein (8.3820)-YFP (YFP-Tb927.8.3820) using anti-GFP antibody. V5-ZC3H30 pulled down hypothetical protein (8.3820)-YFP; as a negative control I used anti-V5 agarose beads with extracts from a cell line expressing 8.3820-YFP only; then The anti-V5-agarose beads did not pull down the 8.3820-YFP. Similarly when I used agarose beads conjugated with GFP binding peptides (GBP-beads), I saw that V5-ZC3H30 was pulled down with 8.3820-YFP but not from cells without 8.3820-YFP. I did each experiment twice in presence and absence of RNase A (200µg/ml) to find out whether this interaction dependent on mRNA. V5-ZC3H30 interacted 8.3820-YFP in both conditions (Figure 3.2.19A and B), indicating that the protein-protein interaction might be RNA-independent. To test if the interaction is RNA-independent, one must study the interaction by Yeast-two-hybrid assay, using either protein as bait or prey or both. This system excludes chances of interaction with trypanosome mRNAs (possible targets of one or both proteins).

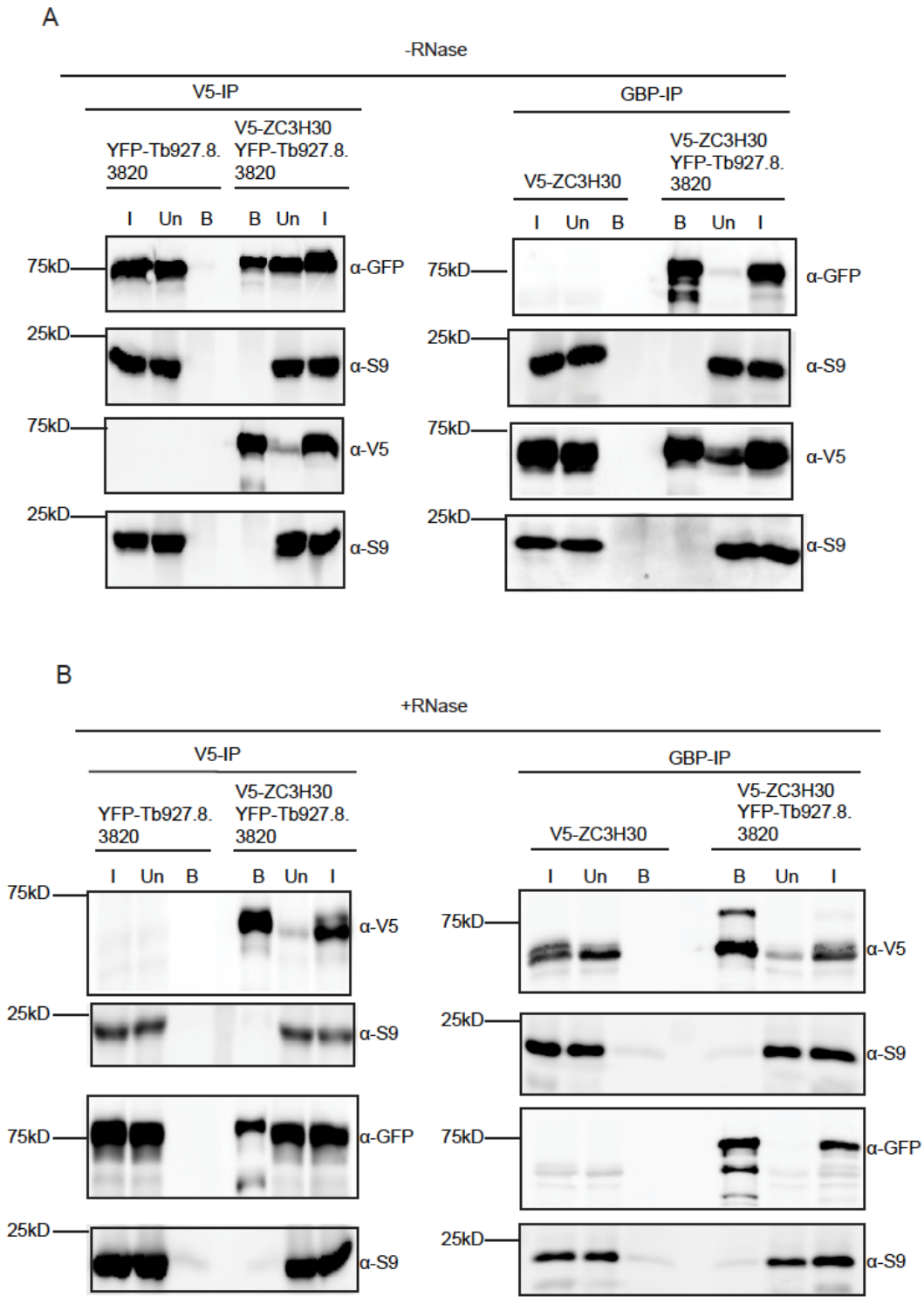


Figure 3.2.19: Co–immunoprecipitation shows that ZC3H30 interacts with the hypothetical protein (Tb927.8.3820). (A). Western blot to show that V5-Z3H30 pulls down hypothetical protein and vice versa. Co-immunoprecipitation was done with cell extracts from *in situ* N-terminal V5-tagged ZC3H30, C-terminal YFP-tagged Hypothetical protein (Tb927.8.3820), or both, using anti-V5 antibody conjugated agarose beads (V5-IP), or GFP binding protein (GBP) conjugated agarose beads (GBP-IP). The lanes indicate, input (I), unbound (Un- lysate after incubation with beads) and bound (B- fraction bound to beads post incubation and rigorous washing). The proteins ZC3H30 and the hypothetical protein were probed using anti-V5 (α-V5) and anti-GFP (α-GFP) antibody; anti-S9 antibody used to probe ribosomal protein S9 (α-S9) was as a loading control. **(B).** ZC3H30 and hypothetical protein

encoded by Tb927.8.3820 interaction RNA-independent. The co-immunoprecipitations of the cell-lines were done as in A, only here the lysis and wash buffer contained 200μg/ml of RNase A.

RNAi against Tb927.8.3820 gene

Next, I wondered whether expression of normal levels of Tb927.8.3820 was required for cell growth and proliferation. I made a cell line expressing in situ V5-tagged ZC3H30 and YFP-tagged hypothetical protein (8.3820-YFP) and tetracycline (tet) inducible stem-loop RNAi against the mRNA of the hypothetical protein (Tb927.8.3820). I induced RNAi construct targeting Tb927.8.3820 using 200ng/ml of tet. The induction of RNAi did not have any effect on parasite growth (Figure 3.2.20B), although protein was depleted by 48 hours (Figure 3.2.20A-left). There was no effect on the V5-ZC3H30 protein level (Figure 3.2.20A- right). Since 100% protein depletion did not occur, I cannot comment on the essentiality of the hypothetical protein (Tb927.8.3820). To answer this question, one must try to obtain knockout cell lines of Tb927.8.3820. If one obtains cell lines lacking the gene (Tb927.8.3820) that grows identical to the wild type cells, then it is not essential for normal growth of trypanosomes in a certain life stage.

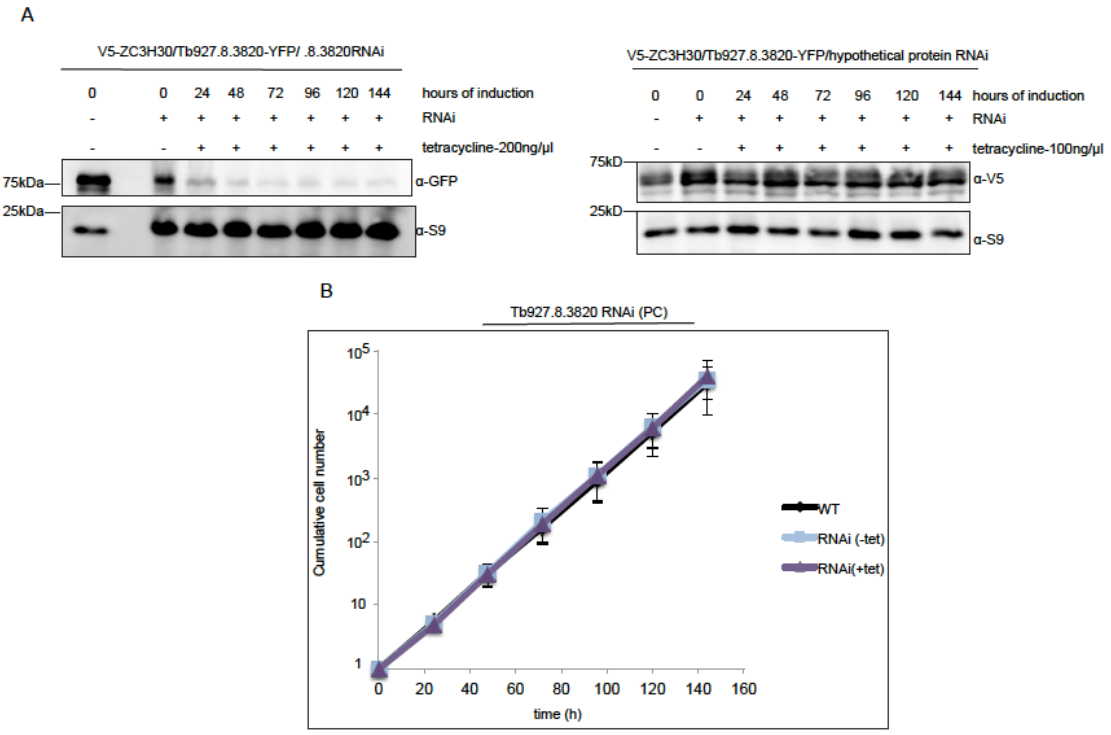


Figure 3.2.20: Depletion of hypothetical protein with RNAi has no effect on trypanosome growth and on homeostatic protein levels of V5-ZC3H30. (A). Western blots showing depletion of hypothetical protein (Tb927.8.3820) (left), and the effect of depletion on protein levels of V5-ZC3H30 (right). Procyclic trypanosomes expressing V5-ZC3H30, Tb927.8.3820-YFP (hypothetical protein) carrying stem loop construct for RNAi against the hypothetical protein were used for this experiment. RNAi was induced with 200ng/ml of tetracycline. Cells were regularly diluted after 24 hours to maintain log phase. Equal amount of cells were loaded on SDS-PAGE gel from uninduced (0 hour, - tet) and induced (+ tet) samples from several time points. The proteins ZC3H30 and the hypothetical

protein were probed using anti-V5 (α-V5) and anti-GFP (α-GFP) antibody; anti-S9 antibody used to probe ribosomal protein S9 was as a loading control. **(B).** Cumulative assessment on growth upon depletion of Tb927.8.3820. Log phase growing cultures was diluted every 24 hours and tetracycline (200ng/ml) was used to induce RNAi. The uninduced RNAi cells and the cell line not harbouring stem loop RNAi construct served as control. The plot shows the result from 2 independent experiments, using two biological replicates for each cell line. If no error bar is visible, it is because the standard deviation value was too small.

Hypothetical protein localises in the stress granules upon heat shock

Tb927.8.3820 protein is enriched the starvation granules [112]. Since it interacted with ZC3H30, which is also localised to heat shock granules, I investigated whether Tb927.8.3820-YFP, also migrated to the heat-shock granules. I heat shocked the cells at 37°C and isolated the granules by following previously published protocols [106, 112] and found that indeed, upon heat shock the hypothetical protein colocalised with Dhh1 and SCD6 in the stress granules (Figure 3.2.21).

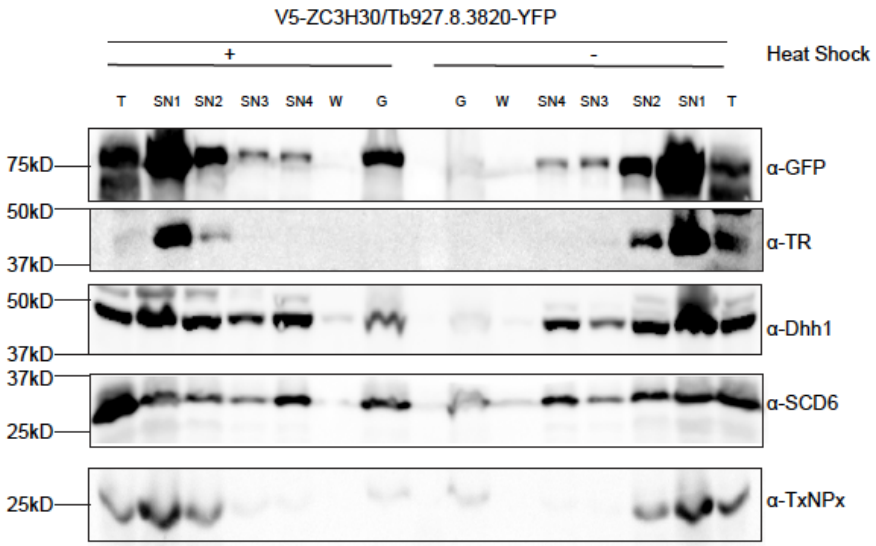


Figure 3.2.21: Hypothetical protein encoded by Tb927.8.3820 localizes in stress granules upon heat shock. About equal amount of procyclic cells expressing *in situ* YFP-tagged hypothetical protein either heat shocked (37°C) or not (27°C) were harvested for heat shock stress granule enrichment as previously described [106, 112]. Antibodies against SCD6 and Dhh1 serve as positive controls for stress granule formation, and antibodies against trypanothione reductase (TR) and peroxiredoxin (TxNPx) serve as negative control, as they are cytoplasmic and do not relocate to granules upon stress. Mouse anti GFP antibody was used to investigate the localization of hypothetical protein upon heat shock. The lanes indicate the same fractions with the same designation as in 3.2.16.

Depletion of hypothetical protein has no effect on migration pattern of ZC3H30 to stress granules

The hypothetical protein encoded by (Tb927.8.3820) is spanned with 5 low complexity regions. Proteins with such structure are capable of aggregation and

phase separation [140-144] and concentrates associated proteins to subcellular compartments. Hence I asked myself, whether the hypothetical protein concentrates ZC3H30 to stress granules during heat shock. For this experiment I used cell lines expressing stably the stem-loop RNAi construct against the hypothetical protein (Tb927.8.3820), an endogenously tagged C-terminal YFP-Tb927.8.3820 and a N-terminally tagged V5-ZC3H30. RNAi was induced with 200ng/ml of tetracycline for 96 hours. Then one culture were heat shocked at 37°C for 1 hour and the other culture was kept at normal growth temperature (27°C). Stress granule enrichment was done as previously described [106, 112]. Western blot samples were collected for enrichment from each fraction as previously described and loaded on a 12% SDS-PAGE gel, followed by Western blotting. The blot was probed for V5-ZC3H30, stress granule marker Dhh1 and also for a cytoplasmic protein trypanothione reductase (TR) known not to migrate to stress granule. It was observed that V5-ZC3H30 still colocalised with Dhh1 in the stress granules, upon RNAi of the hypothetical protein followed by heat shock (Figure 3.2.22 C). This indicated that maybe the localisation of ZC3H30 into the stress granules was independent of the interaction with the hypothetical protein (YFP-Tb927.8.3820).

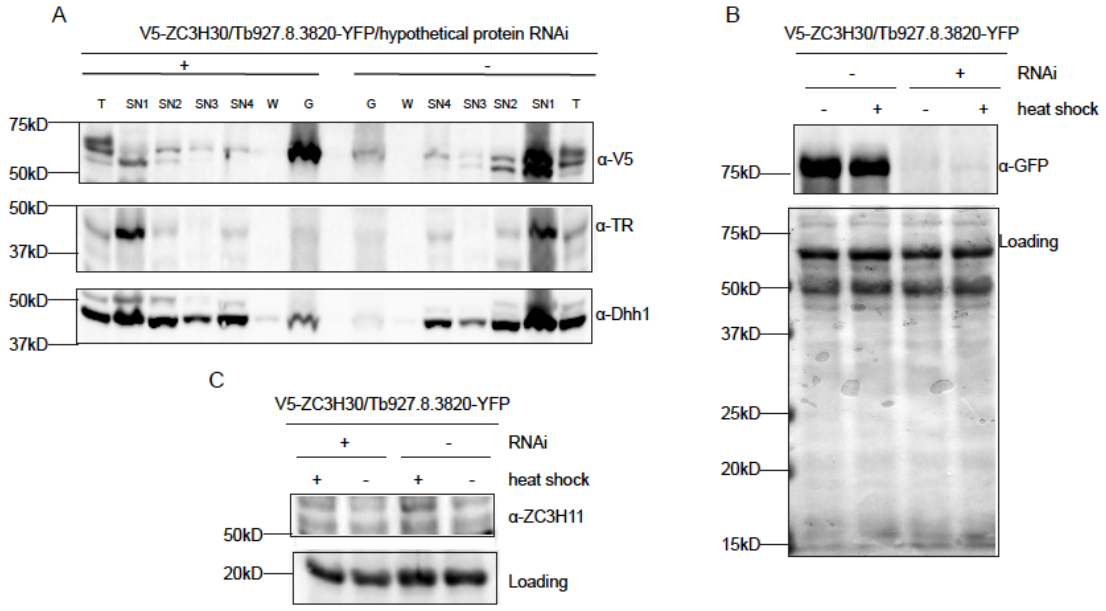


Figure 3.2.22: The migration of ZC3H30 to stress granules is independent of its association with the hypothetical protein (Tb927.8.3820). RNAi against the hypothetical protein was induced for 96 hours prior to the experiment. **(A).** Localization of N-terminal V5-tagged ZC3H30 on RNAi of C-terminal YFP tagged hypothetical protein (Tb927.8.3820) upon heat shock at 37°C. Equal number of PCs were heat shocked or not for 1 hour at 37°C. Heat shock induced stress granule isolation was done according to previous publication [106, 112]. Protein samples corresponding to equal cell number were collected during stress granule isolation, and were analyzed on a Western. The lanes indicate the same fractions with the same designation as in 3.2.16. Localization of N-terminal V5-ZC3H30 was determined using anti V5 antibody as probe. Antibodies against Dhh1 and SCD6 served to confirm the localization of the stress granule markers under stress. Antibodies against cytoplasmic markers like trypanothione reductase (TR) and peroxiredoxin (TxNPx) served as negative controls. **(B).** A Western blot showing efficiency of hypothetical protein knockdown post 96

hours and expression of the hypothetical protein upon heat shock (37°C). RNAi was induced with 200ng of tetracycline (tet). Protein depletion was checked by probing the blot with anti-GFP (α-GFP) antibody. Ponceau stained blot is used as a loading control. **(C).** ZC3H11 induction upon heat shock after 96 hours of RNAi of hypothetical protein. RNAi was induced with 200ng/ml of tet, and heat shock was done at 37°C.

Role of the hypothetical protein in combating stress conditions

The association of the Tb927.8.3820 hypothetical protein with heat shock stress granules and starvation granules and its interaction with ZC3H30 made me wonder if the protein had a role in combating stress. Cells expressing stem-loop RNAi construct (post 96 hours of RNAi induction with 200ng/ml of tetracycline or tet) against the hypothetical protein were stressed with heat shock (37°C) and with oxidative stress reagent Sodium arsenite (10μM- final concentration in MEM-Pros medium) for 1 hour. Then they were either shifted back at normal growth temperature or were returned to normal medium. The recovery and growth of the wild type (Wt), uninduced RNAi (RNAi), and induced RNAi (RNAi +tet) cell lines were monitored for 48 hours. A minimal growth defect for cell lines with RNAi constructs against the wild type cells (Figure 3.2.23 A and B) was observed. Both the uninduced and induced RNAi cell lines were equally affected by stress, possibly because of the leakiness of the construct in the uninduced RNAi cell lines. Since the depletion is not 100%, the remaining protein could have been enough to relieve the trypanosomes from heat shock and oxidative stress.

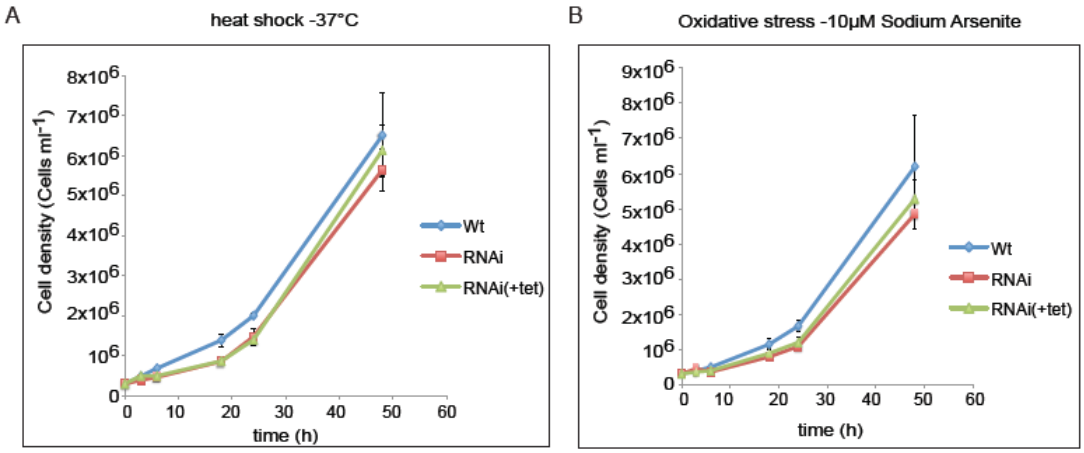


Figure 3.2.23: Effect of stress on procyclic trypanosomes with depleted hypothetical protein. (A). Recovery of PC with or without tetracycline induced RNAi construct after 1 hour of heat shock at 37°C followed by return to 27°C. The density of the culture was monitored for 48 hours. Results from two experiments and two biological replicates are represented as arithmetic mean ± standard deviation. **(B).** Recovery of PCs with or without tetracycline induced RNAi construct after 1 hour of treatment with 10μM Sodium arsenite followed by recovery and growth in MEM-Pros medium. The recovery and growth were monitored as in A.

Transcriptomic studies

To find out how ZC3H30 prevents stress induced lethality in trypanosomes, I decided to look for mRNAs that are bound and regulated by ZC3H30. I used tandem affinity purification method with the lysate from the cell line expressing N-terminal *insitu* TAP-ZC3H30, and lacking the other endogenous copy. The efficiency of the RNA immunoprecipitation is shown by a Western blot in figure 3.2.24A (Top). In order to assess whether the pull down was contaminated by rRNAs, and whether the eluate contained mRNAs, Northern blots were prepared (Figure 3.2.24A- bottom). The integrity of the mRNA pulled down by ZC3H30 was checked using an abundant mRNA encodin alpha-tubulin (*TUBA*), and presence of mRNAs was confirmed by the signal obtained from splice leader (*SL*). The rRNA from the unbound fraction was depleted by RNase H treatment using anti-rRNA oligos, and all samples were set for RNA-Seq. During analysis of the RNA-Seq data, I found that one of the unbound fractions RNA was distinctly different from the other, and now it is being reanalysed with the RNA from the Wt culture.

To identify which mRNAs are effected or regulated directly or indirectly by ZC3H30, I decided to look at the total mRNA levels in double knockouts and double knockouts expressing ectopic copy of ZC3H30, we did RNA-Seq and data analysis on total RNA from non heat shocked (27°C) and heat shocked samples (39°C). I isolated total RNA from heat shocked trypanosomes and cells, which were healthily growing at normal temperature. Figure 3.2.24B (top- left and right) shows a Northern blot and shows efficient rRNA depletion from total RNA. The blot was probed against *tubulin* (*TUBA*) to check the integrity of stable mRNAs. Since rRNA depletion is carried out at room temperature, there is a high probability of RNA degradation. Too much degradation can cause bias while data analysis. Figure 3.2.24 (bottom) shows a Western blot to demonstrate induction of ZC3H11 upon heat shock at 39°C for the same samples used for RNA Seq. A non- specific band at 20 kD was used as a loading control. Genes with very low read counts were removed.

Presently the reads from the samples are being compared with the reads of Wt PC RNA-Seq data at 27°C and at 39°C. I expect that mRNAs bound by ZC3H30, should be more abundant in DKOs compared to the DKO ectopics and Wt, since ZC3H30 is a repressor. A detailed study on the functions of the target mRNAs is expected to illuminate why trypanosomes lacking ZC3H30 are sensitive to stress.

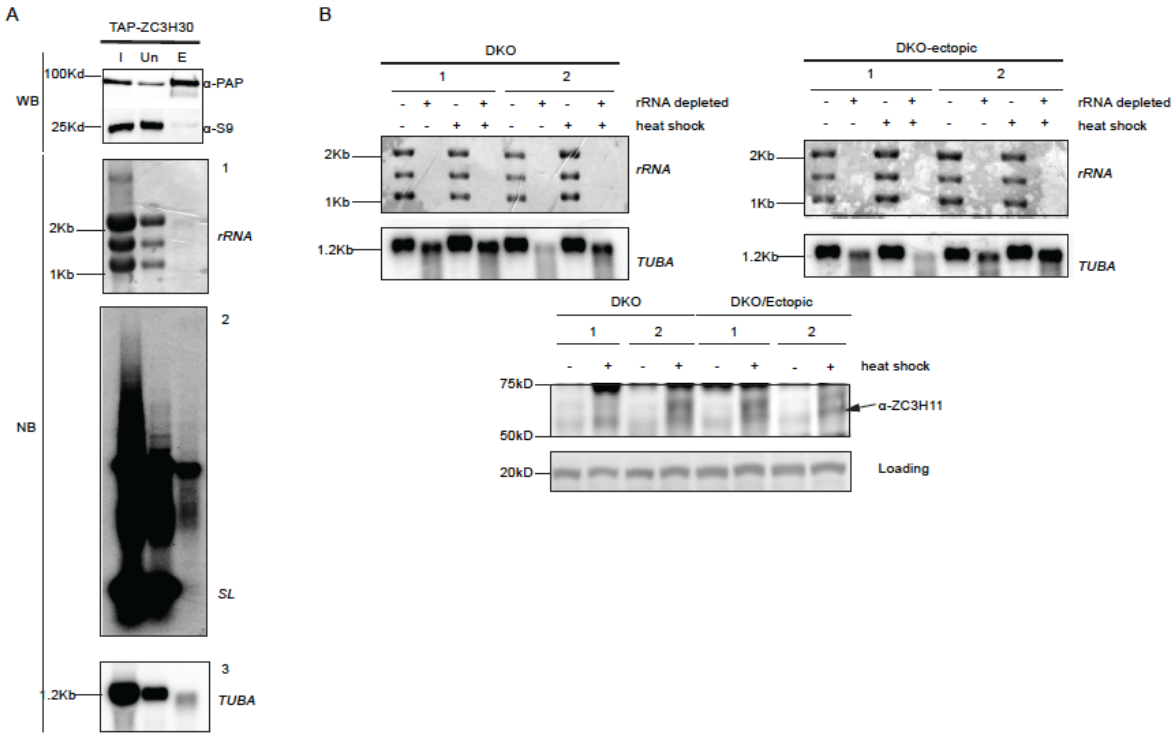


Figure 3.2.24: ZC3H30 transcriptomic studies. (A) TAP purification to pull down mRNAs bound to ZC3H30. Log phase PC TAP-ZC3H30 cells were harvested and processed forpulling down bound mRNAs as described in'RNA-Immunoprecipitation' (Materials and Methods). On top is a western blot (WB) to show efficiency of purification. TAP-ZC3H30 was detected using anti-PAP (α-PAP) antibody, and antibody against ribosomal protein S9 (α-S9) was used as a loading control. The lanes on WB and northern blot (NB) represent: I- input, Un-unbound and B- bound ZC3H30 to beads fraction. The NB (bottom) shows (1) efficiency of mRNA pull-down specifically (rRNAs were used as loading control), (2) presence of mRNAs in I, Un and E (SL was used to probe for total mRNA), and (3) integrity of mRNAs after purification (Tubulin probe-TUBA was used to check the integrity of the mRNAs). **(B)** Total RNA from non-heat shocked and heat shocked cells (DKO and DKO-ectopic) for RNA-Seq. Top (right and left) shows Northern blots of total RNAs extracted from log phase growing PC cultures heat shocked (+) (39°C) and non-heat shocked (-) cells, before (-) and after rRNA depletion (+). Tubulin probe-TUBA was used to check the integrity of the mRNAs after rRNA depletion. The two DKO and DKO-ectopic clones are designated as 1 and 2 respectively. Bottom shows a WB to show induction of ZC3H11 protein upon heat shock. Cytoskeleton free total protein was extracted and proceeded for western blotting. Anti-ZC3H11 (α-ZC3H11) was used to identify ZC3H11 endogenous protein (indicated by arrow). A non-specific band was used as a loading control.

Chapter 4-Discussion

Part1:

Putative partners of the CAF1-NOT complex of trypanosomes

Methods resulting in better understanding of protein-protein interactions are relevant to discover the biological function of the protein. A complete understanding of a certain biological or cellular process requires the map of binary protein interactions. Also, such maps can predict the novel function of a yet uncharacterized protein. Although methods for detecting protein-protein interactions from a huge group provide a lot of information, they cannot detect all interactions and are also not free from false positives. One of the most widely used methods to detect protein-protein interaction is by yeast-two-hybrid [145]. One of the interacting partners is fused with the DNA-binding domain (BD) of GAL4 transcription factor, and the other with its activation domain (AD). The BD domain binds to the GAL4 promoter and the AD domain is responsible for activation of transcription of the downstream gene. If the two partners interact they bind to the promoter and activate the transcription of downstream reporter genes. Usually it is done by mating a strain expressing the BD fusion protein with another yeast strain expressing AD fusion protein, so that The interaction is detected in diploids [146]. The strategy of yeast-two-hybrid mating has been used in countless genome wide screens to identify protein-protein interactions, for example in T7 bacteriophage and Vaccinia virus [147, 148]. I believed that yeast-two-hybrid mating screen would give me an overview as which novel RBPs might interact with the CAF1-NOT complex and target mRNAs for deadenylation in both of the life-cycle stages. Therefore, I applied the strategy of identifying only those genes to be potential candidates that have a high RPM, and that has not been identified in other screens. For better detection of a potential candidate gene, I looked at the list of genes that were identified in all screens, and from these I identified genes that were common in all the screens. I named these genes as contaminants and identified them in the list of my candidate genes identified in the screens and removed them.

Then I looked only for candidates that appeared more than once in the screen and their reads were found in more than one location in the ORF. These are supposed to be the potential interacting partners of the subunits of the CAF1-NOT complex. After removing genes which appeared only once per screen, I obtained 6, 3, 158 and 15 putative partners for NOT2, CAF40, NOT10 and NOT11. Published data from the genome-wide tethering screen [95] was used to look for candidates, which can positively or negatively regulate reporter gene expression. I found 4 proteins that specifically interacted with NOT2, 8, which interacted with NOT10 specifically, 11 interacting only with NOT11 and only 1 protein, which interacted specifically with CAF40. I compared the data with published data on differentiation, RNA biology and stress. Out of these genes, ZC3H30, hypothetical proteins expressed by Tb927.9.7690, and Tb927.6.5010 also bind to mRNAs [94]. The RBP ZC3H29, the gene identified in the screen for NOT10, was enriched in the granule fraction of starvation granule [112] and is essential in BS and parasites capable of

differentiation [125]. The NOT10 interactome data from YTH was compared with the affinity purified protein pull-down data of NOT10. We found 4 common genes namely, NOT1, NOT5, NOT11 (components of the CAF1-NOT complex), and 3-ketoacyl CoA thiolase (Tb927.8.2540). The appearance of a gene related to lipid metabolism, associating with the deadenylation machinery, is strange, as we cannot decipher the function of such an association. The only explanation I have is probably, this protein is abundant, and its primary structure might be hydrophobic in nature, and hence its association to other proteins is random, rather than functional.

The screen was done with the idea of obtaining genome-wide protein interacting partners of the CAF1-NOT complex. There can be several other factors that might interfere with protein-protein interaction. In the humans, CNOT10 has intrinsic RNA binding property [85], and in Drosophila, the interaction with Nanos and the CCR4-NOT complex developmentally regulates the fate of mRNA. This implies that interaction with many of the binding partners of the deadenylation complex is RNA-dependent. Therefore it is absolutely possible, that the trypanosome CAF1-NOT complex subunits and its protein partners might not find a mediating partner mRNA via which the two proteins come in close vicinity and interact. This is perhaps the most important reason, why I obtained so few specific interacting partners. Although, the yeast-two-hybrid mating screen is a quick method to obtain protein-protein interaction partners, it does not have the advantage of identifying all interaction partners, especially those interactions that are weak and transient or those that involve protein modifications (eg., acetylation, methylation, phosphorylation etc). A post-translational modification, like methylation, acetylation and glycosylation can play a significant role in protein-protein interactions. Acetylation of CAF1 by HATs, p300 and CBP induce degradation is important for adipocyte differentiation [48]. The high probability of the NOT complex subunits not properly modified for functionality might also be the ground for not obtaining interactors, which interact only with the modified isoform. The putative specific interactors of the complex that have been listed are the ones, which interact with the subunits in their unmodified isoform. Most of the proteins with WW-domain, spectrin, pterin domains are promiscuous. Usually they serve as cytoskeletal interaction coordinators [149], or cofactors of enzymes. They are abundant and interact without specificity; so whether they have an actual role in NOT interactions and whether this interaction is biologically relevant is questionable. Nevertheless, to know further answers, I suggest, that co-immunoprecipitations in trypanosomes, with lysates from both bloodstream and procyclic form should be carried out, and experiments on the biological relevance of this interaction should be carried out. Also, I believe that high throughput experiment should be corroborated with at least 2 different methods. I believe that an experiment like affinity precipitation followed by mass-spectrometry could be helpful in scrutinizing and identifying actual interactors of a complex.

Conclusion

I have screened for genome-wide interacting partners of the CAF1-NOT complex for the subunits, CAF40, NOT2, NOT10 and NOT11. I have provided a list of potential partners that might interact with the complex in vivo either in the bloodstream form or in the procyclic form cells or both. Further investigation should be carried out to validate the interaction and find function to it; either in complex formation, regulation of subunit function, targeting to mRNAs for degradation etc.

Part2

ZC3H30 is a cytoplasmic RBP involved in repression of mRNA

The results from this study show that zinc finger RNA binding protein ZC3H30 localises in the cytoplasmic fraction. I used BS cells, which had an in situ N-terminally TAP-tagged ZC3H30. I used cell fractionation by SiC where I separated the organellar fractions from the cytoplasmic soluble fraction and found ZC3H30 localising with the cytoplasmic marker trypanothione reductase (TR) and not in the organelles. Since, ZC3H30 was previously identified as a post-transcriptional regulator involved in reporter mRNA repression [94, 95], I expected that my protein would be found in the cytoplasmic fraction, unless it was also involved in splicing, then it would be present both in cytoplasm and in nucleus.
The tethering screen identified the C-terminal part of ZC3H30 was involved in mRNA repression [95]. Since artificial tethering with a partial fragment of a protein may give rise to faulty results, I did tethering assay with the full ZC3H30 protein in the BS. This gave me the result that the whole ZC3H30 protein could repress reporter mRNA (*CAT*) expression both by destabilisation (*CAT* mRNA levels were down) and translational repression (CAT protein levels were low). What was intriguing was that the C-terminal half of the protein had the same affect on the *CAT* gene expression as the full-length protein, whereas the N-terminal half, which had two zinc finger domains and two LCRs could destabilise the mRNA. What has not been investigated is, whether ZC3H30 dimerizes under normal conditions, and which part of the protein is responsible for dimerization. Because, if the C-terminal half of ZC3H30 dimerizes with the endogenous full-length protein, then that would explain why only the C-terminal half of the protein has the same effect on reporter mRNA like the full length protein. I also tried to replicate the tethering assay in PCs, but ZC3H30 had no effect on the reporter expression. In our lab it has been observed that the tethering assay for putative RNA binding proteins with milder effect on reporter mRNA expression in PCs show no detectable effect in tethering assay. Probably ZC3H30 has lower reporter mRNA repression activity in procyclics than in bloodstream form, and hence I could not get a decent effect on CAT reporter protein and mRNA.

ZC3H30- role in differentiation or infection?

ZC3H30 is not essential for the survival of trypanosomes either in BS or PC life-cycle stages. Another important piece of information we know about this protein (besides being a repressor [94, 95]) is derived from the proteomic studies to identify all protein of the starvation granule [112]. In a recently published article from our laboratory, we also showed that it interacted with the CAF1-NOT deadenylation complex partner CAF40 [93]. Therefore, I was interested to find out what was the role of this protein in the trypanosomes. In the differentiation experiment I saw that the BS knockouts of ZC3H30 and the BS Wt could be induced to express EP using 6μM cis-Aconitate. I do not have the knockouts in pleomorphic EATRO1125 cells, so I cannot comment whether ZC3H30 DKOs could transform to PCs. Whether ZC3H30 has any role during differentiation from BS to PC can only be answered if ZC3H30 DKO

pleomorphic cells are subjected to differentiation, followed by studying of various differentiation markers like PIP39 [150], RBP10 [100] and RBP6 [101]. Another recent study was of gene expression regulation in metacyclic trypanosomes (mammal-infective parasites found in the salivary gland expressing metacyclic VSGs). To induce metacyclic stage they induced the expression of RBP6 and then compared the total RNA and proteins with the RNA and proteins from PC. I looked into the data to see if ZC3H30 expression was different in PC and the metacyclics, and found that there was no difference [151]. Although detail studies like infection of ZC3H30 DKO metacyclic in mice can be done, but after comparison of several published data [125, 151] I can vaguely suggest that probably ZC3H30 does not play a role in differentiation.

Expression of ZC3H30 is essential to survive stress

Since I did not obtain any defect in EP induction in the DKOs, I decided to investigate whether ZC3H30 had any role in stress regulation in trypanosomes. In the meantime, a study on the composition of starvation granules was published. This study identified for over 400 proteins, which localises in the granule fraction of the parasite cell under short-term starvation [112]. One of the RBP identified in this study was ZC3H30. Hence I stressed the BS ZC3H30 DKO cells at 41°C and 42°C and found that the DKOs could not recover as fast as the Wt BS cells after the heat shock. Also, it was clear that at 41°C the DKOs took longer (48 hours) to recover and reach similar cell numbers as the Wt. I tried to induce unfolded protein response (UPR) accompanied by splice leader silencing (SLS- leads to cell death [113, 114]) at a very low dose of DTT (0.1mM) for a very short time (1 hour), and noticed that the ZC3H30 BS DKOs suffered much more than the wild type cells and never recovered and attained the similar cell numbers like the Wt Bs or the BS DKO-ectopics, indicating that truly the lack of ZC3H30 expression resulted in more cell death during SLS stress. Although, I tried to replicate the SLS with higher concentrations of DTT, but this led to cell death of my Wt BS cells and their recovery was delayed over 24 hours, hence I decided to work at a lower concentration of DTT.
Next, I decided to study whether ZC3H30 had any role in stress regulation in PCs. Stress regulation in PC trypanosomes has been very well studied. We know that the RBP ZC3H11 is a common stress marker, which is upregulated and phosphorylated upon stress [104, 105]. Also, it is inessential for normal growth of PCs and is important for their survival under heat shock conditions, and RNAi against ZC3H11 relocalises HSP70 mRNA from cytoplasm to the heat shock granules [106]. Since ZC3H30 also was not essential for normal growth of PC trypanosomes, I was curious to know the behaviour of the DKOs upon stress. Upon heat shock at 37°C, there was a marked growth defect in the PC DKOs, and severe stress at 41°C was completely lethal for the DKOs. I checked the induction of ZC3H11 at both temperatures and found that the protein was up in Wt, DKOs and DKO-ectopic, marking stress. This trend of DKOs being susceptible to stress was also found in Ethanol stress and oxidative stress with Sodium arsenite. In Ethanol stress a higher concentration from 1-2% Ethanol addition to the medium did not induce a greater growth and recovery defect, but increase in concentration of Sodium arsenite from 10 to 20μM resulted in a huge growth defect. Ethanol is known to cause oxidation and fragmentation of

proteins and disturb the metabolic balance [152, 153]. Probably application of a higher concentration of Ethanol would show increase in growth defect.
Previous studies on another RBP DRBD3 in PCs showed that the protein migrates from cytoplasm to nucleus upon oxidative stress induction with 50μM Sodium arsenite [109]. I tried to replicate this condition, and saw that this concentration of sodium arsenite was lethal for the DKOs, but Wt PCs and DKO-ectopics also took longer than 24 hours to recover and so I decided to use lower concentration of Sodium arsenite. I found out that at both 10 and 20μM Sodium arsenite induced oxidative stress, PC Wt and DKO-ectopics recovered completely within 24 hours and resumed normal cell multiplication rate, but DKOs suffered extensively. At 10μM Sodium arsenite concentration, the cells suffered and died and so never reached the same number of cells after 48 hours of recovery period. But, the condition was even worse at 20μM, since the cells did not double even after 48 hours of recovery. Probably there was more damage and cell death at this condition. All these results indicated that indeed ZC3H30 expression was required in both life stages under stress conditions.

Chaperones and stress markers are not affected due to the lack of ZC3H30 expression

I wanted to know if ZC3H30 had any effect on chaperones both at protein and mRNA levels. Hence I did northern blot analysis to know if lack of ZC3H30 affected mRNA levels of HSP70, since at heat shock it is known to be stabilized and up regulated in trypanosomes [105, 138]. The RBP ZC3H11 is known to bind the AU riche element of the 3'UTR of HSP70 and stabilize it for translation during heat shock. Depletion of ZC3H11 results in lowering the abundance of *HSP70* mRNA in steady state conditions [105], also, upon heat shock, *HSP70* mRNA localises in the heat shock granules [112]. I thought since the DKOs are susceptible to stress, probably there is a dysfunction in chaperone mRNA regulation during stress. But, I saw nearly two-fold increase in *HSP70* mRNA (Tb927.11.11330) in ZC3H30 DKOs, both in BS and PC (at least twice). But, the RNA-Seq analysis from the total transcriptomic analysis for knockouts and knockouts expressing ectopic ZC3H30 did not show the same result. It might be that with more time the DKOs get adapted to the culture and therefore the increase in *HSP70* mRNA goes away. We also believe that *HSP70* mRNA is not a direct target of ZC3H30, as it was not enriched in the bound fractions in the RNA-immunoprecipitation experiment for TAP-ZC3H30. Nevertheless, if an RBP that is responsible for decay of *HSP70* is depleted or deleted, those cells should be more resistant to stress. Also increase in *HSP70* mRNA might result in stress resistance unless these *HSP70* mRNAs were not translated. The antibody against HSP70 protein was not helpful as it identified a lot many unspecific bands.
Although I found expression of ZC3H30 to be essential under ER stress in BS, I could not infer the same for PC. I used dithiothreitol (DTT) as a chemical ER stress agent. Although 0.1mM DTT in medium for 1 hour was sufficient to stress bloodstream form cells followed by rescue, it was a lethal dose for the procyclic forms. At this dosage, our procyclics could not be rescued, and longer exposure period to DTT in bloodstream and procyclic form turned out to be lethal. Maybe this was also, one of the reasons that I could not see an up-regulation of ER stress chaperone BiP, because previously published data suggests that BiP requires a

higher dosage of DTT for a longer period of exposure, like 4 hour so that BiP is expressed in quantitatively higher amounts [113].
Also, lack of ZC3H30 expression had no impact on common stress regulating factors like ZC3H11, HSP60, BiP (ER HSP70), and SCD6 protein levels upon stress. This indicated that probably ZC3H30 did not participate in the regular stress regulation pathway involving chaperones and RBPs regulating chaperone expression. Probably, the growth defect phenotype is a secondary effect due to ZC3H30 gene deletion. Hence I decided to identify other phenotypes of the DKOs (translation) and interacting partners of ZC3H30. I expected that probably these would help me to identify what role does ZC3H30 play in the trypanosomes, dysfunction of which results in lethality of ZC3H30 DKOs under stress conditions.

Does ZC3H30 affect translation?

Since the lack of ZC3H30 expression resulted in cell death in trypanosomes under heat shock, oxidative stress, Ethanol stress and SLS, I decided to look for other stresses, where the DKOs would not be lethally affected. These would then indicate the pathways not affected by ZC3H30. With this hypothesis in background, I stressed the PC DKOs and the Wt with Hygromycin, a translational inhibitor. Under the stress of Hygromycin (50µg/ml) PC ZC3H30 DKOs survived better; also their IC_{50} was higher (about 10x). This made me ask myself how does Hygromycin kill cells, and what possible role does ZC3H30 have in combating Hygromycin mediated cell killing. I supposed that ZC3H30 and Hygromycin have a negative relationship to each other. I hypothesised that since the ZC3H30 DKOs were more resistant to Hygromycin mediated killing, probably ZC3H30 was combating the impact of Hygromycin. Also, what I realized was all my ectopic tet inducible constructs, which expressed ZC3H30 with Hygromycin as a resistant marker, were leaky. They were expressing ZC3H30 even in absence of tet. This is possible if Hygromycin stresses the cells and then they would express ZC3H30, because the protein was helping in combating the negative effect on the cells. In the meantime, I also tested what would happen if I stressed cells and looked at ZC3H30 expression. The expression of ZC3H30 was up in heat shock at 37°C, 41°C, 10µM arsenite and even in presence of Hygromycin, but not when I stressed the cells with another translational inhibitor Puromycin. What was interesting is that at 50µg/ml Hygromycin concentration, the levels of ZC3H30 was still up (when compared to ribosomal protein S9), but the levels of ZC3H30 was not up in presence of 0.5 or 1µg/ml of Puromycin. Hygromycin B inhibits translation by interfering with EF2 function, or displaces it from the translating mRNA, and also stabilizes EF1 dependent Phe-tRNA binding to the acceptor site, which results in the induction of misreading [154]. Puromycin is aminonucleoside, which mimics aminoacyl-tRNA. It binds to the acceptor site of translating ribosomes, resulting in the formation of amide bond between the C-terminus of the nascent polypeptide and the primary amino group of Puromycin [155]. This made me hypothesize that probably ZC3H30 was involved in regulation of stalled or miss-read mRNAs and not in chain termination regulation during translation. If we think of a situation where the drug Hygromycin binds to amino-acylated tRNA and stabilizes the binding thereby stalling the ribosome, there will be incomplete protein synthesis, and the partial protein translated would be non-functional. If ZC3H30 is a repressor of mRNAs, with such

stalled mRNAs, or mRNAs which are being miss-read then probably the expression of ZC3H30 needs to be up to combat this situation. This is a possible explanation to what ZC3H30 might be involved in trypanosomes. Usually in other eukaryotes faulty mRNAs with pre-termination codon (PTC) undergo non-sense mediated decay (NMD) [156], or mRNA without termination codon are degraded in a tmRNA dependent manner [157-159], or translating mRNP complexes with defect in elongation get degraded by no-go mediated decay [160]. In *Trypanosoma brucei*, the classical NMD mediated mRNA degradation is missing [161]. The non-stop decay (NSD) and no-go decay (NGD) aspects have not been explored in the parasites. So, I cannot actually hypothesize and comment on whether ZC3H30 might be involved in any of these decay processes. In order to be certain, that ZC3H30 might be involved in NSD or NGD, one must treat DKOs, DKO-ectopics and Wt with Cyclohexamide and isolate RNA, and look for accumulation of mRNAs. In presence of ZC3H30, probably stalled ribosome bound mRNAs will be degraded, whereas in its absence there will be an accumulation of mRNAs with stalled ribosomes.
Also, ZC3H30 might be involved in general translation repression of mRNAs. I did RIP-Seq on pulled mRNAs by affinity purification of TAP-ZC3H30, and RNA-Seq on total mRNAs from ZC3H30 DKOs and DKO-ectopics. When the bound mRNAs were compared to the RNA-Seq data from the ectopics, they did not show any difference but, that long mRNAs were enriched. Also, I did not find any sequence motif enriched in the 3'UTRs. Also there was no mRNA class enrichment between the ZC3H30 DKOs and DKO-ectopics. So, I cannot conclude that ZC3H30 regulates a specific group of mRNAs, or whether it regulates translation of mRNAs. Presently, the RNA-Seq data from the DKOs and DKO-ectopics is being analysed with the Wt PCs. Also the project will be further proceeded in doing RIP-Seq and RNA-Seq in BS ZC3H30 DKOs, followed by S^{35}-Methionine pulse labelling assays under stressed conditions to understand if ZC3H30 regulates basal translation. This would explain why in absence of the gene the cells are stressed to SLS, heat shock and oxidative stress.

ZC3H30 and hypothetical protein (8.3830) participate in a pathway that rescues procyclic trypanosomes from stress

A study was recently conducted to find out all mRNAs and proteins associated with the trypanosomal stress granules and P-bodies. Among several novel RBPs, ZC3H30 was also found enriched in the granule fraction [112]. This experiment was done under starvation and so I wanted to find out if ZC3H30 relocalised to the heat shock induced stress granules and if the relocalisation of ZC3H30 was a general phenotype. I followed the protocol as published in [112], [106], and found out that heat shock could also result in the relocalisation of ZC3H30 into the stress granules. Another well known RBP, known to be a translational repressor and localising in the stress granules is SCD6 ([162], [163]). The affinity purification of ZC3H30 also pulled down a hypothetical protein, which was also enriched in the granule fraction from the starvation stress granule enrichment study. The hypothetical protein has no conserved domains; instead it has 5 low complexity regions (LCR) spanning the entire primary structure of the protein. Proteins with LCRs are known to form aggregation under stress stimuli or lead to the formation of potential pathologic aggregates in neuro-degenerative diseases. In yeast, these amyloid aggregates are

solubilised by HSP104 complex [164]. Most RBPs carry low complexity regions, besides their RNA recognition or binding motif. It is the LCR in the RBPs, which is responsible for the formation of in vitro aggregates [165]. Also, these LCRs act like environment sensors for proteins, like the LCR of Pab1, which helps the protein to undergo phase separation under stress [166]. ZC3H30 itself spans 3 LCRs and interacts with hypothetical protein (8.3820). I asked myself if the hypothetical protein is responsible for the phase separation and granule localisation, ZC3H30 undergoes during stress. Therefore, I did knockdown studies on the hypothetical protein and looked at the localisation of ZC3H30 after the granule was formed. I have to admit, that the knockdown was not 100%. Also, because of the detection limit of the antibody, I might not be able to completely assess the amount of the hypothetical protein actually left after the RNAi. If a small quantity of hypothetical protein that has the ability to associate with ZC3H30 and bring about the phase separation, then ZC3H30 will migrate to the stress granules upon heat shock. Also, ZC3H30 has its own LCRs, which might be enough to insolubilize ZC3H30 and relocate it to the stress granules. Therefore, I cannot infer anything upon the function of the interaction of the hypothetical protein and ZC3H30. What I can conclude from the stress induction in the RNAi cell line, is that knockdown of the hypothetical protein affects slightly the fitness of procyclic form trypanosomes under stress conditions. Since, the knockdown was inefficient the effect is also not remarkable. In order to be absolutely sure about the role of the hypothetical protein (8.3820) in stress, one must knockout the gene and examines growth and viability of trypanosomes under stress conditions.

Conclusion

My results conclude that the zinc finger RNA binding protein ZC3H30 is indeed essential in combating stress in both BS and PC form trypanosomes. In the bloodstream form cells the RBP regulates stress by promoting decay and translational repression. It is localised usually in the cytoplasm but relocates to subcellular foci, the stress granules during heat shock in the PCs. Expression of ZC3H30 might be important in regulating basal level translation, the dynamics of which might be essential for cell viability under stress.

Publications

List of publications based on the study

Erben E, Chakraborty C, and Clayton C (2014). The CAF1-NOT complex of Trypanosomes. frontiers of Genetics. doi: 10.3389/fgene.2013.00299

Chakraborty C, Fadda A, Erben E *et al.* Interactions of CAF1-NOT complex components from *Trypanosoma brucei* [version 1; referees: 2 approved with reservations]. *F1000Research* 2017, 6:858(doi: 10.12688/f1000research.11750.1)

Appendix I- A. Yeast-two hybrid screen-Genes involved in biological process

Names	total	Gene Ids	Annotation
Differentiation- RNA metabolism- stress	13	Tb927.5.3850	XRNE
		Tb927.11.6160	ERF3 eukaryotic release factor 3, putative
		Tb927.10.15360	hypothetical protein, conserved, weak DEAD domain/P-loop hydrolase
		Tb927.2.5160	chaperone protein DNAj, putative
		Tb927.3.1920	NOT5
		Tb927.7.2680	ZC3H22
		Tb927.10.13360	EF-Tu Mitochondrial elongation factor Tu
		Tb927.10.1510	NOT1
		Tb927.2.4370	trypanothione synthetase (TRYS)
		Tb927.8.1960	NOT11
		Tb927.2.4930	esterase, putative
		Tb927.6.5010	hypothetical protein, conserved
		Tb927.11.1450	2-oxoglutarate dehydrogenase E1 component, putative
Differentiation -RNA metabolism	30	Tb927.8.5580	VPS13; N-terminal region of Chorein, a TM vesicle-mediated sorter/Integral peroxisomal membrane peroxin,
		Tb927.8.2540	3-ketoacyl-CoA thiolase, putative
		Tb927.11.13360	AAA ATPase, putative
		Tb927.8.2640	ubiquitin-activating enzyme e1, putative (UBA1)
		Tb927.11.12850	oligopeptidase b (OPB)
		Tb927.4.1080	V-type ATPase, A subunit, putative
		Tb927.10.9190	isoleucyl-tRNA synthetase, putative (IleRS)
		Tb927.10.11140	hypothetical protein, conserved, kineto-specific, signal peptide
		Tb927.10.14180	TbSec13 Outer ring nucleoporin
		Tb927.3.5310	paraflagellar rod protein,I2 antigen
		Tb927.10.8540	chaperone protein DNAj, putative
		Tb927.3.5530	Tb-292 membrane associated protein
		Tb927.3.2050	hypothetical protein, conserved, no domains, poly(Q) tracts, K-specific
		Tb927.11.3190	hypothetical protein, conserved, LIS homology & WD40 repeat region matches a human HIV BPR binding protein
		Tb927.6.3930	Mitchondrial LSU ribosomal protein
		Tb927.2.2230	hypothetical protein, conserved, no domains, Kineto-specific
		Tb927.11.14250	Tric complex epsilon subunit (TCP-1-epsilon)
		Tb927.8.8120	hypothetical protein, conserved
		Tb927.7.6640	hypothetical protein, conserved, no domains, no yeast or human match, NADP binding domain Rossmann Fold
		Tb927.11.2640	ras-like small GTPase, putative (TbNST)
		Tb927.7.6270	peptidase t, putative
		Tb927.11.760	protein phosphatase 2C, putative
		Tb927.11.16830	IFT22 (was RABL5) (in UPF1-TAP)
		Tb927.11.15130	chaperone protein DNAj, putative
		Tb927.4.2430	Endonuclease/Exonuclease/phosphatase family, putative
		Tb927.8.3840	hypothetical protein, conserved
		Tb927.9.7690	hypothetical protein, conserved, nucelotidyltransferase and poly(A) polymerase substrate binding domains

		Tb927.1.4690	arginine N-methyltransferase (PRMT1)
		Tb927.11.6590	aminopeptidase, putative,metallo-peptidase, Clan MF, Family M17
		Tb927.9.5210	glutaminyl-tRNA synthetase, putative (28G16.175)
Differentiation-stress	13	Tb927.10.11590	nicotinate phosphoribosyltransferase, putative
		Tb927.7.2070	heat shock protein DNAJ, putative
		Tb927.10.1170	intraflagellar transport protein IFT172
		Tb927.7.1100	hypothetical protein, conserved
		Tb927.2.5030	transcription initiation protein, putative (30M24.65)
		Tb927.1.2570	coatomer beta subunit (beta-coP)
		Tb927.5.2700	otubain cysteine peptidase, Clan CA, family C65, putative
		Tb927.5.2590	Macro domain containing protein, putative
		Tb927.7.5310	YEATS family, putative
		Tb927.6.640	kinetoplastid-specific phospho-protein phosphatase, putative
		Tb927.4.450	coatomer alpha subunit, putative (TbCoatomerAlpha)
		Tb927.5.2320	Mak10 subunit, NatC N(alpha)-terminal acetyltransferase, putative
		Tb927.10.3500	RBSR4 = U2AF65
RNA metabolism-stress	2	Tb927.4.4060	hypothetical protein, conserved
		Tb927.10.1540	ZC3H30
Differentiation	138	Tb927.6.620	hypothetical protein, conserved
		Tb927.11.14490	RNA polymerase subunit, putative (RPB7)
		Tb927.7.6260	TPR repeat, putative
		Tb927.10.6190	aldehyde dehydrogenase, putative (ALDH)
		Tb927.7.5560	glucosamine-fructose-6-phosphate aminotransferase, putative
		Tb927.5.2930	Mitochondrial ATP synthase subunit, putative
		Tb927.7.4410	hypothetical protein, conserved
		Tb927.3.1190	hypothetical protein, conserved
		Tb927.10.9900	ABC1 protein, putative
		Tb927.9.13050	hypothetical protein, conserved
		Tb927.11.9190	protein kinase, putative
		Tb927.11.4030	hypothetical protein, conserved, no clear yeast or human matches
		Tb927.6.1430	hypothetical protein, conserved
		Tb927.7.1910	pyridoxal phosphate containing glycine decarboxylase, putative
		Tb927.9.15290	CHAT domain containing protein, putative
		Tb927.3.2150	protein phosphatase 2C, putative
		Tb927.1.1530	protein kinase, putative
		Tb927.8.1860	pitrilysin-like metalloprotease
		Tb927.11.1250	kinetoplast poly(A) polymerase complex 1 subunit, MIT ssu-associated
		Tb927.5.2100	RBP30
		Tb927.4.1340	CPSF73
		Tb927.11.10480	PQQ-like domain/WD domain, G-beta repeat/Utp21 specific WD40 associated putative domain containing protein, putative
		Tb927.11.15190	Domain of unknown function (DUF4200), putative
		Tb927.11.6370	leucine-rich repeat protein (LRRP)
		Tb927.3.1290	CUL4 cullin
		Tb927.11.14970	hypothetical protein, conserved
		Tb927.5.1940	hypothetical protein, conserved

Tb927.8.5730	SLK1 protein kinase
Tb927.11.7570	ATP-grasp domain containing protein, putative
Tb927.3.930	OADalpha outer arm dynein heavy chain
Tb927.9.5190	proliferative cell nuclear antigen (PCNA), putative (28G16.165)
Tb927.1.4050	ser/thr protein phosphatase with EF-Hand domains (PPEF)
Tb927.7.6610	hypothetical protein, conserved
Tb927.9.9010	SAC3/GANP/Nin1/mts3/eIF-3 p25 family, putative
Tb927.4.3390	tuzin, putative
Tb927.10.1420	metallo-peptidase, Clan MG, Family M24 (MetAP2)
Tb927.7.2300	TbNup132 Outer ring
Tb927.5.2310	Phosphorylated CTD interacting factor 1 WW domain containing protein, putative
Tb927.11.11740	membrane-bound acid phosphatase, putative
Tb927.6.4990	ATP synthase, epsilon chain, putative
Tb927.7.6800	Mitchondrial LSU ribosomal protein
Tb927.11.11140	hypothetical protein
Tb927.8.7570	tRNA-dihydrouridine synthase 1, putative
Tb927.11.2530	Mitochondrial SSU ribosomal protein, putative
Tb927.5.2060	cell division control protein CDC5f, part of PRP19 complex
Tb927.11.9130	replication factor A, 51kDa subunit, putative
Tb927.9.12730	chaperone protein DNAj, endoplasmatic reticulum
Tb927.11.10860	Nin one binding (NOB1) SSU pre-rRNA processing
Tb927.11.11850	splicing factor 3B subunit 1, SF3b(SAP)155
Tb927.11.6460	hypothetical protein, conserved
Tb927.7.4440	NAD dependent epimerase/dehydratase family, putative
Tb927.6.3630	Sphingosine-1-phosphate lyase
Tb927.11.14210	conserved protein
Tb927.11.5150	hypothetical protein, conserved, no homologues
Tb927.6.2920	hypothetical protein, conserved
Tb927.8.5090	DNA-directed RNA polymerase I largest subunit (RPA190)
Tb927.11.240	hypothetical protein, conserved
Tb927.6.840	hypothetical protein, conserved
Tb927.9.10370	TAX-1, flagellar protein
Tb927.7.7210	POMP37
Tb927.10.12980	methyltransferase, NOL1/NOP2/sun family
Tb927.11.14570	LicD family, putative
Tb927.7.1220	hypothetical protein, conserved
Tb927.6.3670	paraflagellar rod protein PFC8
Tb927.8.3250	IAD-1beta inner arm dynein heavy chain
Tb927.9.1750	Fibronectin type III domain containing protein, putative
Tb927.3.3360	acyltransferase, putative
Tb927.11.14450	hypothetical protein, conserved
Tb927.9.12770	TPR repeat protein
Tb927.3.5070	hypothetical protein, conserved
Tb927.10.13510	zinc metallopeptidase, putative
Tb927.10.13740	calcium-dependent lipid binding protein, putative,synaptotagmin, putative (CaLB)

Tb927.4.3400	hypothetical protein, conserved, ARM repeat
Tb927.11.10640	DNA-(apurinic or apyrimidinic site) lyase, putative
Tb927.8.5250	coatomer delta subunit, ,delta-COP
Tb927.11.7040	pterin-4-alpha-carbinolamine dehydratase, putative
Tb927.8.3550	MAPK3 protein kinase
Tb927.7.5820	Monooxygenase, putative
Tb927.10.6230	hypothetical protein, conserved,MORN repeat containing protein
Tb927.9.12780	Leucine carboxyl methyltransferase/Cupin-like domain containing protein, putative
Tb927.8.5140	DAG kinase
Tb927.9.10680	RNA polymerase III RPC4, putative
Tb927.8.1050	hypothetical protein, conserved, WD40 repeat
Tb927.3.1670	hypothetical protein, conserved
Tb927.4.870	dynein heavy chain DNAH10 IAD-1alpha inner arm dynein heavy chain
Tb927.5.3970	adenylate kinase, putative (ADKE)
Tb927.3.5180	cofilin/actin depolymerizing factor, putative
Tb927.10.11210	hypothetical protein, conserved
Tb927.11.16660	Dopey, N-terminal, putative
Tb927.6.2420	p22 protein precursor, interacts with UBP1 in T. cruzi
Tb927.11.3790	hypothetical protein, conserved
Tb927.8.6330	hypothetical protein, conserved, similar to yeast Doa1p, 7WD repeats,PFU domain binding to Ubiquitin
Tb927.11.7930	RNA polymerase B subunit RPB8, putative (RPB8)
Tb927.11.8310	class I transcription factor A, subunit 4 (CITFA-4)
Tb927.11.10130	hypothetical protein, conserved
Tb927.7.3050	Mitchondrial SSU ribosomal associated
Tb927.8.1140	hypothetical protein, conserved
Tb927.11.5270	ubiquitin carboxyl-terminal hydrolase, Clan CA cysteine peptidase, family C12, putative
Tb927.11.11730	MRP protein, putative
Tb927.10.2780	DNA-directed RNA polymerase iii largest subunit,t (RPC160)
Tb927.11.2490	Kinesin-3
Tb927.5.1140	Ethanolamine kinase 1
Tb927.5.3600	ATP-dependent DEAD/H RNA helicase, putative Prp5, splicing?
Tb927.10.4930	protein phosphatase 2C, putative
Tb927.4.560	Cytoplasmic dynein 2 heavy chain (DYNC2H1), putative
Tb927.8.1780	protein kinase, putative
Tb927.10.230	proteasome alpha 5 subunit, putative,20S proteasome subunit alpha 5
Tb927.11.2730	UDP-galactose 4-epimerase (galE)
Tb927.10.3670	hypothetical protein, conserved
Tb927.1.1720	cyclophilin 15, putative
Tb927.11.7530	long-chain-fatty-acid-CoA ligase, putative
Tb927.10.11220	procyclic form surface glycoprotein (PSSA-2)
Tb927.7.3910	NADH-ubiquinone oxidoreductase complex I subunit, putative
Tb927.6.2020	FG-GAP repeat protein, putative,intergrin alpha chain protein, putative
Tb927.4.4210	ATP-dependent zinc metallopeptidase, putative,metallo-peptidase, Clan MA(E) Family M41

		Tb927.1.3000	amidohydrolase, putative
		Tb927.11.8160	dynein heavy chain
		Tb927.11.1430	CMF2, T. brucei components of motile flagella (TbCMF2)
		Tb927.4.860	hypothetical protein, conserved
		Tb927.10.14860	Complex 1 protein (LYR family), putative
		Tb927.10.7120	hypothetical protein, conserved
		Tb927.2.5050	protein phosphatase 2C, putative
		Tb927.2.6050	coatamer beta-prime subunit beta prime COP
		Tb927.11.8090	protein phosphatase 1, putative
		Tb927.10.12590	hypothetical protein, conserved
		Tb927.4.1600	AAA domain containing protein, putative
		Tb927.7.6670	hypothetical protein, conserved
		Tb927.8.5280	Mitchondrial SSU ribosomal associated TbMRPS34
		Tb927.7.2720	eIF3c
		Tb927.6.1270	PITH domain-containing protein
		Tb927.11.9350	phosphonopyruvate decarboxylase-like protein, putative
		Tb927.2.4400	Mitchondrial SSU ribosomal protein
		Tb927.9.6040	E1-like ubiquitin-activating enzyme, putative
		Tb927.7.3290	WD domain, G-beta repeat, putative
		Tb927.11.260	PIFTF6 ,predicted WD40 repeat
		Tb927.11.15430	U5 small nuclear ribonucleoprotein U5-116K, putative
		Tb927.4.4490	multidrug resistance protein E,p-glycoprotein (MRPE)
		Tb927.7.6450	tRNA-dihydrouridine synthase 3, putative
RNA metabolism	11	Tb927.8.4200	hypothetical protein, conserved, no domains, no yeast.human match
		Tb927.11.15800	CIF1
		Tb927.6.1010	cysteine peptidase precursor, Clan CA, family C1, Cathepsin L-like (CP)
		Tb927.9.9520	ZC3H29
		Tb927.11.16280	60S ribosomal protein L2 L8
		Tb927.5.4620	ESAG9, putative
		Tb927.10.14900	hypothetical protein, conserved
		Tb927.11.1460	phosphatidylinositol (3,5) kinase, putative
		Tb927.6.3220	hypothetical protein, conserved, no domains, no yeast or human match
		Tb927.3.580	leucine-rich repeat protein (LRRP)
		Tb927.1.1890	Replication Factor C Subunit 1-related protein
stress	9	Tb927.2.2450	ribosomal RNA methyltransferase, putative (25N14.40)
		Tb927.7.4560	histone acetyltransferase HAT1
		Tb927.5.530	hypothetical protein, conserved
		Tb927.8.6840	mismatch repair protein MLH1
		Tb927.9.1340	TbNup92 Basket
		Tb927.2.2860	hypothetical protein, conserved
		Tb927.5.1640	hypothetical protein, conserved
		Tb927.3.4180	hypothetical protein
		Tb927.3.5450	BBP96

B. Yeast-two-hybrid screen-Genes common in screens

Names	total	Gene Ids	Category
CAF40, NOT10 and NOT2	6	Tb927.5.1890	hypothetical protein, conserved
		Tb927.7.7210	POMP37
		Tb927.8.3540	hypothetical protein, conserved
		Tb927.8.8120	hypothetical protein, conserved
		Tb927.5.3960	arginine N-methyltransferase, type I (PRMT6)
		Tb927.6.2540	DREV methyltransferase, putative, Zinc finger domains
CAF40, NOT11 and NOT2	2	Tb927.7.5590	Domain of unknown function (DUF3883), putative
		Tb927.7.6670	hypothetical protein, conserved
CAF40 NOT10 NOT11	1	Tb927.5.2440	CEP164
NOT10 and NOT2	2	Tb927.11.6160	ERF3 eukaryotic release factor 3, putative
		Tb927.11.2490	Kinesin-3
NOT11 and NOT2	2	Tb927.10.13360	EF-Tu Mitochondrial elongation factor Tu
		Tb927.2.2230	hypothetical protein, conserved, no domains, Kineto-specific
CAF40 and NOT2	16	Tb927.10.6000	hypothetical protein, conserved
		Tb927.8.7960	hypothetical protein
		Tb927.8.940	hypothetical protein, conserved, no clear yeast or human matches
		Tb927.3.1670	hypothetical protein, conserved
		Tb927.10.9050	pseudouridine synthase PUS7
		Tb927.6.890	hypothetical protein, conserved
		Tb927.5.650	GRESAG 4, putative receptor-type adenylate cyclase
		Tb927.4.2440	hypothetical protein, very weak match to yeast ribosomal protein RSM23
		Tb927.5.3490	hypothetical protein, conserved, LRR-RI like domain
		Tb927.2.2260	ATM kinase
		Tb927.10.11170	hypothetical protein, conserved, P-loop, AAAdomain, NF-Zinc finger with Ring
		Tb927.4.4510	protein phosphatase 2C, putative
		Tb927.10.6930	hypothetical protein. cullin, cullin-Nedd88 bs site like
		Tb927.10.10310	hypothetical protein, conserved. metallo protease like
		Tb927.9.3670	Endonuclease/Exonuclease/phosphatase family, putative
		Tb927.3.4180	hypothetical protein
NOT10 and NOT11	27	Tb927.11.14970	hypothetical protein, conserved
		Tb927.9.9010	SAC3/GANP/Nin1/mts3/eIF-3 p25 family, putative
		Tb927.6.860	hypothetical protein
		Tb927.11.11850	splicing factor 3B subunit 1, SF3b(SAP)155
		Tb927.11.2710	cell division cycle 45 (CDC45), putative
		Tb927.1.3290	hypothetical protein, conserved
		Tb927.7.5710	hypothetical protein, conserved
		Tb927.7.1100	hypothetical protein, conserved
		Tb927.6.3670	paraflagellar rod protein PFC8
		Tb927.8.4320	hypothetical protein, conserved

Tb927.11.7040	pterin-4-alpha-carbinolamine dehydratase, putative
Tb927.11.760	protein phosphatase 2C, putative
Tb927.11.15130	chaperone protein DNAj, putative
Tb927.9.7690	hypothetical protein, conserved, nucelotidyltransferase and poly(A) polymerase substrate binding domains
Tb927.8.1960	NOT11
Tb927.6.5010	hypothetical protein, conserved
Tb927.6.730	hypothetical protein, conserved
Tb927.11.9350	phosphonopyruvate decarboxylase-like protein, putative
Tb927.3.5290	hypothetical protein, conserved
Tb927.5.4290	hypothetical protein, conserved
Tb927.7.2680	ZC3H22
Tb927.6.4990	ATP synthase, epsilon chain, putative
Tb927.11.14210	conserved protein
Tb927.11.4640	calpain-like protein, putative
Tb927.8.5250	coatomer delta subunit, ,delta-COP
Tb927.4.4060	hypothetical protein, conserved
Tb927.8.4880	DNA polymerase alpha catalytic subunit,DNA polymerase I

Appendix II: List of protein interactors of ZC3H30 (Tandem Affinity Purification)

Gene ID	Description	Number of Peptides (-Rnase)	Number of Peptides (+ Rnase)	Percentage coverage (1,2,3)
Tb927.3.5610	ribosomal protein L3 mitochondrial, putative	26	NA	46.1,19.2, 5.92
Tb927.4.2790	hypothetical protein, conserved	13	NA	13.4,13, 2.99
Tb927.8.2760	Mitchondrial LSU ribosomal protein	23	NA	30.8,17.8, 4.4
Tb927.8.3300	Mitchondrial LSU ribosomal protein	35	NA	32.6, 22.9, 3.47
Tb927.8.3820	hypothetical protein, conserved, kinetoplastid-specific	18	9	51.6, 50.1, 24.8 (-Rnase); 50.1, 18 (+Rnase)
Tb927.9.5900	glutamate dehydrogenase (GDH)	48	12	60.3, 3.53, 6.15 (-Rnase); 5.65, 11.4 (+Rnase)
Tb927.9.8880	actin A	NA	5	58.5, 0, 0 (-Rnase);12.5, 13.8 (+Rnase)
Tb927.9.15150	60S ribosomal protein L5, putative	NA	6	25, 0, 0 (-Rnase);8.12, 14.3 (+Rnase)
Tb927.10.1540	ZC3H30	60	26	66.7, 51.2, 19 (-Rnase); 46.5, 19.7 (+Rnase)
Tb927.10.5330	40S ribosomal protein, S18	21	7	51.6, 34.6, 28.8 (-Rnase); 25.5, 20.9 (+Rnase)
Tb927.11.3600	40S ribosomal protein S4,putative	NA	7	56.4, 26, 3.66 (-Rnase); 14.3, 11 (+Rnase)
Tb927.11.6280	pyruvate phosphate dikinase (PPDK)	55	NA	65.6, 4.71, 5.04 (-Rnase); 1.2, 0 (+Rnase)
Tb927.11.10790	40S ribosomal protein, SA, putative	NA	4	43.4, 8.2, 3.28 (-Rnase); 18, 8.61 (+Rnase)

References

1. Maudlin, I., *African trypanosomiasis.* Annals of Tropical Medicine & Parasitology, 2006. **100**(8): p. 679-701.
2. Simarro, P.P., et al., *The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases.* Int J Health Geogr, 2010. **9**: p. 57.
3. Mair, G., et al., *A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA.* RNA, 2000. **6**(2): p. 163-9.
4. Garcia, A., et al., *Follow-up of Card Agglutination Trypanosomiasis Test (CATT) positive but apparently aparasitaemic individuals in Côte d'Ivoire: evidence for a complex and heterogeneous population.* Tropical Medicine & International Health, 2000. **5**(11): p. 786-793.
5. Kaboré, J., et al., *First evidence that parasite infecting apparent aparasitemic serological suspects in human African trypanosomiasis are Trypanosoma brucei gambiense and are similar to those found in patients.* Infection, Genetics and Evolution, 2011. **11**(6): p. 1250-1255.
6. Koffi, M., et al., *Aparasitemic serological suspects in Trypanosoma brucei gambiense human African trypanosomiasis: A potential human reservoir of parasites?* Acta Tropica, 2006. **98**(2): p. 183-188.
7. Jamonneau, V., et al., *Untreated Human Infections by Trypanosoma brucei gambiense Are Not 100% Fatal.* PLoS Neglected Tropical Diseases, 2012. **6**(6): p. e1691.
8. Peacock, L., et al., *The influence of sex and fly species on the development of trypanosomes in tsetse flies.* PLoS Negl Trop Dis, 2012. **6**(2): p. e1515.
9. Siegel, T.N., et al., *Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei.* Genes Dev, 2009. **23**(9): p. 1063-76.
10. Preußer, C., et al., *Pre-mRNA Splicing in Trypanosoma brucei: Factors, Mechanisms, and Regulation*, in *RNA Metabolism in Trypanosomes*, A. Bindereif, Editor. 2012, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 49-77.
11. Radomski, M.W., et al., *Disruptions in the secretion of cortisol, prolactin, and certain cytokines in human African trypanosomiasis patients.* Bull Soc Pathol Exot, 1994. **87**(5): p. 376-9.
12. McLean, K.J. and M. Jacobs-Lorena, *Plasmodium falciparum Maf1 Confers Survival upon Amino Acid Starvation.* Mbio, 2017. **8**(2).
13. Stijlemans, B., et al., *Immune Evasion Strategies of Trypanosoma brucei within the Mammalian Host: Progression to Pathogenicity.* Frontiers in Immunology, 2016. **7**: p. 233.
14. Rudenko, G., *Epigenetics and transcriptional control in African trypanosomes.* Essays Biochem, 2010. **48**(1): p. 201-19.
15. Hertz-Fowler, C., et al., *Telomeric expression sites are highly conserved in Trypanosoma brucei.* PLoS One, 2008. **3**(10): p. e3527.
16. Bailey, J.W. and D.H. Smith, *The quantitative buffy coat for the diagnosis of trypanosomes.* Trop Doct, 1994. **24**(2): p. 54-6.
17. Lumsden, W.H.R., D.A. Evans, and C.D. Kimber, *Miniature anion-exchange/centrifugation technique for the diagnosis of microfilaraemia in*

the field. Transactions of The Royal Society of Tropical Medicine and Hygiene, 1980. **74**(1): p. 40-42.
18. Lumsden, W.H.R., et al., *Field diagnosis of sleeping sickness in the Ivory Coast. I. Comparison of the miniature anion-exchange/centrifugation technique with other protozoological methods.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 1981. **75**(2): p. 242-250.
19. Mugo, E., F. Egler, and C. Clayton, *Conversion of procyclic-form Trypanosoma brucei to the bloodstream form by transient expression of RBP10.* Mol Biochem Parasitol, 2017. **216**: p. 49-51.
20. Zhang, J., W. Ruyechan, and N. Williams, *Developmental regulation of two nuclear RNA binding proteins, p34 and p37, from Trypanosoma brucei.* Mol Biochem Parasitol, 1998. **92**(1): p. 79-88.
21. Ruan, J.P., et al., *Functional characterization of a Trypanosoma brucei TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes.* Mol Cell Biol, 2004. **24**(21): p. 9610-8.
22. Stevens, J.R., et al., *The taxonomic position and evolutionary relationships of Trypanosoma rangeli.* Int J Parasitol., 1999. **29**.
23. Jensen, B.C., et al., *Extensive stage-regulation of translation revealed by ribosome profiling of Trypanosoma brucei.* BMC Genomics, 2014. **15**: p. 911.
24. Antwi, E.B., et al., *Integrative analysis of the Trypanosoma brucei gene expression cascade predicts differential regulation of mRNA processing and unusual control of ribosomal protein expression.* BMC Genomics, 2016. **17**: p. 306.
25. Djikeng, A., et al., *Characterization of a candidate Trypanosoma brucei U1 small nuclear RNA gene.* Mol Biochem Parasitol, 2001. **113**(1): p. 109-15.
26. Uilenberg, G., *A Field Guide for the Diagnosis, Treatment and Prevention of African Animal Trypanosomosis.* 1998: p. 158.
27. Inacio, P., et al., *Parasite-induced ER stress response in hepatocytes facilitates Plasmodium liver stage infection.* EMBO Rep, 2015. **16**(8): p. 955-64.
28. Bruhn, D.F., M.P. Sammartino, and M.M. Klingbeil, *Three Mitochondrial DNA Polymerases Are Essential for Kinetoplast DNA Replication and Survival of Bloodstream Form Trypanosoma brucei.* Eukaryotic Cell, 2011. **10**(6): p. 734-743.
29. Steinmann, M.E., et al., *TbIRK is a signature sequence free potassium channel from Trypanosoma brucei locating to acidocalcisomes.* Scientific Reports, 2017. **7**: p. 656.
30. Freire, E.R., et al., *Trypanosoma brucei translation initiation factor homolog EIF4E6 forms a tripartite cytosolic complex with EIF4G5 and a capping enzyme homolog.* Eukaryot Cell, 2014. **13**(7): p. 896-908.
31. Kramer, S., *The ApaH-like phosphatase TbALPH1 is the major mRNA decapping enzyme of trypanosomes.* Plos Pathogens, 2017. **13**(6).
32. Capy, P., et al., *Stress and transposable elements: co-evolution or useful parasites?* Heredity (Edinb), 2000. **85 (Pt 2)**: p. 101-6.
33. Berriman, M., et al., *The genome of the African trypanosome Trypanosoma brucei.* Science, 2005. **309**(5733): p. 416-22.
34. Maslov, D.A., et al., *Phylogeny of trypanosomes as inferred from the small and large subunit rRNAs: implications for the evolution of parasitism in the*

trypanosomatid protozoa. Molecular and Biochemical Parasitology, 1996. **75**(2): p. 197-205.
35. Simpson, A.G.B., J. Lukeš, and A.J. Roger, *The Evolutionary History of Kinetoplastids and Their Kinetoplasts.* Molecular Biology and Evolution, 2000. **19**(12): p. 2071-2083.
36. Lin, R.-H., et al., *Analysis of the mitochondrial maxicircle of Trypanosoma lewisi, a neglected human pathogen.* Parasites & Vectors, 2015. **8**(1): p. 665.
37. Akiyoshi, B. and K. Gull, *Evolutionary cell biology of chromosome segregation: insights from trypanosomes.* Open Biology, 2013. **3**(5): p. 130023.
38. Ponte-Sucre, A., *An Overview of Trypanosoma brucei Infections: An Intense Host–Parasite Interaction.* Frontiers in Microbiology, 2016. **7**: p. 2126.
39. Jackson, D.G., M.J. Owen, and H.P. Voorheis, *A new method for the rapid purification of both the membrane-bound and released forms of the variant surface glycoprotein from Trypanosoma brucei.* Biochemical Journal, 1985. **230**(1): p. 195-202.
40. Hutchinson, O.C., et al., *VSG structure: similar N-terminal domains can form functional VSGs with different types of C-terminal domain.* Molecular and Biochemical Parasitology, 2003. **130**(2): p. 127-131.
41. Marcello, L. and J.D. Barry, *Analysis of the VSG gene silent archive in Trypanosoma brucei reveals that mosaic gene expression is prominent in antigenic variation and is favoured by archive substructure.* Genome research, 2007. **17**(9): p. 1344-1352.
42. Horn, D., *Antigenic variation in African trypanosomes.* Molecular and Biochemical Parasitology, 2014. **195**(2): p. 123-129.
43. Matthews, K.R., R. McCulloch, and L.J. Morrison, *The within-host dynamics of African trypanosome infections.* Philosophical Transactions of the Royal Society B: Biological Sciences, 2015. **370**(1675): p. 20140288.
44. Rudenko, G., *Mechanisms mediating antigenic variation in Trypanosoma brucei.* Mem Inst Oswaldo Cruz, 1999. **94**(2): p. 235-7.
45. Krstin, S., H.S. Peixoto, and M. Wink, *Combinations of Alkaloids Affecting Different Molecular Targets with the Saponin Digitonin Can Synergistically Enhance Trypanocidal Activity against Trypanosoma brucei brucei.* Antimicrobial Agents and Chemotherapy, 2015. **59**(11): p. 7011-7017.
46. Pavio, N., D.R. Taylor, and M.M.C. Lai, *Detection of a novel unglycosylated form of hepatitis C virus E2 envelope protein that is located in the cytosol and interacts with PKR.* Journal of Virology, 2002. **76**(3): p. 1265-1272.
47. Christoffels, A., et al., *International Glossina Genome Initiative 2004–2014: A Driver for Post-Genomic Era Research on the African Continent.* PLoS Neglected Tropical Diseases, 2014. **8**(8): p. e3024.
48. Sharma, S., et al., *Acetylation-Dependent Control of Global Poly(A) RNA Degradation by CBP/p300 and HDAC1/2.* Mol Cell, 2016. **63**(6): p. 927-38.
49. Roditi, I. and C. Clayton, *An unambiguous nomenclature for the major surface glycoproteins of the procyclic form of Trypanosoma brucei.* Molecular and Biochemical Parasitology, 1999. **103**(1): p. 99-100.
50. Vassella, E., et al., *Expression of a Major Surface Protein of Trypanosoma brucei Insect Forms Is Controlled by the Activity of Mitochondrial Enzymes.* Molecular Biology of the Cell, 2004. **15**(9): p. 3986-3993.

51. Ruepp, S., et al., *Survival of Trypanosoma brucei in the Tsetse Fly Is Enhanced by the Expression of Specific Forms of Procyclin.* The Journal of Cell Biology, 1997. **137**(6): p. 1369-1379.
52. Vassella, E., et al., *A major surface glycoprotein of Trypanosoma brucei is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia.* Genes & Development, 2000. **14**(5): p. 615-626.
53. Graham, S.V. and J.D. Barry, *Transcriptional regulation of metacyclic variant surface glycoprotein gene expression during the life cycle of Trypanosoma brucei.* Molecular and Cellular Biology, 1995. **15**(11): p. 5945-5956.
54. Peel, E., *Identification of metacyclic trypanosomes in the hypopharynx of tsetse flies, infected in nature or in the laboratory.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 1962. **56**(4): p. 339-341.
55. Langousis, G. and K.L. Hill, *Motility and more: the flagellum of Trypanosoma brucei.* Nat Rev Micro, 2014. **12**(7): p. 505-518.
56. Smith, T.K., et al., *Metabolic reprogramming during the Trypanosoma brucei life cycle.* F1000Res, 2017. **6**.
57. Smith, T.K., et al.
58. Gualdrón-López, M., et al., *Ubiquitination of the glycosomal matrix protein receptor PEX5 in Trypanosoma brucei by PEX4 displays novel features.* Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 2013. **1833**(12): p. 3076-3092.
59. Colasante, C., et al., *Comparative proteomics of glycosomes from bloodstream form and procyclic culture form Trypanosoma brucei brucei.* PROTEOMICS, 2006. **6**(11): p. 3275-3293.
60. Chaudhuri, M., R.D. Ott, and G.C. Hill, *Trypanosome alternative oxidase: from molecule to function.* Trends in Parasitology, 2006. **22**(10): p. 484-491.
61. Brown, S.V., et al., *ATP Synthase Is Responsible for Maintaining Mitochondrial Membrane Potential in Bloodstream Form Trypanosoma brucei.* Eukaryotic Cell, 2006. **5**(1): p. 45-53.
62. Guerra-Giraidez, C., L. Quijada, and C.E. Clayton, *Compartmentation of enzymes in a microbody, the glycosome, is essential in Trypanosoma brucei.* Journal of Cell Science, 2002. **115**(13): p. 2651-2658.
63. Nolan, D.P. and H.P. Voorheis, *Factors that determine the plasma-membrane potential in bloodstream forms of Trypanosoma brucei.* European Journal of Biochemistry, 2000. **267**(15): p. 4615-4623.
64. Priest, J.W. and S.L. Hajduk, *Developmental regulation of Trypanosoma brucei cytochrome c reductase during bloodstream to procyclic differentiation.* Mol Biochem Parasitol, 1994. **65**(2): p. 291-304.
65. Haanstra, J.R., et al., *A domino effect in drug action: from metabolic assault towards parasite differentiation.* Molecular Microbiology, 2011. **79**(1): p. 94-108.
66. Mantilla, B.S., et al., *Proline Metabolism is Essential for Trypanosoma brucei brucei Survival in the Tsetse Vector.* Plos Pathogens, 2017. **13**(1).
67. Millerioux, Y., et al., *The threonine degradation pathway of the Trypanosoma brucei procyclic form: the main carbon source for lipid*

biosynthesis is under metabolic control. Molecular Microbiology, 2013. **90**(1): p. 114-129.
68. Deramchia, K., et al., *Contribution of Pyruvate Phosphate Dikinase in the Maintenance of the Glycosomal ATP/ADP Balance in the Trypanosoma brucei Procyclic Form.* Journal of Biological Chemistry, 2014. **289**(25): p. 17365-17378.
69. Daniels, J.-P., K. Gull, and B. Wickstead, *Cell Biology of the Trypanosome Genome.* Microbiology and Molecular Biology Reviews : MMBR, 2010. **74**(4): p. 552-569.
70. Ersfeld, K., S.E. Melville, and K. Gull, *Nuclear and Genome Organization of Trypanosoma brucei.* Parasitology Today. **15**(2): p. 58-63.
71. Ogbadoyi, E., et al., *Architecture of the Trypanosoma brucei nucleus during interphase and mitosis.* Chromosoma, 2000. **108**(8): p. 501-513.
72. Das, A., M. Banday, and V. Bellofatto, *RNA Polymerase Transcription Machinery in Trypanosomes.* Eukaryotic Cell, 2008. **7**(3): p. 429-434.
73. Hertz-Fowler, C., et al., *Telomeric Expression Sites Are Highly Conserved in Trypanosoma brucei.* PLOS ONE, 2008. **3**(10): p. e3527.
74. Wedel, C., et al., *GT-rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes.* The EMBO Journal: p. e201695323-n/a.
75. Ullu, E. and C. Tschudi, *Trans splicing in trypanosomes requires methylation of the 5' end of the spliced leader RNA.* Proc Natl Acad Sci U S A, 1991. **88**(22): p. 10074-8.
76. Dostalova, A., et al., *The nuclear mRNA export receptor Mex67-Mtr2 of Trypanosoma brucei contains a unique and essential zinc finger motif.* Molecular Microbiology, 2013. **88**(4): p. 728-739.
77. Inoue, A.H., et al., *Identification of a Novel Nucleocytoplasmic Shuttling RNA Helicase of Trypanosomes.* PLOS ONE, 2014. **9**(10): p. e109521.
78. Freire, E.R., et al., *The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties.* Molecular and Biochemical Parasitology, 2011. **176**(1): p. 25-36.
79. Freire, E.R., et al., *Trypanosoma brucei Translation Initiation Factor Homolog EIF4E6 Forms a Tripartite Cytosolic Complex with EIF4G5 and a Capping Enzyme Homolog.* Eukaryotic Cell, 2014. **13**(7): p. 896-908.
80. Antwi, E.B., et al., *Integrative analysis of the Trypanosoma brucei gene expression cascade predicts differential regulation of mRNA processing and unusual control of ribosomal protein expression.* BMC Genomics, 2016. **17**: p. 306.
81. Fadda, A., et al., *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, 2014. **94**(2): p. 307-326.
82. Collart, M.A., *Global control of gene expression in yeast by the Ccr4-Not complex.* Gene, 2003. **313**: p. 1-16.
83. Preissler, S., et al., *Not4-dependent translational repression is important for cellular protein homeostasis in yeast.* Embo Journal, 2015. **34**(14): p. 1905-1924.

84. Benson, J.D., et al., *Association of distinct yeast Not2 functional domains with components of Gcn5 histone acetylase and Ccr4 transcriptional regulatory complexes.* Embo Journal, 1998. **17**(22): p. 6714-6722.
85. Bawankar, P., et al., *NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain.* RNA Biol, 2013. **10**(2): p. 228-44.
86. Sandler, H., et al., *Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin.* Nucleic Acids Research, 2011. **39**(10): p. 4373-4386.
87. Menna-Barreto, R.F.S. and S.L. de Castro, *The Double-Edged Sword in Pathogenic Trypanosomatids: The Pivotal Role of Mitochondria in Oxidative Stress and Bioenergetics.* BioMed Research International, 2014. **2014**: p. 614014.
88. Sharma, S., et al., *Acetylation-Dependent Control of Global Poly(A) RNA Degradation by CBP/p300 and HDAC1/2.* Molecular Cell, 2016. **63**(6): p. 927-938.
89. Schwede, A., et al., *The role of deadenylation in the degradation of unstable mRNAs in trypanosomes.* Nucleic Acids Res, 2009. **37**(16): p. 5511-28.
90. Utter, C.J., et al., *Poly(A)-Specific Ribonuclease (PARN-1) Function in Stage-Specific mRNA Turnover in Trypanosoma brucei.* Eukaryotic Cell, 2011. **10**(9): p. 1230-1240.
91. Erben, E., C. Chakraborty, and C. Clayton, *The CAF1-NOT complex of trypanosomes.* Front Genet, 2014. **4**: p. 299.
92. Farber, V., et al., *Trypanosome CNOT10 is essential for the integrity of the NOT deadenylase complex and for degradation of many mRNAs.* Nucleic Acids Res, 2013. **41**(2): p. 1211-22.
93. Chakraborty, C., et al.
94. Lueong, S., et al., *Gene expression regulatory networks in Trypanosoma brucei: insights into the role of the mRNA-binding proteome.* Molecular Microbiology, 2016. **100**(3): p. 457-471.
95. Erben, E.D., et al., *A genome-wide tethering screen reveals novel potential post-transcriptional regulators in Trypanosoma brucei.* PLoS Pathog, 2014. **10**(6): p. e1004178.
96. Clayton, C. and A. Estevez, *The Exosomes of Trypanosomes and Other Protists*, in *RNA Exosome*, T.H. Jensen, Editor. 2010, Springer US: New York, NY. p. 39-49.
97. Estévez, A.M., T. Kempf, and C. Clayton, *The exosome of Trypanosoma brucei.* The EMBO Journal, 2001. **20**(14): p. 3831-3839.
98. Archer, S., et al., *Chapter 18 Trypanosomes as a Model to Investigate mRNA Decay Pathways.* Methods in Enzymology, 2008. **448**: p. 359-377.
99. Clayton, C., et al., *Control of mRNA degradation in trypanosomes.* Biochemical Society Transactions, 2008. **36**(3): p. 520.
100. Wurst, M., et al., *Expression of the RNA recognition motif protein RBP10 promotes a bloodstream-form transcript pattern in Trypanosoma brucei.* Mol Microbiol, 2012. **83**(5): p. 1048-63.
101. Kolev, N.G., et al., *Developmental Progression to Infectivity in Trypanosoma brucei Triggered by an RNA-Binding Protein.* Science (New York, N.Y.), 2012. **338**(6112): p. 1352-1353.

102. Das, A., et al., *The essential polysome-associated RNA-binding protein RBP42 targets mRNAs involved in Trypanosoma brucei energy metabolism.* RNA, 2012. **18**(11): p. 1968-1983.
103. Droll, D., et al., *Post-transcriptional regulation of the trypanosome heat shock response by a zinc finger protein.* PLoS Pathog, 2013. **9**(4): p. e1003286.
104. Minia, I. and C. Clayton, *Regulating a Post-Transcriptional Regulator: Protein Phosphorylation, Degradation and Translational Blockage in Control of the Trypanosome Stress-Response RNA-Binding Protein ZC3H11.* Plos Pathogens, 2016. **12**(3).
105. Droll, D., et al., *Post-Transcriptional Regulation of the Trypanosome Heat Shock Response by a Zinc Finger Protein.* Plos Pathogens, 2013. **9**(4).
106. Minia, I., et al., *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, 2016. **10**(9).
107. Singh, A., et al., *Trypanosome MKT1 and the RNA-binding protein ZC3H11: interactions and potential roles in post-transcriptional regulatory networks.* Nucleic Acids Research, 2014. **42**(7): p. 4652-4668.
108. Lott, K., et al., *Arginine methylation of DRBD18 differentially impacts its opposing effects on the trypanosome transcriptome.* Nucleic Acids Research, 2015. **43**(11): p. 5501-5523.
109. Fernández-Moya, S.M., et al., *Alterations in DRBD3 Ribonucleoprotein Complexes in Response to Stress in Trypanosoma brucei.* PLoS ONE, 2012. **7**(11): p. e48870.
110. Franks, T.M. and J. Lykke-Andersen, *The Control of mRNA Decapping and P-Body Formation.* Molecular cell, 2008. **32**(5): p. 605-615.
111. Hurto, R.L. and A.K. Hopper, *P-body components, Dhh1 and Pat1, are involved in tRNA nuclear-cytoplasmic dynamics.* RNA, 2011. **17**(5): p. 912-924.
112. Fritz, M., et al., *Novel insights into RNP granules by employing the trypanosome's microtubule skeleton as a molecular sieve.* Nucleic Acids Research, 2015. **43**(16): p. 8013-8032.
113. Goldshmidt, H., et al., *Persistent ER Stress Induces the Spliced Leader RNA Silencing Pathway (SLS), Leading to Programmed Cell Death in Trypanosoma brucei.* Plos Pathogens, 2010. **6**(1).
114. Goldshmidt, H. and S. Michaeli, *Induction of ER Stress Response Leading to Programmed Cell Death in Trypanosoma brucei.* Methods in Enzymology, 2011. **489**: p. 189-205.
115. Krauth-Siegel, R.L., S.K. Meiering, and H. Schmidt, *The Parasite-Specific Trypanothione Metabolism of Trypanosoma and Leishmania*, in *Biological Chemistry*2003. p. 539.
116. Melchers, J., et al., *Glutathionylation of trypanosomal thiol redox proteins.* J Biol Chem, 2007. **282**(12): p. 8678-94.
117. Ulrich, K., et al., *Stress-Induced Protein S-Glutathionylation and S-Trypanothionylation in African Trypanosomes-A Quantitative Redox Proteome and Thiol Analysis.* Antioxid Redox Signal, 2017.

118. Erben, E.D., et al., *A Genome-Wide Tethering Screen Reveals Novel Potential Post-Transcriptional Regulators in Trypanosoma brucei.* PLOS Pathogens, 2014. **10**(6): p. e1004178.
119. Singh, A., et al., *Trypanosome MKT1 and the RNA-binding protein ZC3H11: interactions and potential roles in post-transcriptional regulatory networks.* Nucleic Acids Res, 2014. **42**(7): p. 4652-68.
120. Alibu, V.P., et al., *A doubly inducible system for RNA interference and rapid RNAi plasmid construction in Trypanosoma brucei.* Molecular and Biochemical Parasitology, 2005. **139**(1): p. 75-82.
121. Cannata, J.J.B., et al., *Subcellular localization of phosphoenolpyruvate carboxykinase in the trypanosomatids Trypanosoma cruzi and Crithidia fasciculata.* Molecular and Biochemical Parasitology, 1982. **6**(3): p. 151-160.
122. Rigaut, G., et al., *A generic protein purification method for protein complex characterization and proteome exploration.* Nat Biotech, 1999. **17**(10): p. 1030-1032.
123. Kibona, S.N., et al., *Drug-resistance of Trypanosoma b. rhodesiense isolates from Tanzania*
Résistance aux médicaments de Trypanosoma b. rhodesiense isolé en Tanzanie
Resistencia a medicamentos en aislados de Trypanosoma b. rhodesiense provenientes de Tanzania. Tropical Medicine & International Health, 2006. **11**(2): p. 144-155.
124. Mishina, Y.V., et al., *Artemisinins Inhibit Trypanosoma cruzi and Trypanosoma brucei rhodesiense In Vitro Growth.* Antimicrobial Agents and Chemotherapy, 2007. **51**(5): p. 1852-1854.
125. Dejung, M., et al., *Quantitative Proteomics Uncovers Novel Factors Involved in Developmental Differentiation of Trypanosoma brucei.* PLoS Pathog, 2016. **12**(2): p. e1005439.
126. Klein, C., et al., *Polysomes of Trypanosoma brucei: Association with Initiation Factors and RNA-Binding Proteins.* PLoS One, 2015. **10**(8): p. e0135973.
127. Dean, S., J.D. Sunter, and R.J. Wheeler, *TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource.* Trends in Parasitology, 2017. **33**(2): p. 80-82.
128. Berg, J.M. and Y. Shi, *The Galvanization of Biology: A Growing Appreciation for the Roles of Zinc.* Science, 1996. **271**(5252): p. 1081.
129. Wang, D., et al., *Genome-wide analysis of CCCH zinc finger family in Arabidopsis and rice.* BMC Genomics, 2008. **9**: p. 44-44.
130. Coletta, A., et al., *Low-complexity regions within protein sequences have position-dependent roles.* BMC Systems Biology, 2010. **4**: p. 43-43.
131. Schultz, J., et al., *SMART, a simple modular architecture research tool: Identification of signaling domains.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(11): p. 5857-5864.
132. Baron-Benhamou, J., et al., *Using the λN Peptide to Tether Proteins to RNAs*, in *mRNA Processing and Metabolism: Methods and Protocols*, D.R. Schoenberg, Editor. 2004, Humana Press: Totowa, NJ. p. 135-153.
133. Conway, C., et al., *Two pathways of homologous recombination in Trypanosoma brucei.* Molecular Microbiology, 2002. **45**(6): p. 1687-1700.

134. Estévez, A.M., et al., *Knockout of the glutamate dehydrogenase gene in bloodstream Trypanosoma brucei in culture has no effect on editing of mitochondrial mRNAs.* Molecular and Biochemical Parasitology, 1999. **100**(1): p. 5-17.
135. Bondy, S.C., *Ethanol toxicity and oxidative stress.* Toxicology Letters, 1992. **63**(3): p. 231-241.
136. Comporti, M., et al., *Ethanol-induced oxidative stress: basic knowledge.* Genes & Nutrition, 2010. **5**(2): p. 101-109.
137. Li, F.J., et al., *A role of autophagy in Trypanosoma brucei cell death.* Cellular Microbiology, 2012. **14**(8): p. 1242-1256.
138. Kramer, S., et al., *Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2(alpha) phosphorylation at Thr169.* J Cell Sci, 2008. **121**(Pt 18): p. 3002-14.
139. Shannon, P., et al., *Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks.* Genome Research, 2003. **13**(11): p. 2498-2504.
140. Brangwynne, C.P., *Phase transitions and size scaling of membrane-less organelles.* Journal of Cell Biology, 2013. **203**(6): p. 875-881.
141. Buchan, J.R. and R. Parker, *Eukaryotic Stress Granules: The Ins and Outs of Translation.* Molecular Cell, 2009. **36**(6): p. 932-941.
142. Nott, T.J., et al., *Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles.* Molecular Cell, 2015. **57**(5): p. 936-947.
143. Stoecklin, G., et al., *MK2-induced tristetraprolin:14-3-3 complexes prevent stress granule association and ARE-mRNA decay.* The EMBO Journal, 2004. **23**(6): p. 1313-1324.
144. Tompa, P., *The interplay between structure and function in intrinsically unstructured proteins.* Febs Letters, 2005. **579**(15): p. 3346-3354.
145. Fields, S. and O.-k. Song, *A novel genetic system to detect protein protein interactions.* Nature, 1989. **340**(6230): p. 245-246.
146. Finley, R.L. and R. Brent, *Interaction mating reveals binary and ternary connections between Drosophila cell cycle regulators.* Proceedings of the National Academy of Sciences of the United States of America, 1994. **91**(26): p. 12980-12984.
147. Bartel, P.L., et al., *A protein linkage map of Escherichia coli bacteriophage T7.* Nat Genet, 1996. **12**(1): p. 72-77.
148. McCraith, S., et al., *Genome-wide analysis of vaccinia virus protein–protein interactions.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(9): p. 4879-4884.
149. Djinovic-Carugo, K., et al., *The spectrin repeat: a structural platform for cytoskeletal protein assemblies.* FEBS Lett, 2002. **513**(1): p. 119-23.
150. Szöőr, B., et al., *Independent Pathways Can Transduce the Life-Cycle Differentiation Signal in Trypanosoma brucei.* PLOS Pathogens, 2013. **9**(10): p. e1003689.
151. Christiano, R., et al., *The proteome and transcriptome of the infectious metacyclic form of Trypanosoma brucei define quiescent cells primed for mammalian invasion.* Molecular Microbiology: p. n/a-n/a.

152. Bidel, S., et al., *Effect of the ulcerogenic agents ethanol, acetylsalicylic acid and taurocholate on actin cytoskeleton and cell motility in cultured rat gastric mucosal cells.* World Journal of Gastroenterology : WJG, 2005. **11**(26): p. 4032-4039.
153. Stanley, D., et al., *The ethanol stress response and ethanol tolerance of Saccharomyces cerevisiae.* Journal of Applied Microbiology, 2010. **109**(1): p. 13-24.
154. Gonzalez, A., et al., *Studies on the Mode of Action of Hygromycin-B, an Inhibitor of Translocation in Eukaryotes.* Biochimica Et Biophysica Acta, 1978. **521**(2): p. 459-469.
155. Nathans, D., *Puromycin Inhibition of Protein Synthesis - Incorporation of Puromycin into Peptide Chains.* Proceedings of the National Academy of Sciences of the United States of America, 1964. **51**(4): p. 585-&.
156. Popp, M.W.-L. and L.E. Maquat, *Organizing Principles of Mammalian Nonsense-Mediated mRNA Decay.* Annual review of genetics, 2013. **47**: p. 139-165.
157. Karzai, A.W., E.D. Roche, and R.T. Sauer, *The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue.* Nat Struct Mol Biol, 2000. **7**(6): p. 449-455.
158. Venkataraman, K., et al., *Non-stop mRNA decay: a special attribute of trans-translation mediated ribosome rescue.* Frontiers in Microbiology, 2014. **5**: p. 93.
159. Withey, J.H. and D.I. Friedman, *A Salvage Pathway for Protein Synthesis: tmRNA and Trans-Translation.* Annual Review of Microbiology, 2003. **57**(1): p. 101-123.
160. Doma, M.K. and R. Parker, *Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation.* Nature, 2006. **440**(7083): p. 561-564.
161. Delhi, P., et al., *Is There a Classical Nonsense-Mediated Decay Pathway in Trypanosomes?* Plos One, 2011. **6**(9).
162. Cristodero, M., et al., *Functional characterization of the trypanosome translational repressor SCD6.* Biochem J, 2014. **457**(1): p. 57-67.
163. Kruger, T., M. Hofweber, and S. Kramer, *SCD6 induces ribonucleoprotein granule formation in trypanosomes in a translation-independent manner, regulated by its Lsm and RGG domains.* Mol Biol Cell, 2013. **24**(13): p. 2098-111.
164. Glover, J.R. and S. Lindquist, *Hsp104, Hsp70, and Hsp40: A novel chaperone system that rescues previously aggregated proteins.* Cell, 1998. **94**(1): p. 73-82.
165. Kato, M., et al., *Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels.* Cell, 2012. **149**(4): p. 753-767.
166. Riback, J.A., et al., *Stress-Triggered Phase Separation Is an Adaptive, Evolutionarily Tuned Response.* Cell, 2017. **168**(6): p. 1028-+.