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

MSc. Elisha Muchunga Mugo
Born in: Murang'a, Kenya
Oral-examination: 31.07.17

The role of RBP10, a key post-transcriptional regulator in the development of *Trypanosoma brucei*

Referees: Prof. Dr. Christine Clayton (ZMBH)

Prof. Dr. Luise Krauth-Siegel (BZH)

Acknowledgements

First I would like to extend my sincere gratitude and appreciation to my PhD supervisor Prof. Christine Clayton for her mentorship, encouragement and valuable suggestions through the course of the study. Also, many thanks to my Thesis Advisory Committee members (TAC) Prof. Luise Krauth- Seigel and Dr. Martin Kös for their advice and guidance that ensured the study objectives were met.

Special gratitude to my former colleagues Esteban, Valentin, Bhaskar, Doro, Conny, Diana, Aditi and Igor; for their immense support at the beginning of my PhD and the many fun-filled activities playing bowling, movie nights, barbecue and ensuring that I stick to the lab tradition of learning how to make Feuerzangenbowle. In addition, I would like to thank all the current lab members; Daniela, Chaitali, Kevin, Kathrin, Bin and Larissa for creating a friendly and supportive environment in the lab. Special thanks to Franzika and Kathrin for the help in writing my Zusammenfassung. To Franzika, Alex, Manuela, Sabrina, and Mariane thank you for assisting with the RBP10 story. Also, I am immensely grateful to Ute and Claudia for their kind and unconditional support.

Apart from my colleagues I would like to thank Charlotta Funaya who helped me with electron microscopy, David Ibberson for assistance with RNA-seq library preparations and deep sequencing. I am also grateful to lab of Keith Matthews, Isabel Roditi, Shula Michaeli, Paul Michels, Minu Chauduri, Peter Overath, and George Cross for the kind donation of antibody.

Lastly, I would like to thank my spouse Fiona and my family members for their unconditional love, care and support.

Table of Contents

Summary

In African trypanosomes, the control of transcription initiation by RNA pol II is absent at the level of individual mRNAs. Nevertheless, gene expression changes dramatically during life cycle transitions in response to the changing host environments. In the bloodstream of the mammalian host, *Trypanosoma brucei* exist as proliferative long slender forms or as a non-dividing stumpy forms; the latter differentiate to procyclic forms in the midgut of the tsetse fly. Differential gene expression between life cycle stages is achieved through regulation of mRNA degradation and translation, and mainly relies on RNA binding proteins. The work focuses on the RNA binding protein RBP10. RBP10 is a bloodstream form specific cytoplasmic protein with a single RRM domain. Depletion of RBP10 in bloodstream forms or forced expression in the procyclic forms is lethal due to mis-regulation of developmentally expressed mRNAs. Bloodstream forms cells depleted of RBP10 differentiate to procyclic forms after transfer into procyclic growth media and incubation at 27^{o}C; within three days, >80% of the cells express GPEET procyclin, and reposition their kinetoplast. Conversely, expression of RBP10 in procyclic cells for two days converts the cells to bloodstream forms. In such cells, eight VSG transcripts, including three with metacyclic promoters, were strongly up regulated, and ~16% of the cells had acquired a VSG surface coat. More importantly, a subset of the cells survived after transfer into bloodstream form growth media and incubation at 37^{o}C, resulting in proliferating cells in about ten days.

Tethering of RBP10 to a reporter mRNA inhibits translation and promotes mRNA degradation. RBP10 from bloodstream forms co-precipitated many procyclic specific mRNAs that are normally unstable in bloodstream forms. Indeed, 39% of the mRNAs up regulated after RBP10 depletion were bound by it; these included the transcript encoding procyclic surface coat EP procyclin, several enzymes needed for procyclic energy metabolism, regulatory proteins ZC3H21, ZC3H20, two kinases and a phosphatase. The UA(U)$_6$ motif was found to be highly enriched in the *3' UTR* of RBP10 mRNA targets. Binding of RBP10 to *EP 3' UTR* was lost when the UA(U)$_6$ motif was deleted from a reporter mRNA, and the motif was also necessary for its regulation by RBP10. In bloodstream forms RBP10 target mRNAs are likely to be blocked from translation hence degraded; this is important for survival. Perturbation of RBP10 expression therefore triggers a regulatory cascade that is sufficient to modulate *T. brucei* developmental capacity.

Zusammenfassung

Bei afrikanischen Trypanosomen gibt es keine Kontrolle der Transkriptionsinitiation durch RNA Pol II auf der Ebene einzelner mRNAs. Nichtsdestotrotz verändert sich die Genexpression während der Übergänge im Lebenszyklus als Reaktion auf die unterschiedlichen Wirts-Umgebungen dramatisch. Im Blutstrom des Säugerwirts existiert *Trypanosoma brucei* als proliferative lange Form (*slender*) oder als sich nicht teilende gedrungene Form (*stumpy*); letztere differenziert im Mitteldarm der Tsetsefliege zur prozyklischen Form. Unterschiede der Genexpression in den verschiedenen Lebenszyklusstadien werden durch die Regulierung der Degradation und Translation von mRNAs erreicht und sind vor allem auf RNA-bindende Proteine zurückzuführen. Diese Arbeit konzentriert sich auf das RNA-bindende Protein RBP10. RBP10 ist ein zytoplasmatisches Protein, das spezifisch für die Blutstrom-Form ist und eine einzelne RRM-Domäne besitzt. Die Depletion von RBP10 in der Blutstrom-Form oder die erzwungene Expression in der prozyklischen Form sind letal, da mRNAs, die entwicklungsbedingt exprimiert werden, fehlreguliert werden. Die Depletion von RBP10 in der Blutstrom-Form führt nach dem Transfer dieser Zellen in prozyklisches Medium und Inkubation bei 27°C zur Differenzierung in die prozyklische Form. Innerhalb von drei Tagen exprimieren >80% dieser Zellen das Prozyklin GPEET und re-positionieren den Kinetoplasten. Wird RBP10 dagegen zwei Tage in prozyklischen Zellen exprimiert, führt dies zur Bildung der Blutstrom-Form. In diesen Zellen sind acht VSG-Transkripte stark hoch-reguliert, darunter drei mit metazyklischen Promotoren, und ~16% der Zellen haben eine VSG-Oberflächenschicht. Darüber hinaus überlebte ein Bruchteil der Zellen den Transfer in Blutstrom-Form Wachstumsmedium und Inkubation bei 37°C. Dies führte nach etwa 10 Tagen zu teilungsaktiven Zellen.

Die Bindung von RBP10 an eine Reporter-mRNA inhibiert deren Translation und fördert deren Degradation. Die Präzipitation von RBP10 in der Blutstrom-Form hat viele mRNAs mitausgefällt, die spezifisch für die prozyklische Form sind und normalerweise in der Blutstrom-Form nicht stabil sind. Tatsächlich waren 39% der mRNAs, die nach Depletion von RBP10 hoch-reguliert waren, an RBP10 gebunden; dies umfasste das Transkript, welches das prozyklische Oberflächenprotein EP Prozyklin kodiert, sowie einige Enzyme, die für den prozyklischen Metabolismus benötigt werden, die regulatorischen Proteine ZC3H21 und ZC3H20, zwei Kinasen und eine Phosphatase. In den 3'UTRs der mRNAs, die an RBP10 binden, wurde vermehrt das Motiv $UA(U)_6$ gefunden. Die Entfernung des $UA(U)_6$ Motives in der EP 3'UTR einer Reporter-mRNA führte zu einem Verlust der Interaktion mit RPB10. Außerdem ist dieses Motiv für die Regulation der mRNA durch RBP10 nötig. In der Blutstrom-Form wird die Translation der Ziel-mRNAs von RBP10 vermutlich blockiert und diese daher abgebaut, was für das Überleben wichtig ist. Eine Störung der Expression von RBP10 löst daher eine regulatorische Kaskade aus, die ausreicht, um das Entwicklungspotential von *T. brucei* zu verändern.

1 Introduction

1.1 Background

Trypanosomes are protists that belong to the class Kinetoplastida. The class comprises of flagellated unicellular organisms that infect humans, animals and plants; many Kinetoplastids are transmitted by arthropods. Several *Trypanosoma brucei* subspecies cause a poverty-related vector-borne infectious disease in either humans (sleeping sickness) or animals (nagana in cattle) in sub-saharan Africa. Close relatives of trypanosomes include *Leishmania* spp. (causes various cutaneous and visceral leishmaniasis) and *Trypanosoma cruzi* (causes Chagas disease).

Trypanosoma brucei was first detected more than a century ago as the parasitic agent responsible for what was then called 'tsetse fly disease' or nagana in African cattle [1]. Later on trypanosomes were linked to a human disease called sleeping sickness after detection of the parasite in patients blood [2] and cerebral spinal fluids [3]. Follow-up studies revealed that *T. brucei* is cyclically transmitted between mammalian hosts by the tsetse fly [4]. Currently, Human African Trypanosomiasis (HAT) or sleeping sickness affects 36 countries in sub-Saharan Africa where the tsetse fly is found (http://www.who.int/mediacentre/factsheets/fs259/en/).

1.2 Diseases caused by *T. brucei*

Two *T. brucei* subspecies, *T. b. gambiense* and *T. b. rhodesiense,* infect both humans and animals. *T. b. gambiense* is mainly present in West and Central Africa, and usually causes a chronic form of infection; pigs have been identified as a potential reservoir but the epidemiological role for the animal reservoir is not well understood. In contrast, *T. b. rhodesiense* is mainly zoonotic and only occasionally infects humans. It is found in Eastern and Southern Africa, where it is known to cause an acute form of HAT. The *gambiense* form of HAT accounts for more than 95% of the reported cases [5]. In 2009 the reported HAT cases were <10,000 for the first time in 50 years, and in 2015 only 2804 cases were recorded (http://www.who.int/mediacentre/factsheets/fs259/en/). This is attributed to the continuous active surveillance for HAT, vector control and treatment for HAT cases [6]. Such efforts need to be well sustained in order to achieve the WHO's target to eliminate HAT as a public health burden by 2030. For *T. b. rhodesiense*, a major

challenge is the parasite reservoir in wild animals which acts as a source for new human infection.

HAT occurs in two clinical stages [6-8]. In the early/first stage, parasites are found mainly in the bloodstream and the lymphatic system, also known as the hemo-lymphatic stage. The late/second stage of the disease begins when the parasites cross the blood- brain barrier and invade the central nervous system. Early stage HAT patients develop unspecific clinical signs including intermittent fever, headache, pruritus, joint pains, weakness and weight loss [9, 10]. The onset of the late stage is characterised by the appearance of neuro-psychiatric symptoms including mental changes and sleep disturbances [6]. If left untreated at this point, the disease leads to coma, multiple organ failure and eventually death.

Unfortunately, vaccination is not an option for HAT treatment due to the antigenic variation phenomenon in trypanosomes; a process where bloodstream form trypanosomes escape host immune response by switching their variant surface glycoproteins (VSGs) that coat the outer membrane. The choice of drugs for treating HAT depends on the stage of the disease and species of the parasite. Early stage *T. b. gambiense* HAT is treated using pentamidine administered by intramuscular injections once daily for 7 days [11] while the late stage is treated with nifurtimox-eflornithine combination therapy (NECT) [12]; nifurtimox is administered orally three times a day for ten days in combination with intravenous eflornithine given every 12h for 12 days. For *T. b. rhodesiense* HAT, the early stage is treated with suramin administered by slow intravenous infusions every 3-7 days for four weeks [13]; the late stage is treated using melarsoprol (an arsenical drug) administered intravenously in a 10-day regimen [11]. Melarsoprol treatment produces a reactive encephalopathy in ~10% of the patients, and half of those cases are fatal [11].

All the current anti-trypanosomal agents are considered undesirable due to low efficacy, difficulty in administration, dangerous side effects, and emerging resistance. However, there is hope for new drugs against HAT with two compounds fexinidazole and benzoxaborole (SCYX-7158), currently in advanced stages of clinical trials under direction by the Drugs for Neglected Diseases initiative (DNDi) [11, 14].

The major trypanosome species responsible for Animal African trypanosomiasis (AAT) are *T. congolense* and *T. vivax* [15, 16]. AAT is recognized as a major economic constraint especially in the rural areas where it is estimated to cause more

than $1300 million annual losses to livestock producers and consumers [17]. Other minor pathogens for AAT include *T. godfreyi*, *T. simae*, and *T. b. brucei*. Transmission mainly involves the tsetse fly but mechanical transmission by biting flies has been reported for some of the species under laboratory conditions [18-20]. In most primates including humans, trypanosome species causing AAT get killed by trypanolytic factors (TLFs) present in the serum [21]. *T. b. gambiense* and *T. b. rhodesiense* which infect both humans and animals developed mechanisms to avoid lysis by the human serum factors. *T. b. rhodesiense* express a serum resistance associated gene, SRA, which binds TLF-1 protecting the parasite from lysis [22] while *T. b. gambiense* uses a multifactorial strategy involving expression of specific glycoprotein (TgsGP) that prevents the parasite killing effect of apolipoprotein L1 APOL1 (part of the TLFs), and a mutation that inactivate haptoglobin-hemoglobin receptor resulting to decreased TLF-1 uptake [23, 24].

As early diverging eukaryotes, trypanosomes share features common in most eukaryotes, but also they have very unusual features which makes them interesting model organisms. For example they have a single mitochondrion and the mitochondrial genome is organized into a disk like structure called the kinetoplast [25]. The mitochondrion respiratory activity is also developmentally regulated; in bloodstream stage, the size and activity of the mitochondrion is significantly reduced because the parasite depends on glycolysis for energy generation. In contrast, a fully functional and more elaborate mitochondrion exists in the insect stage procyclic forms which mainly depend on oxidative phosphorylation of amino acids for energy generation. Another peculiar characteristic is separation of various metabolic enzymes including most glycolytic enzymes into a peroxisome-like organelle known as the glycosome [26, 27], and the use of RNA polymerase I to transcribe genes encoding the major surface proteins, VSG and procyclins [28].

1.3 *T. brucei* cell structure

Trypanosomes have a single flagellum with a canonical axoneme structure plus an associated paraflagellar rod [29]. They swim with the flagellum tip leading. Apart from cell motility, the *T. brucei* flagellum is involved in cytokinesis and might function as a sensory organelle [30]. The flagellum is attached on the entire cell body (apart

from the distal tip) using a filament and a set of four specialized microtubules that form a structure known as the flagellum attachment zone (FAZ) [31]. Trypanosomes have several other organelles typical of eukaryotic cells such as the nucleus with nuclear membrane, Golgi apparatus, endoplasmic reticulum network, endosomes and lysosomes. The single mitochondrial DNA is composed of concatenated maxi and mini circles. The maxi-circles contain genes encoding mitochondrion proteins; some of their pre-mRNAs undergoe editing using guide RNAs (transcribed from both maxi and mini circles) as templates to generate mature mRNAs [32, 33]. The kinetoplast is physically attached to the basal body of the flagellum and this is important for segregation during cell division [34]. A trypanosome cell where the kinetoplast is placed posterior to the nucleus is known as the trypomastigote (see Figure 1.3), if the kinetoplast is placed anterior of the nucleus the cell is referred to as an epimastigote. A sub-pellicular microtubules corset helps to define the cell shape. The cytoskeleton network is enclosed by the plasma membrane and is only interrupted where the flagellum exits the cell body to form the flagellar pocket, which is the only site of endo and exocytosis [35].

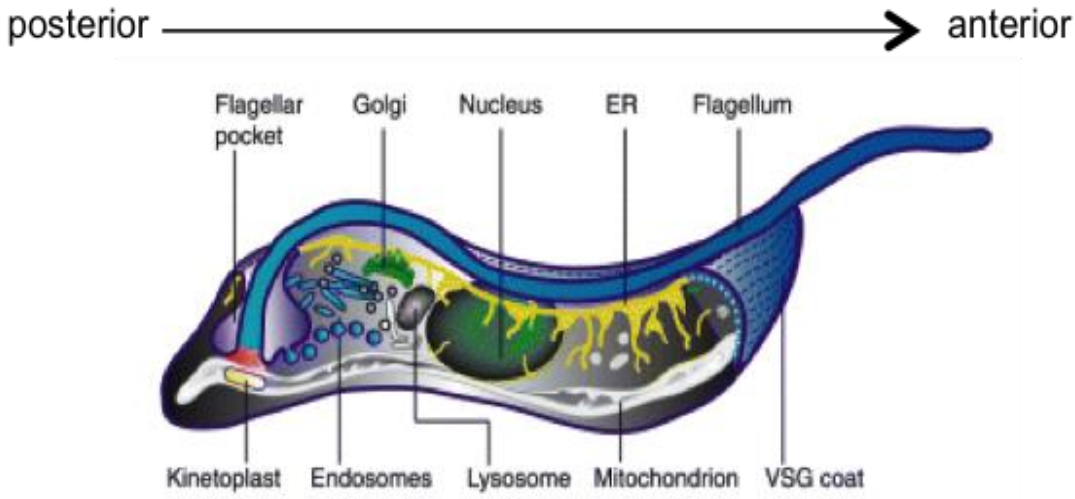

Figure 1.3 Ultrastructure of a *T.brucei* bloodstream form trypomastigote. Taken from [36].

1.4 *T. brucei* life cycle

The *T. brucei* life cycle is digenetic. It involves a vertebrate host and a blood feeding insect, the tsetse fly (Figure 1.4). The environments in the two hosts differ in temperature, nutritional levels, pH and immune response. To survive in such variable

environments, trypanosomes have to sense and respond accordingly. Within various tsetse fly tissues more than 10 distinct morphological forms have been observed [37]. In the mammalian host, pleomorphic strains exist as proliferative long slender forms, which transform to growth arrested stumpy forms at peak parasitaemia. *T. brucei* transmission is initiated when a tsetse fly feeds on an infected mammalian host. The hostile environment in the midgut kills the long slender bloodstream form while the stumpy forms differentiate into proliferative procyclic forms after 2-5 days. This entail cell enlargement, kinetoplast repositioning, increased mitochondrion respiratory activity and the replacement of the variant surface glycoprotein (VSG) coat with a less dense coat composed of procyclins; GPEET procyclin is present only in early procyclic forms, the late procyclic forms express EP procyclin instead [38-40]. After proliferation, migration to the foregut is accompanied by the appearance of elongated mesocyclic forms, and soon after long epimastigotes in the proventriculus [41]. The long epimastigotes undergo asymmetric division giving rise to both long and short epimastigotes [41], usually covered by a surface coat protein called BARP and with the kinetoplast positioned anterior to the nucleus [42]. The short epimastigote goes on to colonize the salivary gland [43, 44] and it is believed the long epimastigote helps to deliver the short epimastigotes (poor swimmers) to the salivary gland. Outgrowth of the flagellar membrane enables the short epimastigotes to get attached to the salivary gland epithelium while they differentiate to pre-metacyclics [30, 45]. Finally after a cell cycle arrest, nascent metacyclic forms covered with the VSG coat get released into the saliva ready to be transferred into the next mammalian host.

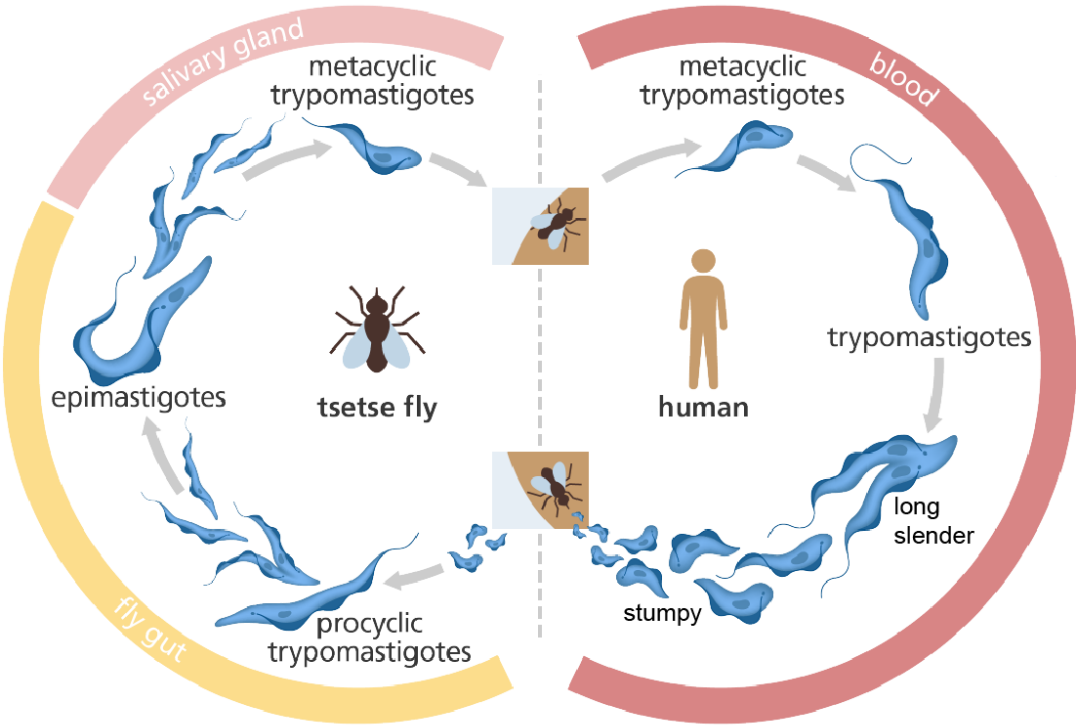

Figure 1.4 A simplified illustration showing the lifecycle of *T. b. brucei*. Image adopted from http://www.yourgenome.org/facts/what-is-african-sleeping-sickness.

Trypanosomes live extracellularly in the blood and tissue fluids of a mammalian host or in several tissues of the tsetse fly. As a result they get exposed to various immune defence systems present in the host. In the vertebrate host, the trypanosome cell surface is covered with a dense (15 nm thick) homogeneous coat [46] made up of 10^7 molecules of a single variant surface glycoprotein (VSG) which is linked to the cell membrane via a glycosylphosphatidylinositol (GPI) anchor [47]. The VSG coat is highly immunogenic but trypanosomes escape adaptive immunity through periodic switching to a different VSG; a phenomenon known as antigenic variation [48]. This occurs in one out of a thousand cells [49]. Trypanosome infection is therefore characterized by waves of parasitaemia where trypanosome populations are periodically eliminated but parasites that switch VSG survive and re-establish an infection. Apart from the variable function, the dense and thick VSG coat protects the less variable or invariant surface proteins such as the nutrient receptors from immune effectors [50]. Since the trypanosome virulence depends on the VSGs, the parasite has dedicated about 30% of its genome to archiving about 2000 VSGs genes and gene fragments; most of them are pseudogenes or have frame shift

mutations. However, productive VSGs are generated via mosaic formation, hence the VSG repertoire is almost unlimited [50-52].

Surface coat exchange during life stage transitions is used as a hallmark for trypanosome life cycle progression, but is only one of the many changes associated with the extensive transcriptome and proteome remodelling necessary for survival in and adaptation to different host environments. It is often initiated by environmental cues such as, pH, nutrient levels, temperature and chemical triggers [53, 54].

1.5 Gene regulation in trypanosomes

1.5.1 Transcription

Unlike other eukaryotes, trypanosomatids have a peculiar mechanism to transcribe and process mRNAs [55]. Transcription in trypanosomatids is polycistronic. In the case of *T. brucei*, genes present in the eleven megabase homologous chromosome pairs (which contains most RNA and protein coding genes) are organized into long unidirectional clusters of tens to hundreds of genes. Introns and individual promoters are absent for almost all of the protein coding genes [56]. Transcription initiation by RNA polymerase II is bidirectional and appears to be regulated by histone variants and modifications rather than transcription factors [57, 58]. The only exception is the splice leader (SL) gene cluster where a specific promoter and specific transcriptional factors are required [59-61].

Genes present in a single transcription unit are mostly functionally unrelated and it is assumed that transcription by RNA polymerase II happens at a uniform rate. However, large differences in mRNA and protein levels are observed between genes belonging to the same polycistronic transcription unit, and between developmental stages [62]. This is due to post-transcriptional regulatory mechanisms operating at the levels of pre-mRNA processing, mRNA transport, localization, stability and translation. Monocistronic mature mRNAs are generated from the long pre-mRNA through trans-splicing and coupled polyadenylation of the upstream transcript [63]. Trans splicing entails addition of a 39 nucleotide splice leader sequence to the 5’ of each mRNA. This provides the mRNA cap (cap 4) plus elements needed to improve mRNA stability and translation. Cis-splicing is rare with only two trypanosome

transcripts having been reported to be cis-spliced [64, 65]. Transcripts required in high amounts exist as multi-copy genes in the genome. The highly expressed genes encoding VSG and procyclins are transcribed by pol I which is ten times more active than pol II [63].

1.5.2 Post transcriptional regulation in trypanosomatids

After maturation in the nucleus, the mRNA is bound by RNA binding proteins (RBPs) which help with the export from the nucleus to the cytoplasm. Here the fate of the mRNA is determined by additional RBPs that get recruited mostly via cis-acting elements present in the 3' untranslated regions (UTRs), or by protein-protein interactions to form a messenger ribonucleoprotein (mRNP) complex. Translating mRNA associates with the cap binding protein eIF4E and interacts with the scaffold protein eIF4G that circularizes the mRNP by interacting with poly A binding protein PABP1 bound to the poly A tail at the 3' UTR of the mRNA. Additional translation initiation factors help to recruit the 40S ribosomal subunit to form the translation pre-initiation complex that begins to scan for the start codon. Some *T. brucei* translation factors are highly expanded with genes encoding six eIF4E and five eIF4G variants being present in the genome [66]. *Tb*eIF4E sequences are highly diverged from those in other eukaryotes, perhaps to accommodate the hyper-methylated cap 4 structure. eIF4E-4 and its partner eIF4G-3 are considered responsible for the global translation; it is not known whether translation initiation is regulated by having different eIF4E and eIF4G combinations. On the other hand, mRNAs that are not translated can be stored in membrane free granules/aggregates in association with different protein components; examples of such aggregates are p-bodies and stress granules [67]. P-bodies are thought to be the sites of mRNA degradation usually containing the decapping enzymes, the enhancer of decapping RNA helicase DHH1 and the 5'-3' exoribonuclease XRN1. The stress granules are aggregates of mRNPs stalled in the process of translation initiation, normally induced by different types of stress [67]. Both types of granules are present in trypanosomes [68, 69]; a recent study identified many RNA metabolism-related proteins as new components of starvation granules in *T. brucei* procycic forms [70].

In yeast and mammalian cells, mRNA translation and degradation are interlinked and considered as competing processes [67]. mRNA degradation is initiated by shortening of the poly A tail by the CCR4/CAF1/NOT complex [71], then the exosome complex degrades the mRNA from 3'-5' direction [72, 73]. After decapping, XRN1 degrades the mRNA from the 5'-3' direction [67, 74]. So far the *T. brucei* decapping protein is still unknown but trypanosome cell extracts are known to contain decapping activity [75]. Unlike some other eukaryotes, kinetoplastids lack a CCR4 homolog, the rest of the NOT complex subunits are present [76]. CAF1 is the major deadenylase in *T. brucei* with CAF1/NOT complex being required for global mRNA degradation [77]; another deadenylase PAN2/PAN3 seems to be important in the degradation of some unstable mRNAs [78, 79]. mRNA degradation is regulated via sequence specific RBPs that interact with proteins of the degradation machinery [80, 81]. In *T. brucei*, such RBPs that interact with the degradation machinery have yet to be discovered. In a yeast two hybrid assay, *T. brucei* CAF1 was shown to interact with the zinc finger proteins ZC3H15 and ZC3H5 and the RRM-containing proteins DRBD5 and RBP31 [82], whether this is true in trypanosomes is yet to be determined. Moreover, a recent screen identified numerous novel proteins that lack classical RNA binding domains (RBDs) but associated with mRNAs in bloodstream form trypanosomes [82], some of the candidates were previously shown to increase or decrease expression of a reporter mRNA when bound on the 3' UTR [83]; for most of them the mode of action and their biological functions are unknown. Follow up studies on those that decreased reporter mRNA might provide hints on how the degradation machinery gets recruited.

To date several RNA binding protein (RBP) families have been studied in *T. brucei*, these mostly include proteins containing conserved RNA binding domains such as RNA recognition motifs (RRM), PUF, CCCH-type zinc finger and ALBA domains [55, 84]. Similar to other organisms, RBPs with RRM domains are the most common in *T. brucei*; there are around 70. They include several RBPs conserved in other eukaryotes such as PABP1 and HnRNPH, while many others are unique to trypanosomatids [85, 86]. The RRM domain binds specific sequences on a single stranded RNA, but can also interact with DNA as well as other proteins. About half of the proteins containing RRM domain are required for *T. brucei* growth in at least one life-cycle stage [87]. However, only a few have been characterized. For example UBP1 and UBP2 whose mRNA targets are not known yet [88], RBP42 which binds

transcripts encoding energy metabolism proteins [89], HnRNPF/H which is expressed in the nucleus and cytoplasm, and can influence both mRNA stability and splicing [90]; finally, RBP10 [91] and RBP6 [92] are both implicated in differentiation regulation (will be described later).

T. brucei has about 40 RBPs with $CX_8CX_5CX_3H$ zinc finger domains. ZFP1, ZFP2 and ZFP3 are involved in differentiation from bloodstream to procyclic forms [93-95]. ZC3H20 is required for growth in procyclic forms; two procyclic specific mRNAs are bound and stabilized by ZC3H20 [96]. ZC3H11 is only essential in bloodstream form. In the procyclic forms ZC3H11 is required for heat shock response; it acts by stabilizing mRNAs encoding chaperone proteins through binding UAU motif present in the 3' UTRs [97], and interaction with MKT1 and PBP1 which then recruits PABP [98].

PUF (Pumilio and FBP) proteins are known to regulate mRNA stability and translation in other organisms. *T. brucei* has 11 PUF proteins. PUF9 stabilizes a small number of transcripts important in cell cycle control; a UUGUACC motif is required for PUF9 mediated regulation [99]. PUF7 and PUF10 are localized in the nucleolus. PUF7 interacts with nuclear cyclophilin 1, and has effect on rRNA maturation in procyclic forms [100]. NRG1, another nucleolar protein, interacts with PUF7 and PUF10 and the three proteins are implicated in the regulation of the mRNA that encodes GPEET procyclin surface protein [101].

There are four *T. brucei* proteins with an ALBA (acetylation lowers binding affinity) domain, which can bind both DNA and RNA. All *Tb*ALBA proteins are present in the cytoplasm, but upon nutritional stress they relocate to stress granules [102, 103]. ALBA1/2 recognizes regulatory elements present in the GPEET procyclin mRNA, and can form complexes with ALBA3 and ALBA4. ALBA3 interacts with eIF4E4 and both ALBA2 and 3 partially co-migrate with polysomes [103]. This, taken together with the stress granule relocation indicates involvement in translation control. ALBA3/4 may be required for trypanosome differentiation regulation in the tsetse fly [102]. However, ALBA protein target mRNAs and their mode of action remain to be determined.

1.6 Control of trypanosome differentiation

T. brucei differentiation is triggered by external stimuli present in the different host environments. In other eukaryotes, the ultimate target of the signal transduction

pathways is usually a transcription factor which function to regulate specific gene expression. In case of *T. brucei* and other trypanosomatids, this is not possible since transcription control is absent [104]. Currently, studies investigating the molecular mechanisms regulating *T. brucei* differentiation are limited to the transition from long slender bloodstream to procyclic forms; which until very recently was the only one available. Some of the signaling pathways that are known or speculated to operate during bloodstream to procyclic life stage transitions can be classified into two phases.

1.6.1 Bloodstream to stumpy forms transition

Transition from bloodstream to stumpy forms is irreversible [105]. It involves well-coordinated developmental programs characterized by cell cycle arrest and profound changes in cellular morphology and metabolism [53, 54]. In the mammalian host a parasite-specific signal acts as trigger for stumpy cells formation; it is commonly referred to as stumpy induction factor (SIF). So far the SIF molecule (or mixture of molecules) still remains to be identified. It is known that SIF gets released by proliferating long slender bloodstream forms at peak parasitaemia (>10^8 cells/ml) and acts in a mechanism similar to quorum sensing in microbial communities [106, 107] Once SIF reaches a critical threshold, it triggers long slender cells to differentiate to cell-cycle arrested stumpy forms (adapted for survival in tsetse fly). How the parasite perceives SIF is yet to be determined since the identity of SIF is unknown. The possibility that cAMP signaling pathways are involved in SIF signal transduction has been widely investigated in *T. brucei* [106, 108, 109]. This was based on the reports that cAMP levels change during bloodstream to stumpy form differentiation [110]. It turned out the parasite do not recognise cAMP directly but rather the products of cAMP; since non-hydrolysable cAMP failed to induce differentiation while hydrolysis products of cAMP analogues induced stumpy formation more efficiently [111]. The downstream effector proteins targeted by the AMP signal in trypanosome are unknown. The closest *T. brucei* homologue for protein kinase A is not activated by cAMP *in vitro* [109, 112], and genes encoding G-protein coupled receptors and heterotrimeric G-proteins are absent in *T. brucei* genome [109]. It is speculated the large family of transmembrane receptor-like adenylate cyclases (ACs) present in *T. brucei* could compensate for the missing receptor via their highly variable N-terminal extracellular domain. Several receptor-

type flagellar ACs are developmentally regulated and show differential localization on the flagellum [113, 114]. So far no putative ligands for ACs are known, but depletion of insect stage specific adenylate cyclase 6 affected social motility [115]. This demonstrates a possible role for cAMP signaling in response to extracellular stimuli. In bloodstream forms, ACs belonging to the ESAG4 or ESAG4-like subfamily are required for cytokinesis [116].

Additional cAMP signaling components driving stumpy formation were recently described in a genome-wide RNAi screen that selected trypanosomes resistant to SIF mimetics [117]. A putative RNA binding protein RBP7 was identified. When overexpressed in pleomorphic bloodstream forms, RBP7 promoted stumpy formation demonstrating that it is involved in normal quorum sensing. RBP7 mRNA targets and mechanism of action remains to be determined. Other factors identified from the screen included genes encoding purine metabolism proteins, kinases, phosphatases and several proteins of unknown function.

1.6.2 Stumpy to procyclic forms transition

The stumpy to procyclic form transition naturally occurs in the midgut of a tsetse fly. The decrease in temperature (from 37^{o}C to ~20^{o}C) experienced a few hours post ingestion sensitizes stumpy cells for signals to differentiate to procyclic forms. TCA cycle intermediate citrate/*cis*-aconitate (CCA) was identified [36, 118] as a possible physiological signal. The signal is perceived via a family of surface transporters known as the PAD proteins [119] expressed in stumpy forms. Cold shock at 20^{o}C leads to increased (4x) PAD2 protein levels and relocation of PAD2 from the flagellar pocket to the cell surface [119]. This results to increased signal recognition when stumpy cells are incubated at 20^{o}C, which enables differentiation induction with much less cis-aconitate (micromolar concentration instead of the usual 6mM at 27^{o}C [36]); possibly reflecting conditions encountered in the tsetse fly midgut.

A tyrosine phosphatase PTP1 prevents pre-mature differentiation of the short stumpy (Figure 1.6.2) by inhibiting a serine/threonine phosphatase PIP39 [120, 121]. The differentiation trigger citrate/*cis*-aconitate acts by inactivating PTP1, this together with cold shock leads to upregulation of phosphorylated PIP39 (active form) which in turn activates PTP1, creating a positive feedback loop [120]. Since PTP1 is located

in the cytoplasm/cytoskeleton, relocation of phosphorylated PIP39 to the glycosome prevents inactivation and ensures differentiation is irreversible. PIP39 phosphorylation and glycosomal localization are required for differentiation since catalytically dead or localization impaired mutants are unable to rescue differentiation defects caused by PIP39 depletion [120]. The kinase responsible for PIP39 activation is unknown. Also, the substrates for PIP39 remain to be determined; perhaps PIP39 acts by regulating some of the metabolic activities in the glycosome.

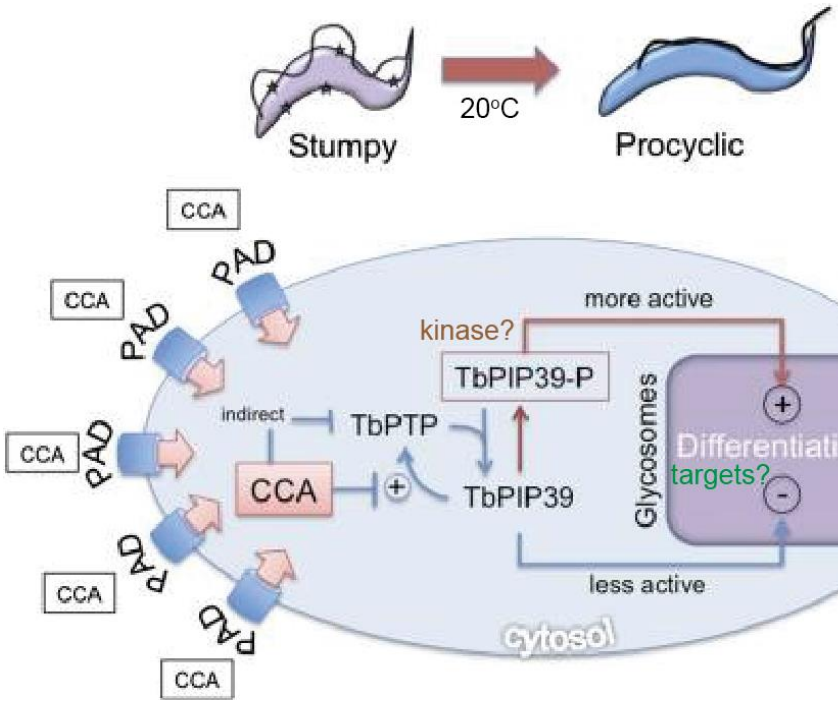

Figure 1.6.2 Model illustrating *T. brucei* phosphatase signaling cascade driving stumpy to procyclic form differentiation. The regulatory interactions are mostly based on *in vitro* data. Image taken from [120]. PAD= protein associated with differentiation, CCA= citrate/*cis*-aconitate.

Apart from PTP1, the two kinases RDK1 and RDK2 act as repressors of differentiation in long slender bloodstream forms [122]. In contrast, the protein kinases MAPK5, ZFK, and TOR4 prevent premature differentiation to stumpy forms [123-125]; it remains to be determined whether they act individually or are part of a cascade. Stumpy to procyclic form differentiation can also be triggered by pronase, acting independently of PIP39 signaling [126].

1.7 RBPs regulating developmental changes in *T. brucei*

Gene expression changes precede the profound morphological and metabolic changes observed during trypanosome differentiation. In the *T. brucei* stumpy to procyclic form transition, 69 genes are up regulated within one hour of differentiation initiation [127]; some of these genes encode proteins linked to mRNA metabolism, which is in accordance with post transcriptional mechanisms regulating gene expression in trypanosome. Several RNA binding proteins that are tightly regulated between life cycle stages have been shown to modulate developmental capacity in *T. brucei* [91, 92, 128]. Selected examples are described below.

ALBA3 and ALBA4 are expressed in all *T. brucei* lifecycle stages apart from those found in the proventriculus of the tsetse fly [102]. Depletion of ALBA3/4 in procyclic forms slowed cell growth and resulted to a cellular morphology resembling that of the epimastigote while overexpression of ALBA3 in the tsetse fly impaired normal differentiation in the proventriculus [102]. Three zinc finger RBPs have been implicated in the *T. brucei* life cycle transition from bloodstream to procyclic form: ZFP1 is procyclic specific and is required for repositioning of the kinetoplast during stumpy-to-procyclic form differentiation [93]; ZFP2 is present in both bloodstream and procyclic forms and is important for efficient differentiation to the procyclic form [94]; and ZFP3 was found to bind EP1 and GPEET mRNAs that encode procyclic surface coat proteins, ZFP3 depletion decreased target mRNA levels and overexpression increased their translation [129]. RBP6 mRNA is up regulated in trypanosomes found in the proventriculus of the tsetse fly. Ectopic expression of RBP6 for ten days in procyclic forms allows production of epimastigotes and infective metacyclic forms [92]. In case of stumpy formation, overexpression of RBP7 in pleomorphic bloodstream form cells promoted pre-mature cell cycle arrest and increased capacity for differentiation to procyclic forms [117]. REG1, a putative RNA binding protein acts as a repressor of stumpy enriched mRNAs in the long slender bloodstream forms. Depletion of REG1 in long slender bloodstream form promoted stumpy formation *in vivo* independent of quorum sensing, whereas overexpression primed the cells to differentiate to procyclic forms [130]. REG1 and RBP7 mRNA targets remains to be determined.

RBP10, which is the focus of this study; is a cytosolic RNA binding protein with a single RRM domain. RBP10 is expressed only in long slender bloodstream forms

where it is essential [91]; in stumpy and procyclic forms, RBP10 protein is undetectable [91, 105, 131]. Over-expression of RBP10 in pleomorphic bloodstream forms for 24 hours blocked *in vitro* differentiation to procyclic forms [91]. Transcriptome analysis revealed a widespread effect on gene expression upon alteration of RBP10 levels. In bloodstream forms, depletion of RBP10 led to up-regulation of procyclic specific mRNAs whereas mRNAs more abundant in bloodstream form cells were significantly down-regulated. A partial reversal of the transcriptome changes seen in bloodstream form occurred when RBP10 was over expressed in procyclic forms [91]. In both experiments, secondary effects due to growth inhibition could not be ruled out since mRNAs encoding several RNA binding proteins, as well as cytoskeleton and flagellum proteins were also significantly affected; these made it difficult to identify potential RBP10 mRNAs targets. Nonetheless, the transcriptome data indicated three possibilities; i) RBP10 acts as a positive regulator for the mRNAs more abundant in bloodstream forms, ii) RBP10 could be a negative regulator for procyclic specific mRNAs that are highly unstable in long slender bloodstream forms, and iii) RBP10 does both functions mentioned above. The second option was strongly supported by two main findings: first, RBP10 promoted reporter mRNA degradation and translation suppression when attached to the reporter 3' UTR, second, RBP10 failed to comigrate with the translating polyribosomes after a sucrose density gradient centrifugation [91]. For the current study, we decided to investigate further how RBP10 acts as a negative regulator, and why that function is relevant in the bloodstream form life cycle stage.

1.8 Aim of the study

The main goal for this project was to investigate RBP10 mechanism of action in order to understand how RBP10 modulates trypanosome gene expression and development.

The specific objectives included:

1. Identification of protein partners of RBP10 using genome wide yeast two hybrid screen and TAP purification.

2. Investigating the early effects of RBP10 perturbation on global gene expression and translation using polysome profiling coupled to RNA sequencing.

3. Identifying the target mRNAs and *in vivo* mRNA binding site of RBP10.

4. Analyzing the role of RBP10 in differentiation regulation in *Trypanosoma brucei*.

2 Materials and methods

2.1 Trypanosome cell culture and manipulation

The experiments in this study were carried out using monomorphic *T. brucei* Lister 427 bloodstream or procyclic form cells. The only exception is the differentiation studies where pleomorphic EATRO1125 cells were used. All the cell lines constitutively expressed the Tn10 tet repressor [132, 133].

2.1.1 Bloodstream form culture

Bloodstream form parasites were cultured at a density between 0.2-1.5x10^6 in HMi-9 medium at 37^{o}C with 5% CO2.

The pleomorphic bloodstream EATRO1125 cells were maintained in the same conditions but diluted after every 24h to keep the cell density below 5x10^5 cells/ml. For stumpy forms generation, the pleomorphic bloodstream EATRO1125 cells were grown to high density (2-3x10^6 cells/ml) in HMI-9 medium with 1.1% methylcellulose and then left at the stationary phase for several days.

HMi-9 medium [134]**:** 17.66 g/l Iscove's modified Dulbecco's medium, 3.024 g/l $NaHCO_3$, 136 mg/l hypoxanthine, 110 mg/l sodium pyruvate, 39 mg/l Thymidine, and 28 mg/l Bathocuprono disulfonic acid disodium salt. The media was filter sterilized and stored at 4^{o}C in 450 ml aliquots. Prior to use the media was supplemented with 10% (v/v) heat inactivated (30 min at 55^{o}C) fetal bovine serum, 50 U/ml penicillin/streptomycin, 1.5 mM L-Cysteine-HCL.H_2O, and 7 µl β-mercaptoethanol (14.1 M stock, sigma).

HMI-9 medium with 1.1% methylcellulose, 1 litre: [123]

2X HMI-9	560 ml (complete with FCS, L-Cys, βME & PenStrep)
Methylcellulose	440 ml (2.5% in water)

2.5% methylcellulose was prepared in water then dissolved by stirring at 4^{o}C for 1-2 days. After autoclaving and stirring at 4^{o}C to re-dissolve, the volume was topped up using sterile water then further stirred (1-2 days, 4^{o}C) to make sure it is completely dissolved.

2.1.2 Procyclic form culture

The procyclic forms cells were maintained at a density between 0.5-5x10^6 cells/ml in MEM-Pros medium at 27^{o}C.

MEM-Pros medium: 16.55 g/l MEM-pros mixture (Biochrom), 1% (v/v) MEM non-essential amino acids (sigma), 1% (v/v) MEM vitamins (sigma) and 10mg/l phenol red. The pH was adjusted to 7.4 using NaOH, the media was filter sterilized and stored in 450 ml aliquots at 4^{o}C. Prior to use the media was supplemented with 10% (v/v) heat inactivated (30 min at 55^{o}C) fetal bovine serum, 7.5mg/l Hemin and 50U/ml penicillin/streptomycin (sigma).

To dilute cells, the complete media was first pre-warmed in a water bath at 37^{o}C for 30min.

2.1.3 Transfection of Bloodstream/procyclic trypanosomes

For each transfection approximately 1-2x10^7 cells were used. The cells were washed twice using 0.4ml of the transfection buffer; for bloodstream forms (BS) the Amaxa-buffer (90 mM NaH_2PO_4, 5 mM KCl, 0.15 mM $CaCl_2$, 50 mM HEPS, pH 7.3 and for procyclics (PC) the ZPFM buffer (132 mM NaCl, 8 mM KCl, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 1.5 mM MgOAc.$4H_2O$, 90 µM $CaOAc_2$, pH 7.0).

After washing, the cells were resuspended in the appropriate transfection buffer (100 µl for BS, 400 µl for PC) and mixed with 10µg of linearized plasmid in a 2mm gap electroporation cuvette. Transfection of the bloodstream trypanosomes was performed using program X-001 of the Amaxa Nucleofactor (Lonza Cologne AG, Germany). For the procyclic forms the BTX electroporation machine (Harvard apparatus) was used with the settings of 1.5 KV and resistance R2. The cells were then transferred to 25 ml of pre-warmed medium and left overnight to recover. The following day, the selection antibiotics were added and the cells plated in serial dilution on a 24 well plate. For the pleomorphic EATRO1125 cells, 8 µg/ml hygromycin and 2 µg/ml blasticidin were used for selection.

For monomorphic bloodstream and procyclic forms the antibiotics concentrations are listed below.

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

Positive clones were identified 7-10 days later. After scaling up the cell numbers, the transgene expression for at least three independent clones was confirmed using Western blotting. Aliquots of the generated cell lines were frozen in appropriate medium with 10% glycerol and stored in liquid nitrogen.

2.1.4 Cloning and Plasmids constructs

Primers used for cloning (see section 2.9.2) were ordered from Biomers.net. Genomic DNA from *T.brucei* was isolated using illustra tissue and cells genomicPrep Mini Spin Kit (GE Healthcare). PCR was done using high fidelity Q5 DNA polymerase and buffer (NEB) according to the manufacturer's instructions. The PCR product was digested with the appropriate restriction enzymes (NEB or Fermentas) and ligated to a plasmid with the same sticky ends. The ligation product was then transformed into competent *E.coli* DH5-alpha cells using heat shock at 42^{o}C for 1min. Positive clones were identified by colony PCR using GoTaq DNA polymerase (promega) and further confirmed by restriction digests after plasmid purification (Macherey-Nagel mini prep kit). Only one plasmid was selected for the final verification by Sanger DNA sequencing (GATC Biotech).
The details of the plasmids and oligonucleotides used in this study are provided in section 2.9.

2.1.5 Stable cell lines for inducible expression

The tetracycline-inducible constructs for *RBP10* RNAi and over expression have been described in [91]. The 4E-IP was ectopically expressed using the plasmid pHD2533. For the tethering assays, a lambdaN peptide was fused to RBP10 or six different fragments of the RBP10 open reading frame (ORF). The constructs were

separately transfected in a cell line constitutively expressing the *CAT* reporter with *boxB actin* 3' UTR. Expression of the fusion protein was induced for 24h using tetracycline (100 ng/ml) and the levels of the *CAT* mRNA and CAT protein quantified using Northern blot and CAT assay respectively.

2.1.6 Endogenous tagging of RBP10

A cell line with *in-situ* TAP-RBP10 gene was generated by replacing one endogenous copy of *RBP10* with a gene encoding N-terminally TAP tagged RBP10. To do this, a construct with puromycin resistance gene plus TAP tag cassette was flanked on the 5' end with a fragment of *RBP10 5' UTR*. Also, downstream on the 3' end, the N terminal region of *RBP10 ORF* was cloned in frame with the TAP tag. Prior to transfection, the plasmid (pHD2506) was cut with *Sac I* and *Apa I* enzymes to allow homologous recombination. Using cells with *in situ* TAP-RBP10, the other copy of *RBP10* was replaced with a knock out cassette (pHD2061) containing blasticidin resistance gene, and the absence of untagged RBP10 was confirmed by Western blotting.

A similar approach was used for V5 *in situ* tagging; the construct used pHD1914 had a blasticidin resistance gene and in this case the knockout was not done.

2.2 Tandem affinity purification (TAP)

Approximately $2x10^{10}$ bloodstream form cells expressing either in-situ TAP-RBP10 (pHD2506) or the tet-inducible GFP-TAP (pHD1743) were harvested and used for TAP as previously described [135]. Three biological replicates were done, with or without RNAse A treatment; triplicate results from the cells with inducible GFP-TAP served as background control. The cell pellet was lysed in 6 ml of lysis buffer (20 mM Tris pH7.5, 10 mM NaCl, 1× complete protease Inhibitor without EDTA (Roche), 0.1% (v/v) IGEPAL) by passing 15-30 times through a 21G needle. The cell lysate was cleared by centrifugation at 15000 g for 10 minutes at 4^{o}C. The supernatant was transferred to a new tube and NaCl adjusted to 150 mM. An aliquot (input sample) equivalent to ~$5x10^{6}$ cells was taken and 20 µl protein loading buffer added.

250µl IgG sepharose beads were transferred to a Bio-Rad column and pre-washed with 10 ml IPP150 buffer (20mM Tris pH 7.5, 150mM NaCl, 0.1% IGEPAL). The beads were then incubated (on a rotator) with the cleared lysate for 2h at 4^{o}C. The unbound fraction was collected by gravity and an aliquot (flow through sample) equivalent to ~5x10^{6} cells taken for further analysis. The beads were washed three times in IPP150 buffer and once using TEV cleavage buffer (IPP150 buffer plus 0.5 mM EDTA and 1 mM DTT). The TEV cleavage was done overnight at 4^{o}C using 150 units of TEV protease ('homemade') in 1 ml TEV cleavage buffer. On the next day, the eluate was recovered by gravity and an aliquot (~1.5% of TEV eluate sample) taken for further analysis.

250 µl of calmodulin affinity beads were transferred to a Bio-Rad column and washed with 10 ml calmodulin binding buffer (IPP150 buffer plus 10 mM β-mercaptoethanol, 1 mM magnesium acetate, 1 mM imidazole, 2 mM $CaCl_2$). 3 ul of 1M $CaCl_2$ was added to 3 ml of calmodulin binding buffer and then mixed with the eluate (~1ml) from the TEV cleavage. This sample was incubated with the pre-washed calmodulin beads for 1.5h at 4^{o}C. The beads were washed three times with the calmodulin binding buffer and the bound proteins eluted using calmodulin elution buffer (IPP150 buffer plus 10 mM β-mercaptoethanol, 1 mM magnesium acetate, 1 mM imidazole, 2 mM EGTA). The eluate was concentrated by TCA precipitation, separated on a 10% SDS-polyacrylamide gel for only 2 cm and the proteins visualized by colloidal Coomassie staining. The gel area was excised and analysed by mass spectroscopy.

2.3 Yeast two-hybrid screen

The AH109 yeast strain (Clontech) was used. The cells were grown at 30^{o}C in YPDA medium (20 g/l peptone, 10 g/l yeast extract, 2% gucose, 18 g/l agar (for plates), 0.2% Adenine hemisulfate salt). For protein-protein interaction studies, the Match Maker Yeast-Two hybrid system (Clontech) was used according to manufacturer's instructions. As bait, *RBP10 ORF* was cloned in frame with the yeast GAL4 DNA binding domain, with a myc tag on the N-terminus, using the pGBKT7 plasmid (pHD2361). For the prey, the RBP10 ORF was cloned into pGADT7 plasmid (pHD2368) which has GAL4 activation domain and an HA tag on the N-terminus.

Similar constructs were made for pairwise screens using two fragments of RBP10 (F2 & F3) and the translation factors EF1-α, EF2, eRF1, eRF3, eIF-5A and eIF-2B.

2.3.1 Yeast transformation

Fresh competent AH109 yeast cells were prepared by the LiAc method [136]. Pairwise analysis involved co-transformation with the bait and prey plasmids. Briefly, 100 µl of competent AH109 cells was mixed with 200 ng of each plasmid DNA (prey and bait), 10 µl salmon sperm DNA (10mg/ml stock) and 600 µl PEG/LiAC solution (40% PEG 4000, 100mM LiAc, 1X TE buffer, pH 7.5). The cells were incubated at 30^{o}C (with shaking, 200rpm) for 30 min, and then heat-shocked in the presence of 10% DMSO for 15 min at 42^{0}C. Positive clones were selected on SD quadruple drop-out medium (SD-QDO; minimal SD media lacking trypophan, leucine, adenine and histidine) and further confirmed phenotypically by growth (3-5 days) in presence of X-alpha-Gal substrate. The interaction between murine p53 and SV40 large T-antigen served as positive control, with Lamin and SV40 large T-antigen as negative bait and prey controls. Protein expression was checked by Western blotting using the myc and HA tags for detection.

2.3.2 Genome wide yeast two-hybrid screen and bioinformatics analysis

RBP10 was used as a bait to screen a library made up of random genomic DNA fragments from *T. brucei*. The library was prepared by Dr. Esteban Erban [98]. For the screen, the library was transformed into AH109 yeast cells expressing pGBKT7-RBP10. The positive clones were selected by growth on QDO medium for 5 days at 30^{o}C. 50 individual colonies were picked, checked by plasmid retransformation and the interacting partners sequences (prey plasmid) identified by Sanger DNA sequencing (GATC Biotech). The rest of the colonies were harvested en masse and the plasmid DNA isolated using plasmid mini prep kit (Macherey Nagel) after cell wall digestion using lyticase. To identify putative interacting partners, PCR amplification was carried out using pGADT7 vector-specific primers [98]. The PCR conditions; 95^{o}C for 2 min, 22 cycles of 95^{o}C for 1 min, 57.5^{o}C for 30 s, 72^{o}C for 3 min and a final extension of 5 min at 72^{o}C. To check the size distribution of the DNA smear, an aliquot of the PCR product was analyzed by agarose gel electrophoresis. Samples

were purified and analyzed using high-throughput illumina sequencing (CellNetworks Deep Sequencing Core Facility at the University of Heidelberg and EMBL). The sequencing library was prepared using standard illlumina kits; libraries were multiplexed (6 samples per lane) and sequenced for 50 cycles on the illumina Miseq system. The high-throughput data was analyzed as described in [98]. Briefly, the insert-vector junction was identified and the vector sequence removed using custom made PERL scripts. The sequences were aligned to the *T. brucei* TREU 927 reference genome using Bowtie [137], then sorted and indexed using SAMtools [138]. Sequences in-frame with an annotated ATG start codon and in-frame in the 5' UTR were then selected and counted using SAMtools [138] and custom PERL scripts [98].

2.4 Cross-linking and RNA immunoprecipitation

$3x10^9$ bloodstream form cells expressing in-situ TAP-RBP10 were concentrated in 20 ml FCS-free media and then transferred to a petri dish (145 mm radius). The cells were irradiated on ice using UV (400 mj/cm^2), washed in cold PBS and the cell pellet snap frozen in liquid nitrogen. The RNA immunoprecipitation was done as described in [139]. The cell pellet was lysed in 4 ml of the lysis buffer (10mM Tris pH7.5, 10mM NaCl, 2000U RNasin (Promega), 10 µg/ml leupeptin, 1× complete protease Inhibitor without EDTA (Roche), 0.1% (v/v) IGEPAL) by passing 15-30 times through a 21G needle. After centrifugation at 15000 g for 10 minutes at 4^{o}C to remove the cell debris, the NaCl concentration was adjusted to 150 mM and the cell lysate incubated with 200 ul of pre-washed IgG sepharose beads for 2h at 4^{o}C. The unbound fraction was collected, and beads washed three times in 10 ml of IPP150 buffer (20mM Tris pH 7.5, 150mM NaCl, 0.1% IGEPAL) and once in TEV cleavage buffer (IPP150 plus 0.5 mM EDTA, 1 mM DTT, 200U/ml RNasin (Promega)). The TEV cleavage was done using 150 units of TEV protease ('homemade') in 1 ml TEV cleavage buffer and rotating the beads for 2h at 4^{o}C. The eluate was collected by gravity and prior to RNA purification, the cross-linked protein was digested with proteinase K (50 µg proteinase K, 8 mM EDTA, 0.2% SDS per 50 µl of the eluate) at 42^{o}C for 15 minutes and the RNA isolated from both the bound and unbound fractions using Trifast reagent (Peqlab, GMBH). To assess the quality of the purified RNA, an aliquot of the sample was analysed by Northern blotting and the blot hybridized with a splice

leader probe. Total RNA from the unbound fraction was depleted of ribosomal RNA (rRNA) using RNAse H as described in [140] except that a cocktail of 50-base DNA oligos complementary to trypanosome rRNAs was used [141]. The recovered RNA from both bound and unbound samples was then analysed by RNA-seq.

2.5 Polysome fractionation and RNA seq

Approximately 3-5x10^8 cells were collected by centrifugation (850 x g, 10 minutes, 21^{o}C), and then treated in serum-free media with 100µg/ml cyclohexamide for 7 minutes at room temperature. The pellet was washed in 1 ml ice cold PBS, lysed in 350µl lysis buffer (20 mM Tris pH7.5, 20 mM KCl, 2 mM MgCl2, 1 mM DTT, 1000U RNasin (Promega), 10 µg/ml leupeptin, 100 µg/ml cycloheximide, 1× complete protease Inhibitor without EDTA (Roche), 0.2% (vol/vol) IGEPAL) by passing 15-30 times through a 21G needle, followed by centrifugation (15000 x g, 10 minutes, 4^{o}C) to clear the lysate. KCl was adjusted to 120 mM and the clarified lysate loaded on top of a 4 ml continuous linear 15-50% sucrose (w/v) gradient in polysome buffer (20mM Tris pH7.5, 120mM KCl, 2mM MgCl2, 1mM DTT, 10µg/ml leupeptin, 100µg/ml cycloheximide). After 2 hours of ultracentrifugation (40000 rpm, 4^{o}C; Beckman SW60 rotor), the samples were fractionated using Teledyne Isco Foxy Jr. system and the polysome profile recorded at 254 nm. 400 µl per fraction were collected and RNA isolated using Trifast reagent (Peqlab, GMBH). In the case of RBP10 tethering, lambdaN-RBP10 was induced (24h) and the distribution of the *CAT* reporter and *alpha-tubulin* mRNAs in the collected fractions detected by Northern blotting.

For RNAseq, bloodstream form cells plus or minus *RBP10* RNAi for 15h and, procyclic form cells with or without RBP10-myc over-expression for 6h were used. In this case 250 µl fractions were collected from each gradient and RNA was isolated after pooling the fractions into two groups; i) lighter fractions including monosomes, ii) denser fractions with at least two ribosomes. The amount of mRNA in the pooled samples was assessed by Northern blotting using splice leader RNA as probe. Also, total RNA was prepared from the input samples (~10% of total cell lysate) to quantify changes in the steady state mRNAs levels. All samples were depleted of rRNAs (as above) prior to analysis by RNA-seq.

2.5.1 High throughput RNA sequencing and bioinformatic analysis

RNA-seq was done at the CellNetworks Deep Sequencing Core Facility at the University of Heidelberg. For library preparation, NEBNext Ultra RNA Library Prep Kit for Illumina (New England BioLabs Inc.) was used. The libraries were multiplexed (6 samples per lane) and sequenced with a HiSeq 2000 system, generating 50 bp single-end sequencing reads.

The quality of the raw sequencing data was checked using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), and the sequencing primers removed using Cutadapt software [142]. The data was aligned to *T. brucei* TREU 927 reference genome using Bowtie [142], then sorted and indexed using SAMtools [138]. Reads aligning to open reading frames of the TREU 927 genome were counted using custom python scripts. Analysis for differentially expressed genes was done in R software using the DESeq2 package [143] with an adjusted p-value cut-off of 0.05. The wild-type transcriptome datasets from [144] was used to generate the list of developmentally regulated trypanosome mRNAs with a 2-fold cut-off. Comparative analysis was limited to the unique genes list [145]; for the RNA pull down, a minimum of 3-fold enrichment (bound versus unbound) in both replicates was was used as the cut-off. Categories enrichment between the datasets was checked using Fisher's exact test. The 3' UTR motif enrichment search was done using DREME [146]; annotated 3' UTR sequences were downloaded from tritrypDB and we considered only the mRNAs with 3' UTRs >20 nt.

For *de novo* assembly, the SPAdes [147] genome assembler was used. Contigs having the 'TGATATATTTTAAC' motif present in the 3' UTR of all VSG mRNAs were identified and analysed using the VSG identification tool (https://github.com/klprint/IdentifyVSGs). The random shotgun reads from the EATRO1125 genome were obtained from [148]; only contigs of 4 kb and longer were considered for metacyclic promoter identification.

2.6 Analysis of reporters containing *EP1* 3' UTR

The bloodstream form cells expressing the *CAT* reporter with either full length *EP1* 3' UTR (pHD1610) or a mutant version (pHD1611) lacking the 26mer (*EP1*Δ*26*) instability element were used for RNA immunoprecipitation using anti-RBP10 [91]

antibody. For the pull down, $4x10^8$ cells were irradiated on ice using UV (400 mj/cm^2), washed in cold PBS and the cell pellet lysed in 350 µl lysis buffer (10mM Tris pH 7.5, 10mM NaCl, 1000U/ RNasin (Promega), 1× complete protease Inhibitor without EDTA (Roche), 0.1% IGEPAL) by passing 15 times through a 21G needle. The lysate was cleared by centrifugation at 15000 g for 10 minutes at 4°C, NaCl was adjusted to 150 mM followed by incubation with 50 µl anti-RBP10 coupled agarose beads for 2h at 4°C. After washing the beads 5 times with IPP150 buffer (10 mM Tris pH 7.5, 150 mM NaCl, 0.1% IGEPAL), the beads were treated with proteinase K (50 µg proteinase K, 8 mM EDTA, 0.2% SDS per 50 µl of the beads) at 42°C for 15 minutes and RNA was isolated from both bound and unbound fractions using Trifast reagent (peqlab, GMBH). Equal amounts of the recovered RNA (eluate and flow through fractions) were reverse-transcribed into cDNA using RevertAid First Strand cDNA Synthesis Kit (Thermal Scientific) according to the manufacturer's instructions. 2 µl of the cDNA was used as template in a 50 µl PCR reaction to detect the *CAT* and *alpha-tubulin* genes. PCR was done using Q5 DNA polymerase and buffer (NEB) with 0.5µm of the following primers, for *CAT* (CZ5725; CZ689) and *alpha-tubulin* (CZ5725; CZ6168); the forward primer is identical since it anneals to the splice leader. Aliquots (10 µl) were removed after the 27, 31 and 36 cycles and analysed using agarose gel electrophoresis.

To determine if the regulation of the *EP* mRNA by RBP10 depends on the 26mer instability element, a stem-loop construct (pHD1984) targeting *RBP10* was transfected in the two cell lines. *RBP10* RNAi was induced using 100 ng/ml tetracycline for 17h or 24h and the levels of the CAT protein monitored using the CAT assay.

2.6.1 CAT assay

Chrolamphenicol acetyl transferase (CAT) activity was measured in a kinetic assay. It involves the transfer of an acetyl group from radiolabelled ^{14}C butyrl CoA to chrolamphenicol. The ^{14}C butyrl-chrolamphenicol becomes water insoluble and moves from the aqueous to the organic phase of the scintillation fluid.

For the CAT assay, $15x10^6$ cells expressing the CAT reporter were harvested (900 g for 5 min), washed once in PBS and the cell pellet resuspended in 200 µl of CAT

assay buffer (100mM Tris-HCL, pH 7.8). The cells were lysed by freeze-thawing three times in liquid nitrogen. After centrifugation at 15000 g for 3 min at 4^0C, the supernatant was transferred to a new tube and the total protein concentration measured by Bradford method. For longer storage the samples were kept at -80^{o}C.
BSA concentrations 0, 1, 2, 4, 8, 12, 16, 20 µg (100 µg/ml stock) in 800 µl of water were used to generate the Bradford's standard curve. To each, 200 µl of Bradford reagent (BioRad) was added, the samples were incubated for 5 min at room temperature and the OD was measured at 595 nm in a spectrophotometer (BioRad). In the case of the cell lysate samples, 5 µl was used for the Bradford measurement.
To measure the CAT activity 0.5-1 µg of total protein were scaled up to 50 µl using the CAT assay buffer. After addition of 2 µl chloramphenicol (40 mg/ml stock), 10 µl ^{14}C butryl CoA, 200 µl CAT assay buffer and 4 ml scintillation solution (Econfluor-2), the samples were analyzed after every 12-16 min for 1.5h in a scintillation counter (Beckman LS6000IC; using program 7) to detect ^{14}C. Once plotted, the data slope at the linear range (before saturation) was compared between samples to estimate the relative CAT protein levels.

2.6.2 Northern blotting

Total RNA was isolated (for solid and liquid samples) using Trifast reagent (peqlab, GMBH) according to manufacturer's instructions. The purified RNA was separated on formaldehyde agarose gel and blotted onto Nytran membranes (GE Healtcare). Detection was done using radioactively labeled DNA probes (Random Primer Labelling Kit, Stratagene) for *CAT*, *alpha-tubulin* mRNAs, and for total RNA an oligonucleotide (CZ4490) antisense to mini-exon. The signal was measured using phosphorimager (Fuji, FLA7000) and the quantification done using MultiGauge software.

2.7 Trypanosome differentiation

Pleomorphic EATRO1125, bloodstream form were used. For stumpy differentiation, the cells were maintained in HMi-9 with 1.1% methylcellulose and left to grow for 3-5 days without dilution to reach the stationary phase. The cell culture was diluted 1:5 with pre warmed PBS, filtered through MN616 ¼ filter papers (Macherey Nagel)

followed by centrifugation at 1400 g for 10 min. The cells were resuspended in HMi-9 medium without methylcellulose; differentiation was induced using 6mM cis-aconitate and incubation at 27ºC. After 17h, the cells were harvested, resuspended (5-7 x 10^5 cells/ml) in MEM-pros medium and maintained at 27ºC. The expression of the stumpy form marker PAD1 was checked by Western blotting.

2.7.1 RBP10 mediated trypanosome differentiation

Pleomorphic EATRO1125 bloodstream form cells expressing the Tet-repressor were transfected with a stem-loop construct (pHD1984) targeting *RBP10*. The growth of the cells was monitored for 3 days with and without RBP10 RNAi. Differentiation to procyclic forms was done by depleting RBP10 for 17 hours; the cells were pelleted, resuspended (~8x10^5 cells/ml) in procyclic form (MEM-pros) medium and incubated at 27ºC. As positive controls, wild type or uninduced cells (2x10^6 cells/ml) were treated with 6mM cis-aconitate (Sigma) at 27ºC, after 17h the cells were transferred into procyclic form media (~8x10^5 cells/ml) and maintained at 27ºC. The cell density was monitored after 0, 6, 10, 24, 48, and 72 hours, samples (~5x10^6) were taken at each time point and the expression levels of EP, PIP39 and RBP10 proteins analysed by Western blot.

To convert procyclic to bloodstream forms, EATRO1125 bloodstream form cells with an inducible RBP10-myc construct (pHD2098) were differentiated to procyclic forms using cis-aconitate as described above. The cells were cultured in presence of hygromycin (8µg/ml) and phleomycin (0.2µg/ml) for more than 3 months to generate well-established procyclic forms. RBP10-myc was induced using 100ng/ml tetracycline, the growth of the cells was monitored for 3 days and the protein levels of EP, PIP39, RBP10, hnRNPH , TAO and BARP detected by Western blot. The presence of the VSG transcripts was detected by semiquantitative RT-PCR as previously described in [92] using primers CZ6308/CZ6309. To generate blood stream forms, RBP10-myc was induced for 48 hours, the cells were pelleted, resuspended (2x10^5) in HMI-9 medium and incubated at 37ºC with 5% CO2. The cell density was monitored for 6 days; wild type or uninduced cells served as control.

2.7.2 Analyses of cell morphology and differentiation markers expression

DAPI stained cells were analysed by microscopy and the distance between the kinetoplast (K), nucleus (N) and posterior cell end measured using Fiji [149]. The switch of the surface coat proteins EP, GPEET and VSG was analysed by indirect immunofluoresence and FACS. Also, the differential localization of PGK protein was used to distinguish the bloodstream forms cells from the procyclic forms.

2.7.2.1 Indirect Immunofluorescence

$2x10^6$ cells were collected by centrifugation, re-suspended, then fixed in 4% paraformadehyde (in 1x PBS) for 18 minutes, sedimented again for 2 minutes, re-suspended in PBS and allowed to settle on poly-L-lysine coated slides for 30 minutes. Before staining, slides were blocked with 20% fetal calf serum (1x PBS) for 20 minutes. To detect the PGK protein, cells were permeabilised with 0.2% (v/v) Triton-X 100 (in 1xPBS) for 15 minutes at room temperature, washed twice then incubated for one hour with rabbit anti PGK antibody (1:1500; [150]). For the surface coat proteins, the cells were fixed in 100% methanol at -20^{o}C for 15 minutes, rehydrated in 1x PBS for 15 minutes, blocked for 20 minutes with 20% FCS, then labeled with mouse anti EP (1:500; Cedarlane, Canada), mouse anti phospho-GPEET (1:500; Cedarlane, Canada) and rabbit anti VSG-117 (1:500; kind donation from GA. cross) antibodies. The cells were washed 3 times before being stained with fluor-conjugated secondary antibody (1:500; mouse-Cy3 or rabbit-Alexa-488; Molecular probes, Eugene). Cellular DNA was stained with 100ng/ml DAPI in 1x PBS for 15 minutes. Images were taken using Olympus Cell-R microscope.

I took snapshots of random fields then Prof. Christine Clayton did the measurement and analysis of the K-N distances.

2.7.2.2 Flow cytometry

Approximately $5x10^6$ cells were fixed with 2% formaldehyde/0.05% glutaraldehyde at 4^{o}C for at least 1 hour. The cells were pelleted, washed twice with PBS then incubated with 200 µl (2% BSA in PBS) mouse anti-EP (Caderlane, Canada; 1:500) or rabbit anti-VSG-117 (1:500; from either G. cross or P. Overath) for one hour on ice. After washing twice, the cells were stained with the secondary antibody (1:500;

mouse-Cy5 or rabbit-Alexa-488; Molecular probes, Eugene) for one hour. Cells stained only with the secondary antibody and the unstained cells served as the negative controls. Flow cytometry was performed with BD FACSCanto II flow cytometer and the data analysed using FlowJo software (TreeStar Inc.).

2.8 Protein detection

Proteins were detected by Western blotting according to standard protocols. The only exception was PAD1 protein detection. In this case, the samples were not heat denatured and the SDS-PAGE gel was run at 4^{o}C using pre chilled buffers.

The following antibodies were used; anti-c-myc (mouse, 1:1000, Santa Cruz Biotech), anti-v5 (mouse, 1:1000, BioRad), anti-aldolase (rabbit 1:50000), anti-TR (rabbit 1:1000, from L. Siegel lab), anti-RBP10 (rat, 1:2000, [91]), anti-hnRNPH (rabbit,1:5000, [151]), anti-TAO (rabbit,1:100, [150]), anti-EP (mouse,1:2000, Cedarlane, Canada), anti-GPEET (mouse,1:2000, Cedarlane, Canada), anti-BARP (rabbit,1:2500, [42]), anti-PIP39 (rabbit,1:1000, [152]) and anti-PAD1 (rabbit, 1:1000, [153]). The secondary antibodies (1:2000, GE Healthcare) used were coupled to HRP, and the signal was detected using ECL solutions.

For the co-immunoprecipitation (co-IP), the sample lysis and immunoprecipitation was performed as described in section 2.6; without UV irradiation of cells and proteinase K treatment. 50 µl anti-myc or anti-V5 coupled agarose beads was used for immunoprecipitation.

2.9 List of plasmids and Oligonucleotides used

2.9.1 Plasmids list

Plasmid No.	Description	Primers	Resistance marker	Plasmid backbone
Ectopic inducible expression from rDNA locus				
pHD1984	Stem-loop against *RBP10*	*Wurst et al 2012*	HgrR	pHD1146
pHD2098	RBP10-myc	*Wurst et al 2012*	HgrR	pHD1700
pHD2276	LambdaN-RBP10 (full length)	*Wurst et al 2012*	HgrR	pHD1743
pHD2422	LambdaN-RBP10-F1-myc (positions 48-306a.a.)	cz4942; cz4943	HgrR	pHD2202
pHD2421	LambdaN-RBP10-F2-myc (positions 1-218a.a.)	cz4940; cz4941	HgrR	pHD2202
pHD2457	LambdaN-RBP10-F3-myc (positions 218-306a.a.)	cz5008; cz4943	HgrR	pHD2202
pHD2471	LambdaN-RBP10-F4-myc (positions 123-218a.a.)	cz5006; cz5007	HgrR	pHD2202
pHD2504	LambdaN-RBP10-F5-myc (positions 218-262a.a.)	cz5008; cz5085	HgrR	pHD2202
pHD2505	LambdaN-RBP10-F6-myc (positions 262-306a.a.)	cz5086; cz4943	HgrR	pHD2202
pHD2533	4E-IP-myc	cz5253; cz5063	HgrR	pHD1700
pHD2734	DRBD13-myc	*Bhaskar et. al 2015*	Hgr	pHD1700
pHD2813	Stem-loop against *4E-IP*	cz5644; cz5645	HgrR	pHD1146
Single knockout				
pHD2061	3' and 5' UTR RBP10 +Blasticidin resistance cassette	cz5180;cz5181 cz5182; cz5183	BlasR	pHD1748
In situ tagging				
pHD 2506	*In situ* N-TAP-RBP10	cz5128;cz5129 cz5130; cz5131	PuroR	pHD1959
pHD 2845	*In situ* N-BirA-RBP10	cz6197;cz6198	PuroR	pHD2506
pHD 1914	*In situ* V5-RBP10	cz4952;cz4953 cz4954; cz4955	BlasR	Bla-V5
pHD2393	*In situ* V5-RBP26	cz4952;cz4953 cz4954; cz4955	BlasR	Bla-V5
pHD2735	*In situ* V5-DRBD13	*Bhaskar et al 2015*	BlasR	Bla-V5
CAT reporter				
pHD1610	CAT + *EP1* 3'UTR (pT7-*CAT- EP1* 3'UTR)	*Schwede et al*	PuroR	*Schwede et al*
pHD1611	CAT +*EP1*Δ*26* 3'UTR (pT7-CAT-*EP1*Δ*26* 3'UTR)	*Schwede et al*	PuroR	*Schwede et al*
pHD 2277	CAT with *5 boxB + Actin 3'UTR*	*Delhi et al 2011*	NeoR	*Delhi et al 2011*

Plasmid No.	Description	Primers	Resistance marker	Plasmid backbone
Yeast two hybrid				
pHD2361	RBP10 ORF cloned in pGBKT7 (N-terminus myc)	cz4598;cz4599	KanR & TRP1	pGBKT7
pHD2368	RBP10 ORF cloned in pGADT7 (N-terminus HA)	cz4598;cz4599	AmpR & LEU2	pGADT7
pHD2585	RBP10-F3 cloned in pGBKT7 (N-terminus myc)	cz5223;cz5224	KanR & TRP2	pGBKT7
pHD2586	RBP10-F3 cloned in pGADT7 (N-terminus HA)	cz5223;cz5224	AmpR & LEU3	pGADT7
pHD2587	RBP10-F2 cloned in pGBKT7 (N-terminus myc)	cz5254;cz5255	KanR & TRP2	pGBKT7
pHD2588	RBP10-F2 cloned in pGADT7 (N-terminus HA)	cz5254;cz5255	AmpR & LEU3	pGADT7
pHD2490	4E-IP ORF cloned in pGBKT7 (N-terminus myc)	cz5048;cz5049	KanR & TRP3	pGBKT7
pHD2489	4E-IP ORF cloned in pGADT7 (N-terminus HA)	cz5048;cz5049	AmpR & LEU4	pGADT7
pHD2362	EF-1 alpha ORF cloned in pGBKT7 (N-terminus myc)	cz4602;cz4603	KanR & TRP3	pGBKT7
pHD2369	EF-1 alpha ORF cloned in pGADT7 (N-terminus HA)	cz4602;cz4603	KanR & TRP2	pGADT7
pHD2363	EF-2 ORF cloned in pGBKT7 (N-terminus myc)	cz4604;cz4605	AmpR & LEU3	pGBKT7
pHD2370	EF-2 ORF cloned in pGADT7 (N-terminus HA)	cz4604;cz4605	KanR & TRP3	pGADT7
pHD2364	eRF-1 ORF cloned in pGBKT7 (N-terminus myc)	cz4606;cz4607	AmpR & LEU4	pGBKT7
pHD2371	eRF-1 ORF cloned in pGADT7 (N-terminus HA)	cz4606;cz4607	KanR & TRP3	pGADT7
pHD2365	eRF-3 ORF cloned in pGBKT7 (N-terminus myc)	cz4608;cz4609	AmpR & LEU4	pGBKT7
pHD2372	eRF-3 ORF cloned in pGADT7 (N-terminus HA)	cz4608;cz4609	KanR & TRP4	pGADT7
pHD2366	eIF-5A ORF cloned in pGBKT7 (N-terminus myc)	cz4610;cz4611	AmpR & LEU5	pGBKT7
pHD2373	eIF-5A ORF cloned in pGADT7 (N-terminus HA)	cz4610;cz4611	KanR & TRP4	pGADT7
pHD2367	eIF-2B ORF cloned in pGBKT7 (N-terminus myc)	cz4612;cz4613	KanR & TRP3	pGBKT7
pHD2374	eIF-2B ORF cloned in pGADT7 (N-terminus HA)	cz4612;cz4613	AmpR & LEU4	pGADT7
pHD2394	RBP26 ORF cloned in pGBKT7 (N-terminus myc)		KanR & TRP1	pGBKT7
pHD2395	DRBD12 ORF cloned in pGBKT7 (N-terminus myc)		KanR & TRP1	pGBKT7
pHD2396	ZC3H22 ORF cloned in pGBKT7 (N-terminus myc)		KanR & TRP1	pGBKT7
pHD2397	Tb927.10.10050 ORF cloned in pGBKT7 (N-terminus myc)		KanR & TRP1	pGBKT7

Plasmid No.	Description	Primers	Resistance marker	Plasmid backbone
IPTG inducible expression				
pHD2618	pHD359 with Lac repressor + neomycin marker	cz5554; cz5555	NeoR	pHD 389
pHD2700	pHD359 with Lac repressor + blastcidin marker	cz5749; cz5750	BlasR	pHD 389
pHD2659	pHD 1034 with 2x Lac operator + PuroR from pHD1336	cz5582; cz5583	PurR	pHD 1034, 1336
pHD2661	pHD2659 + RBP10 (no tag)		PurR	pHD 2659

2.9.2 Oligonucleotides list

Oligo No.	Description	Restriction site	Sequence	Purpose
Cloning primers				
cz4598	Fwd-RBP10 ORF (pHD2361; pHD2368)	NdeI	TAGACTCATATGGGAGACTCGATATCACCT	Y2H-bait/prey (RBP10)
cz4599	Rev-RBP10 ORF (pHD2361; pHD2368)	EcoRI	TATAGAATTCTCACTCCATTCGAACCGGA	Y2H-bait/prey (RBP10)
cz5223	Fwd-RBP10 ORF-F3 (pHD2585; pHD2586)	NdeI	TATACATATGGGTGCAGGGGTCACACAA	Y2H-bait/prey (RBP10-F3)
cz5224	Rev-RBP10 ORF-F3 (pHD2585; pHD2586)	EcoRI	TATAGAATTCTCCCATTCGAACCGG	Y2H-bait/prey (RBP10-F3)
cz5254	Fwd-RBP10 ORF-F2 (pHD2587; pHD2588)	NdeI	TATACATATGATGGGAGACTCGATATCA	Y2H-bait/prey (RBP10-F2)
cz5255	Rev-RBP10 ORF-F2 (pHD2587; pHD2588)	EcoRI	TATAGAATTCGTGAATTGGCGCTTGCAT	Y2H-bait/prey (RBP10-F2)
cz4602	Fwd-EF1α ORF (pHD2362; pHD2369)	NdeI	TATACATATGGGAAAGGAAAAGGT	Y2H-bait/prey (EF1α)
cz4603	Rev-EF1α ORF (pHD2362; pHD2369)	EcoRI	TATAGAATTCTTATTTCTTCGAAGCCTTC	Y2H-bait/prey (EF1α)
cz4604	Fwd-EF2 ORF (pHD2363; pHD2370)	NdeI	TATACATATGGTCAACTTCACCGT	Y2H-bait/prey (EF2)
cz4605	Rev-EF2 ORF (pHD2363; pHD2370)	EcoRI	TATAGAATTCCTATAGCTTGTCCAGGAAG	Y2H-bait/prey (EF2)
cz4606	Fwd-eRF1 ORF (pHD2364; pHD2371)	NdeI	TATACATATGGCCGACCACGAGT	Y2H-bait/prey (eRF1)
cz4607	Rev-eRF1 ORF (pHD2364; pHD2371)	EcoRI	TATAGAATTCTTACATAAAGTCGTCGTCGAA	Y2H-bait/prey (eRF1)
cz4608	Fwd-eRF3 ORF (pHD2365; pHD2372)	NdeI	TATACATATGTCAGGTTGGGCAC	Y2H-bait/prey (eRF3)
cz4609	Rev-eRF3 ORF (pHD2365; pHD2372)	EcoRI	TATAGAATTCTTAAGCTTTACCTGCGTTC	Y2H-bait/prey (eRF3)
cz4610	Fwd-eIF5A ORF (pHD2366; pHD2373)	NdeI	TATACATATGTCTGACGATGAGGG	Y2H-bait/prey (eIF5A)
cz4611	Rev-eIF5A ORF (pHD2366; pHD2373)	EcoRI	TATAGAATTCTTATCGCTCAGCTGCAT	Y2H-bait/prey (eIF5A)
cz4612	Fwd-eIF2B ORF (pHD2367; pHD2374)	NdeI	TATACATATGCTTTCATGTTCCTGT	Y2H-bait/prey (eIF2B)
cz4613	Rev-eIF2B ORF (pHD2367; pHD2374	EcoRI	TATAGAATTCCTACTGGCATTTCCCC	Y2H-bait/prey (eIF2B)

Oligo No.	Description	Restriction site	Sequence	Purpose
Cloning primers				
cz5253	Fwd-4EIP ORF (pHD2533)	HpaI	TGACGTTAACATGAGAACTACAATTCGA	ectopic-4EIP-myc
cz5063	Rev-4EIP ORF (pHD2533)	BamHI	GACGGATCCGCGGCGCTGGTTGGG	ectopic-4EIP-myc
cz4942	Fwd-RBP10 ORF-F1 (pHD2422)	ApaI	TATAGGGCCCCGTAACGTTTATGTCTCT	tethering_F1
cz4943	Rev-RBP10 ORF-F1 (pHD2422)	BamHI	TATAGGATCCCTCCATTCGAACCGGAGG	tethering_F1
cz4940	Fwd-RBP10 ORF-F2 (pHD2421)	ApaI	TATAGGGCCCATGGGAGACTCGATATCA	tethering_F2
cz4941	Rev-RBP10 ORF-F2 (pHD2421)	BamHI	TATAGGATCCGTGAATTGGCGCTTGCAT	tethering_F2
cz5008	Fwd-RBP10 ORF F3 (pHD2457)	ApaI	TATAGGGCCCGGTGCAGGGGTCACACAA	tethering_F3
cz5006	Fwd-RBP10 ORF F4 (pHD2471)	ApaI	TATAGGGCCCATAGCAGACGCCGCTGCA	tethering_F4
cz5007	Rev-RBP10 ORF F4 (pHD2471)	BamHI	TATAGGATCCGTGAATTGGCGCTTGCAT	tethering_F4
cz5085	Rev-RBP10 ORF F5 (pHD2504)	BamHI	TATAGGATCCAGTCACAAATAGAGGCAT	tethering_F5
cz5086	Fwd-RBP10 ORF F6 (pHD2505)	ApaI	TATAGGGCCCATGCCACAAACCGCGCCC	tethering_F6
cz5128	Fwd-RBP10-5'UTR (pHD2506)	SacI	TATAGAGCTCATCTTGGTGTTCCATACG	*insitu*_N_TAP_RBP10
cz5129	Rev-RBP10-5'UTR (pHD2506)	NdeI	TATACATATGCTTACAGATGTTTGTCAA	*insitu*_N_TAP_RBP10
cz5130	Fwd-RBP10-miniORF (pHD2506)	HindIII	TATCAAGCTTCCATGGGAGACTCGATATCA	*insitu*_N_TAP_RBP10
cz5131	Rev-RBP10-miniORF (pHD2506)	ApaI	TATAGGGCCCCGGCCGCCCATAAGCACT	*insitu*_N_TAP_RBP10
cz5180	Fwd-RBP10-3'UTR (pHD2061)	EcoRI	TATAGAATTCTGGCACAGAGGGTAACGA	SKO_rbp10-3'
cz5181	Rev-RBP10-3'UTR (pHD2061)	BamHI	TATAGGATCCGGCCTGTCTTCCTTCCTC	SKO_rbp10-3'
cz5182	Fwd-RBP10-5'UTR (pHD2061)	XhoI	TATACTCGAGTACCGAATCTGACCTTTC	SKO_rbp10-5'
cz5183	Rev-RBP10-5'UTR (pHD2061)	HindIII	TATTAAGCTTCTTACAGATGTTTGTCAA	SKO_rbp10-5'
cz4952	Fwd-RBP10-miniORF (pHD1914)	XhoI	TATACTCGAGATGGGAGACTCGATATCA	*insitu*_v5_RBP10
cz4953	Rev-RBP10-miniORF (pHD1914)	ApaI	TATAGGGCCCTTAAGTCCTTCGATGCAC	*insitu*_v5_RBP10
cz4954	Fwd-RBP10-5'UTR (pHD1914)	SacII	TATACCGCGGATCTTGGTGTTCCATACG	*insitu*_v5_RBP10
cz4955	Rev-RBP10-5'UTR (pHD1914)	XbaI	TATATCTAGACTTACAGATGTTTGTCAA	*insitu*_v5_RBP10
Cz4804	Fwd-RBP26-miniORF (pHD1914)	XhoI	TATACTCGAGATGGAAACGCACCAGGAT	*insitu*_v5_RBP26
Cz4805	Rev-RBP26-miniORF (pHD1914)	ApaI	TATAGGGCCCAGTGGCTTGCGTGACACG	*insitu*_v5_RBP26
Cz4802	Fwd-RBP26-5'UTR (pHD1914)	SacII	TATACCGCGGTGTCCTGCTTGTGTGGCG	*insitu*_v5_RBP26
Cz4803	Rev-RBP26-5'UTR (pHD1914)	XbaI	TATATCTAGAGACTAAGGGAGGAGAATC	*insitu*_v5_RBP26

Oligo No.	Description	Restriction site	Sequence	Purpose
Cloning primers				
cz5554	Fwd-Neomycin ORF	SpeI	TATAACTAGTGATGGTGGAACAAGATGGATT	subcloning neomycin+ *aldolase*_3' UTR
cz5555	Rev-*Aldolase* 3' UTR	StuI	CCTTCGAATCCCCCCATTTTCTT	subcloning neomycin+ aldolase_3' UTR
cz5749	Fwd-Blasticidin ORF	SpeI	TATCACTAGTATGGCCAAGCCTTTGTCTC	subcloning blasticidin+ *actin*_3' UTR
cz5750	Rev-*Actin* 3' UTR	StuI	TATAAGGCCTTGCAGAATACTGCATAGAT	subcloning blasticidin+ *actin*_3' UTR
cz5582	2x Lac operator forward oligo	SmaI blunted	GGAATTGTGAGCGGATAACAATTCGGAATT-GTGAGCGGATAACAATTCCC	Lac operator 2 copies
cz5583	2x Lac operator complementary oligo	SmaI blunted	CCTTAACACTCGCCTATTGTTAAGCCTTAA-CACTCGCCTATTGTTAAGGG	Lac operator 2 copies (complementary)
cz6197	Fwd-BirA ORF	XhoI	TATACTCGAGCCATGAAGGATAATACTGTTCC	*insitu*_BirA_RBP10
cz8198	Rev-BirA ORF	HindIII	TATTAAGCTTGAGCAGCAGCAGAAATTTGAAC	*insitu*_BirA_RBP10
Primers for Northern blot probes				
cz4615	*CAT* probe Fwd		ATGGAGAAAAAAATCACTGGATAT	*CAT* probe
cz2689	*CAT* probe Rev		GAAAGACGGTGAGCTGGT	*CAT* probe
cz2724	*tubulin* probe Fwd		TGACTCGCCGCAACCTCGAT	*Tubulin* probe
cz2581	*tubulin* probe Rev		CCTTTGGCACAACGTCACCACGG	*Tubulin* probe
cz4490	spliced leader oligo (antisense)		CAATATAGTACAGAAACTGTTCTAATAATA-GCGTTAGTT	splice leader probe
Semi-quantitative RT-PCR primers				
cz5725	*spliced leader* Fwd		ACGCTATTATTAGAACAGTTTCTGTAC	RT-PCR
cz689	*CAT* ORF Rev		GAAAGACGGTGAGCTGGT	RT-PCR
cz6168	*tubulin* ORF Rev		CAGCCTGACCAATGTGGATGCAGAT	RT-PCR
cz6308	*spliced leader* Fwd		GACTAGTTTCTGTACTAT	RT-PCR
cz6309	All *VSGs* Rev		CCGGGTACCGTGTTAAAATATATC	RT-PCR

3 Results

3.1 RBP10 C-terminus promotes reporter mRNA destruction and inhibits translation

To determine the effect that RBP10 has on bound mRNAs, the lambda N peptide was used to artificially tether RBP10 to a *CAT* reporter bearing five *boxB* recognition sites and the *actin 3'UTR* (Figure 3.1 A-i). As a control, we used a cell line lacking the fusion protein. Reduction in CAT protein was observed in some of the uninduced samples, probably because of leaky expression of the fusion proteins (Figure 3.1 A-ii, C). As previously reported [91], full length RBP10 when tethered repressed the CAT reporter expression; the amount of protein was ~90% reduced, *CAT* mRNA abundance was slightly less affected with ~40% being detected. Surprisingly, untagged RBP10 was reduced after 24 hours of λN-RBP10 expression (Figure 3.1 A-ii); cell growth was also affected.

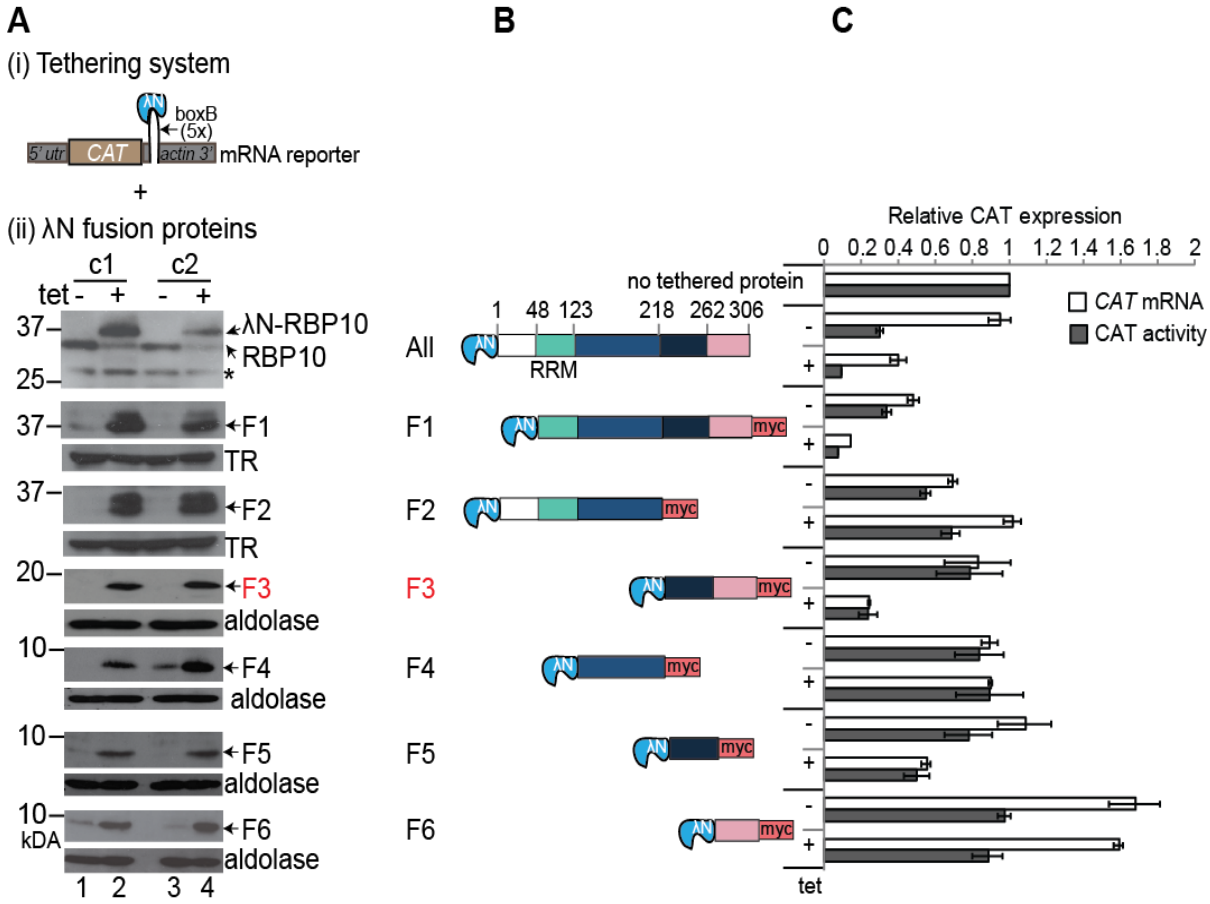

Figure 3.1 RBP10 tethering in bloodstream form trypanosomes. **A)** Part (i) the λN/boxB interaction system used for tethering. The CAT reporter was constitutively expressed, and had 5 copies of *boxB* sequences (for simplicity only one is shown)

upstream of the *actin 3' UTR*. (ii) Protein levels of the λN fused RBP10 or 6 different fragments (F1-F6) of RBP10; expression was tetracycline induced (+ tet for 24h) and the fragments had C-terminus myc tags. Two independent clones (c1, c2) are shown. Aldolase or TR (trypanothione reductase) is the loading conrol; * is non-specific band. **B)** Schematic representations of RBP10 fragments tested; the regions are colour coded and amino acid positions are shown on the full length RBP10 (all). **C)** The relative *CAT* mRNA and protein levels as determined by Northern blot and CAT assay respectively. Results of at least three replicates are shown; the error bar corresponds to the standard deviation. Cells without tethered protein served as control. The minimum fragment (F3) necessary for the repressor effect of RBP10 is highlighted in red.

Using the same approach, six different fragments of RBP10 that were C-terminally myc tagged were tested. The aim was to map RBP10 region necessary for its inhibitory effect. RBP10 N-terminus (Figure 3.1 B, C; F2, F4) including the RRM motif failed to repress CAT expression. The low complexity C-terminal part (Figure 3.1 A-ii; F3) was identified as the minimum region with the repressive effect (Figure 3.1 B, C; F3). In this case, both *CAT* mRNA and protein levels were reduced by ~80%. The repressor effect was lost when this region was deleted (Figure 3.1 B, C; F2), and fragment (Figure 3.1 B, C; F1) retaining the C-terminal part (highlighted in red, Figure 3.1 A, B; F3) as expected remains repressive. An attempt to narrow down further on the F3 fragment was not successful; when divided by half (Figure 3.1 B, C; F6, F5), only F5 retained ~50% of the repressor effect while F6 had an opposing effect with a slightly positive effect on the *CAT* mRNA.

3.2 Tethered RBP10 shifts reporter mRNA from the polysomes

To understand in more detail how RBP10 acts as a repressor, the association of the *CAT* reporter mRNA with ribosomes was examined with and without tethered RBP10 (C-terminus fragment; F3). After expression of λN-RBP10-F3, polysome fractionation was performed using extracts from the cells grown with (Figure 3.2 A-C) or without (Figure 3.2 D-E) tetracycline. Total RNA was isolated from each fraction and the distribution of both *CAT* and α-*tubulin* mRNAs analyzed by Northern blotting (Figure 3.2 B, E). In absence of λN-RBP10-F3 the *CAT* mRNA co-migrated with the polysomes (Figure 3.2 B, C). Tethering of RBP10-F3 led to decreased levels of the *CAT* mRNA (Figure 3.2 E, F). In addition, <10% *CAT* mRNA was observed in the fractions containing translating polyribosomes; most of the quantifiable *CAT* mRNA

co-migrated with the fractions where normally ribosome subunits are found (Figure 3.2 E, F). As a control, expression of λN-RBP10-F3 had no effect on either the amount or distribution of α *tubulin* mRNA (Figure 3.2 B, E). Therefore, RBP10 C-terminus prevents translation initiation and promotes mRNA destruction.

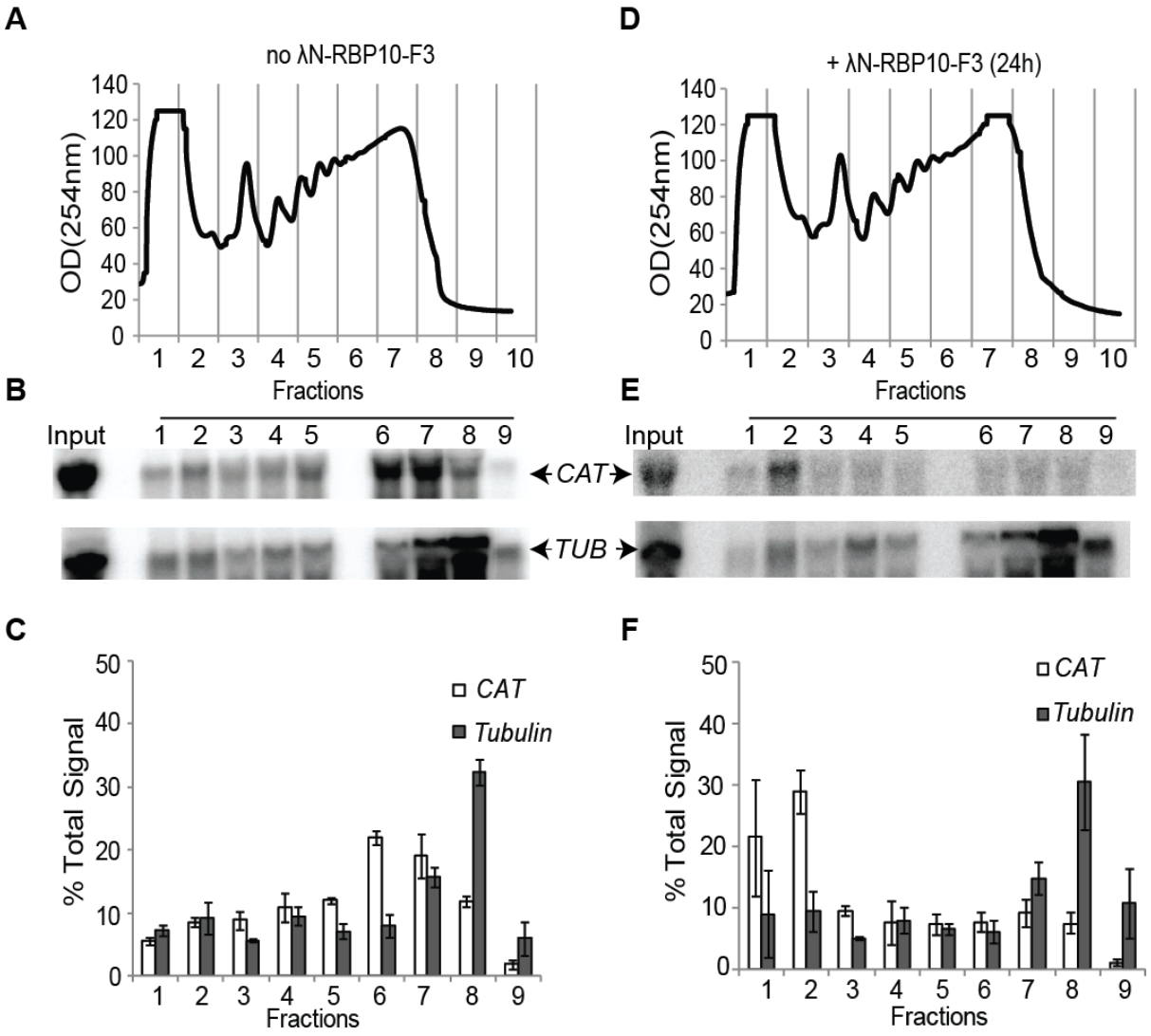

Figure 3.2 Translation repression after tethering RBP10 on reporter mRNA. **(A & D)** An OD254 trace of sucrose gradient (15-50%) of cells with or without λN-RBP10-F3. **(B & E)** Northern blot analysis of the *CAT* and *tubulin* mRNAs isolated from equal amounts of each fraction. One of the three replica blot used for quantification is shown. **(C &F)** Quantification of *CAT* and *tubulin* mRNAs from three replicates.

3.3 Identification of RBP10 interacting proteins by yeast-two hybrid

RBP10 acts as a negative regulator when tethered to reporter mRNA. To test whether the inhibition of the reporter translation by RBP10 occurs via interaction with the translation factors, yeast-two hybrid assay were performed using a few selected translation factors (Figure 3.3). RBP10 failed to interact with EF1-α, EF2, eRF1, eRF3, eIF-5A and eIF-2B in a pairwise assays (Figure 3.3 A). The expression of all the proteins was confirmed (Figure 3.3 B) apart from eRF1 which was undetectable.

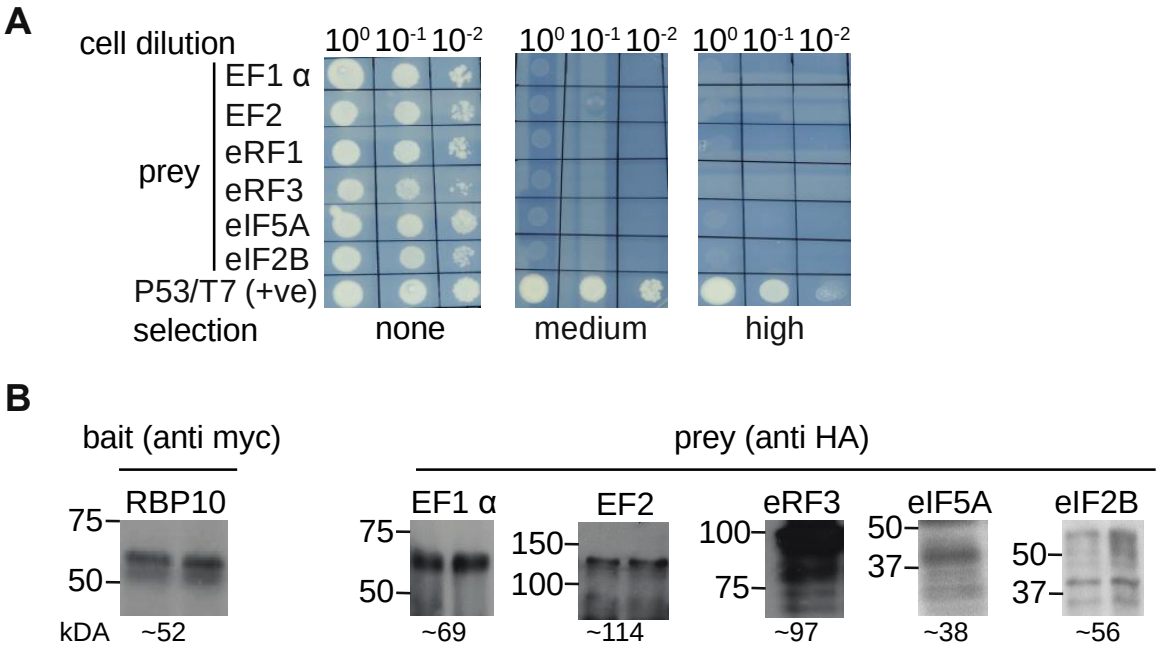

Figure 3.3 Pairwise yeast two-hybrid screen between RBP10 and translation factors. **A)** Plates showing interactions between RBP10 as bait and the translation factors EF1 α, EF2, eRF1 eRF3, eIF5A, eIF2B as the prey. Cells diluted 1:10 or 1:100 were grown on medium and high stringency nutrient selection media. P53 and SV40 T-antigen interaction served as positive control. **B)** Bait and prey proteins expression were detected using anti-myc and anti-HA antibodies respectively.

3.3.1 High throughput yeast two-hybrid screen using RBP10 as bait

To identify RBP10 interactors through an unbiased approach, yeast cells expressing RBP10 as bait were transformed with a prey library (prepared by Dr. Esteban Erben) generated from random genomic DNA fragments of *T. brucei*. ~1.5 x10^6 independent clones were obtained (genome coverage ~7 folds), with ~1x10^5 expected to contain in frame fusion proteins. This complexity was reduced after selection (Figure 3.3.1 A)

for positive interactions; ~1500 colonies survived on the high stringency plates. 50 colonies were randomly selected, and after plasmid rescue the inserts were confirmed by *Xho I* digestion (Figure 3.3.1 B) followed by individual plasmid DNA sequencing. The rest of the colonies were pooled and prey plasmids DNA isolated. After plasmid specific PCR amplification, the complexity of the inserts detected as a smear on agarose gel (Figure 3.3.1 C) was analyzed by high throughput DNA sequencing. The inserts from the input library (9302 ORFs being represented) served as the control.

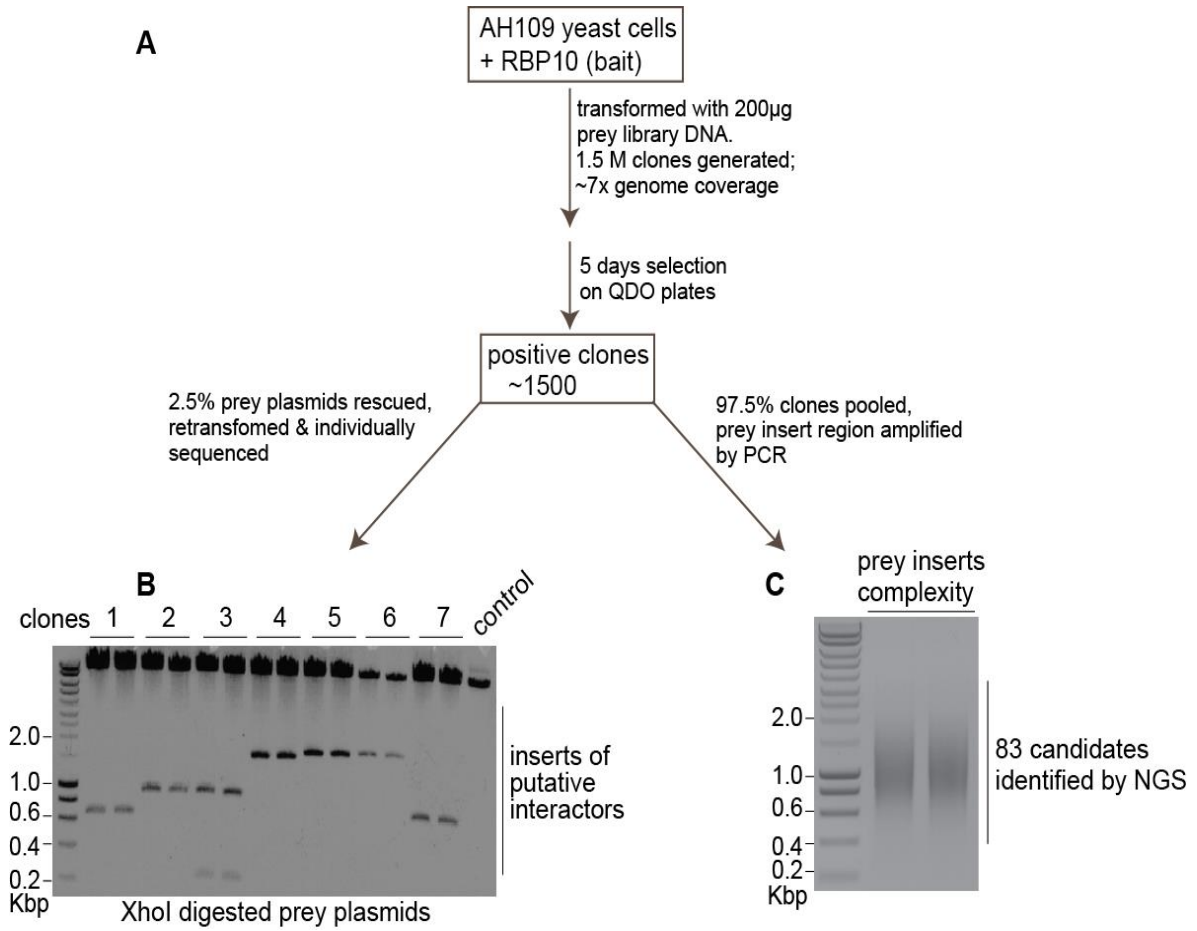

Figure 3.3.1. Genome wide yeast two-hybrid screen with RBP10 as bait. **A)** Number of independent clones obtained before and after selection on quadruple drop out (QDO) plates. **B)** Inserts confirmation from the rescued prey plasmids by XhoI digest. The inserts (2.5% of the selected clones) were identified by Sanger DNA sequencing. Seven clones in duplicate are shown; the control is an empty prey plasmid. **C)** Distribution of the DNA smear after prey plasmid specific PCR of the pooled clones. The inserts were identified using high throughput DNA sequencing (NGS).

83 ORFs with at least 10 sequence reads and more than one fragment representation were obtained. All the candidates identified by individual sequencing were confirmed in the genome-wide approach. Since RBP10 negatively regulates reporter mRNA when tethered, repressor partners of RBP10 with more than one positive fragment and at least 5 sequence reads were considered. To qualify as a repressor, a cut-off of 1.7 fold was used according to [82, 83] datasets. Based on this criterion 25 potential partners were found (Table 1), 16 of these (Table 1 A) had more than one positive fragment; 9 more candidates (Table 1 B) that met two of the requirements but had only one positive fragment are also included. Among the 25 potential partners of RBP10, 8 of them (Table 1 highlighted in grey) are expressed in bloodstream form trypanosomes and binds mRNAs *in vivo* [82]. Comparative analyses using datasets from other similar screens of MKT1, NOT2, and eIF4E-IP (4E-IP) helped to identify common partners. RBP10 came up in the 4E-IP dataset (Table 1 A, B) and interestingly, one 4E-IP fragment (Table 1 B) was found in the RBP10 screen. Also, the RNA binding proteins RBP9, DRBD7, ZC3H14 and one hypothetical protein (Tb927.10.10050) are co-shared between RBP10 and 4E-IP. Surprisingly, RBP10 interacts with itself since 7 positive fragments (Table 1 A) of RBP10 were identified.

Table 1. Potential partners of RBP10 as identified by genome wide yeast two-hybrid screen. **A)** The list shows proteins with >5 reads that had at least two positive fragments and are known [82, 83] as negative regulators when tethered to a reporter mRNA; the ones that met two of these requirements **(B)** are included as well. A 1.7 fold change cut-off for the tethering effect was considered for the studies done using either the short gun library [83] or a complete ORF library of selected proteins [82]. For the RBP10 two-hybrid screen, the numbers of positive fragments and the maximum number of reads are shown. Also in the table are the common partners identified in similar screens with 4E-IP, NOT2 and MKT1; only the maximum number of reads are shown. Highlighted in grey are those known to bind poly A mRNA in bloodstream form trypanosomes according to the findings from [82]. nd = no data; Y2H = yeast two hybrid

Table 1. A

		RBP10 Y2H		Other Y2H screens			*Esteban et.al 2014*	*Lueong et.al 2016*	
Protein ID	**Annotation**	**No. of positive fragments**	**RBP10 (max. reads)**	**4E IP (max. reads)**	**NOT2 (max. reads)**	**MKT1 (max. reads)**	**"Shotgun library" tethering effect**	**"ORFeome library" tethering effect**	**mRNA binder? (pvalue)**
Tb927.3.3940	DRBD11	15	6012	nd	nd	nd	3	6.2	yes (0)
Tb927.8.2780	RBP10	7	1606	205	0	nd	2.6	2.5	yes (0)
Tb927.4.400	DRBD7	5	232	227	0	nd	3.4	6.9	yes (0)
Tb927.11.14220	hypothetical protein	2	31	nd	nd	nd	3.9	nd	yes (0)
Tb927.5.1580	ZC3H13	14	2340	nd	nd	nd	3.2	7.3	yes (0.01)
Tb927.6.5010	hypothetical protein	7	1473	nd	nd	nd	5.6	8.4	yes (0.01)
Tb927.8.910	hypothetical protein	21	7898	nd	nd	nd	3.9	6	unclear (0.05)
Tb927.10.10050	hypothetical protein	17	4960	1129	1	nd	2.1	nd	no (0.1)
Tb927.3.3060	hypothetical protein	5	150	nd	nd	nd	3.2	6	no (0.09)
Tb927.7.2680	ZC3H22	17	10369	nd	nd	nd	2.7	4.9	nd
Tb927.6.4720	ZC3H15	10	2234	nd	nd	nd	4.4	1.5	nd
Tb927.1.1470	hypothetical protein	7	890	nd	nd	nd	3.9	1.5	nd
Tb927.6.4050	ZC3H14	7	2044	897	0	3	nd	4.5	nd
Tb927.7.5380	DRBD12	2	3201	nd	nd	nd	2.1	2.5	nd
Tb927.9.13970	hypothetical protein	2	57	0	0	0	nd	6.7	nd
Tb927.8.6080	hypothetical protein	2	245	nd	nd	nd	1.7	nd	nd

Table 1. B

Repressors, >10 reads but <2 positive fragments									
		RBP10 Y2H		Other Y2H screens			*Esteban et.al 2014*	*Lueong et.al 2016*	
Protein ID	**Annotation**	**No. of positive fragments**	**RBP10 (max. reads)**	**4E IP (max. reads)**	**NOT2 (max. reads)**	**MKT1 (max. reads)**	**"Shotgun library" tethering effect**	**"ORFeome library" tethering effect**	**RNA binder? (pvalue)**
Tb927.10.5250	ZC3H32	1	29	0	0	120196	nd	3.2	yes (0)
Tb927.9.11050	4E-IP	1	14	nd	nd	nd	4.49	17.2	yes (0.01)
Tb927.11.12120	RBP9	1	34	1176	1	nd	10.52	4.6	unclear (0.05)
Tb927.3.1920	NOT5	1	340	nd	nd	nd	2.35	11.9	unclear (0.05)
Tb927.8.7820	Hypothetical protein	1	530	nd	nd	71	2.31	nd	no (0.06)
Tb927.7.2980	Hypothetical protein	1	35	nd	nd	nd	1.79	0.3	nd
Tb927.10.12020	Hypothetical protein	1	367	nd	nd	nd	1.72	0.8	nd
Tb927.9.10280	ZC3H48	1	1693	0	0	1	nd	2.4	nd
Tb927.9.9840	lipoic acid containing carrier protein	1	1232	nd	nd	nd	5.59	nd	nd

3.4 Genome wide yeast two-hybrid data confirmation using complete ORFs

Using full length RBP26, ZC3H22, DRBD12, Tb927.10.10050 and 4E-IP ORFs, the interaction with RBP10 was confirmed (Figures 3.4.1-2) by pairwise yeast two-hybrid. A screen using SV40 T-antigen or Lamin as controls identified ZC3H22, DRBD12 and Tb927.10.10050 as auto activating only when used as baits.

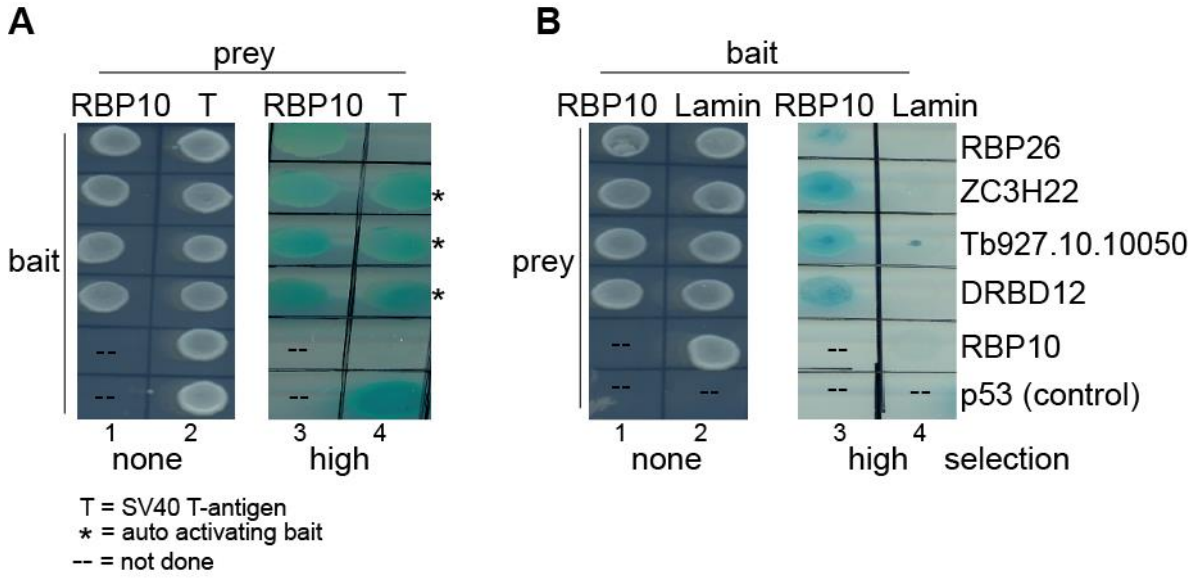

Figure 3.4.1 Confirmation of the genome wide yeast two-hybrid data using full-length prey inserts. **A)** Plates showing pairwise interactions of RBP10 as prey and RBP26, ZC3H22, Tb927.10.10050, DRBD12 as baits. As a negative control a screen using SV40 T-antigen (T) as prey is shown (lanes 2 & 4); the asterisk (*) highlights the auto activating baits. P53/T interaction served as positive control. The cells were selected on quadruple dropout medium (QDO); blue colonies highlight addition confirmation by alpha-galactosidase assay. **B)** Reciprocal interaction. As in (A), but using RBP10 as bait. Screening using Lamin (lanes 2 & 4) as bait served as a negative control.

The interaction between RBP10 and 4E-IP in the yeast two-hybrid (Figure 3.4.2) requires the C-terminal part of RBP10 (Figure 3.4.2 A, B; F3) that has the repressor effect. RBP10 fragment lacking the C-terminal region failed to interact with 4E-IP (Figure 3.4.2 A, B F2). Also, shown in Figure 3.4.2 A (last plate set) is the confirmation that RBP10 can self-interact as identified in the RBP10 genome wide two-hybrid screen (Table 1 A).

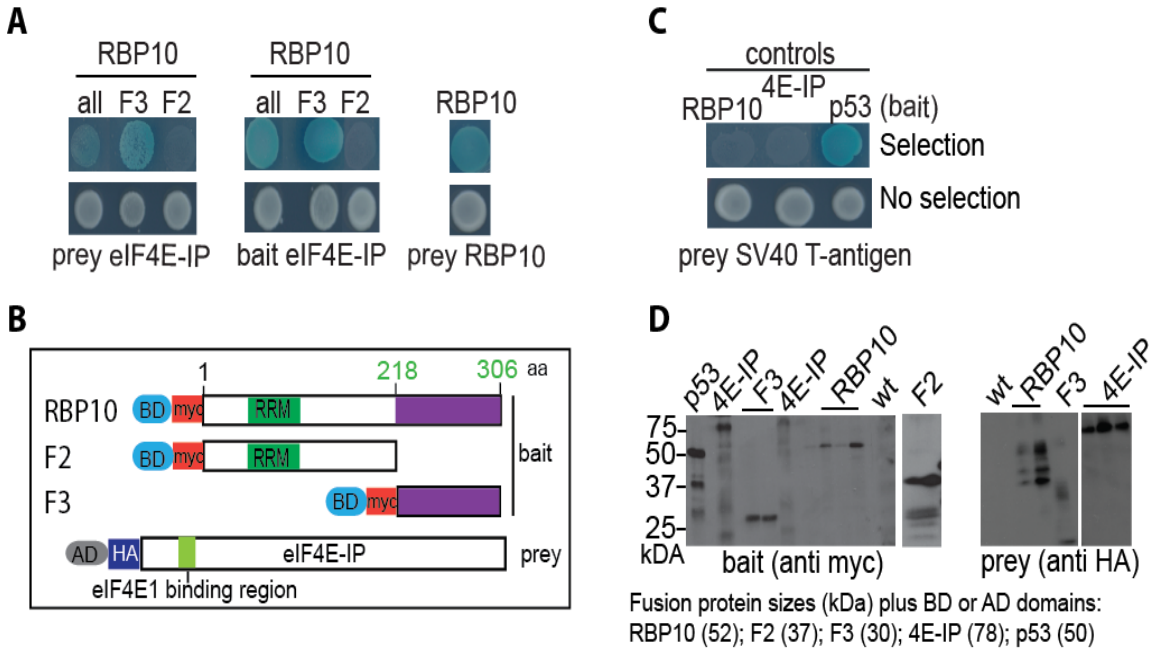

Figure 3.4.2 Interaction of RBP10 and eIF4E-IP in yeast. **A)** Plates showing pairwise interactions between eIF4E-IP and full length RBP10 or the two fragments of RBP10 F3 and F2. RBP10 self-interaction is also shown. The cells were selected on quadruple dropout medium (QDO); blue colonies highlight addition confirmation by alpha-galactosidase assay. **B)** Schematic representation of the RBP10 fragments checked for interaction with eIF4E-IP. **C)** Controls to check for auto activating baits. P53 and SV40 T-antigen (T) interaction served as positive control. **D)** Bait and prey proteins expression were detected using anti-myc and anti-HA antibodies respectively. BD is yeast GAL4 DNA binding domain; AD is the Activation domain.

3.5 RBP10, RBP26 and 4E-IP interactions in bloodstream form trypanosomes

Having confirmed that RBP10 interacts with RBP26 and 4E-IP in yeast-two hybrid (Figure 3.4.1-2), I wanted to find out whether this is true as well in trypanosomes. No follow up studies of ZC3H22, DRBD12 and Tb927.10.10050 were done in trypanosomes because all were detected as auto activating the yeast two-hybrid system when used as baits (Figure 3.4.1 A). Furthermore, the biological relevance for the interactions between RBP10, ZC3H22 and DRBD12 is unclear because these two candidates are not expressed in bloodstream form trypanosomes.

In case of RBP26, extracts from bloodstream form cells expressing in situ V5-RBP26 and inducible RBP10-myc was used for co-immunoprecipitation with α-myc beads. The cell line expressing only V5-RBP26 was used as control (Figure 3.5.1 A; lanes 1-3). After precipitating RBP10-myc, V5-RBP26 was not preferentially co-purified; only a very faint band was detected (Figure 3.5.1 A; lane 6). However, only half of

the eluate was loaded since both tagged proteins have similar sizes. V5-RBP26 did not bind non-specifically to the beads as observed in the control (Figure 3.5.1 A; lane 3). Therefore, the co-purification depended on the presence of RBP10, however it is a very weak interaction.

The reciprocal co-immunoprecipitation using α-V5 beads was also checked. After precipitating V5-RBP26, co-precipitation of RBP10-myc was observed, and again only a weak signal was detected (Figure 3.5.1 B; lane 3). Whether RBP10-myc bound to the beads non-specifically was not tested in this case.

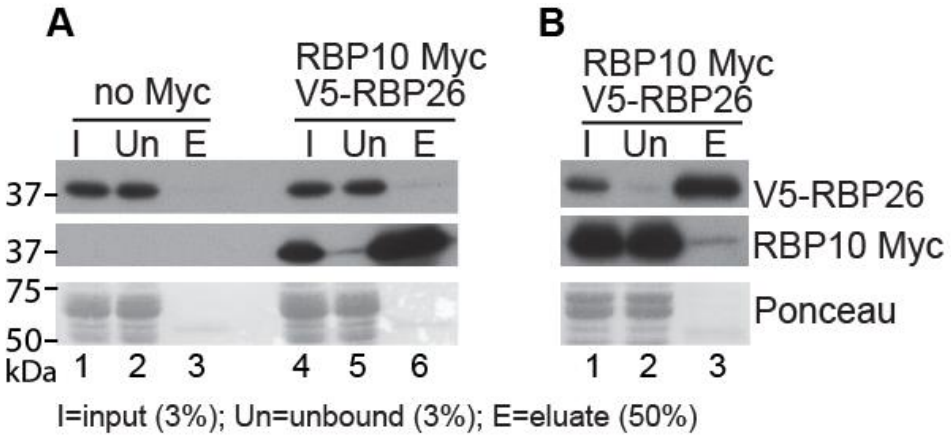

Figure 3.5.1 Interaction of RBP10 and RBP26 in bloodstream form trypanosomes. **A)** Co-immunoprecipitation with α-myc beads using extracts from cells expressing V5-RBP26, with (lanes 4-6) or without (lanes 1-3) additional expression of RBP10 myc. The precipitated proteins were detected by Western blotting using anti-myc, and anti-V5 antibodies. Ponceau staining served as a loading control. I: input, Un: unbound (both 3% of the lysate) and E: eluate (50% of the boiled beads in sample buffer). **B)** Reciprocal co-immunoprecipitation using α-V5 beads.

For 4E-IP, co-immunoprecipitation with α-myc beads was done using extracts from bloodstream form cells expressing in situ V5-RBP10 and inducible 4E-IP-myc (Figure 3.5.2 A; lanes 1-3). As control, the cell line expressing only V5-RBP10 was used (Figure 3.5.2 A; lanes 4-6). Precipitating 4E-IP-myc resulted in co-purification of V5-RBP10 (Figure 3.5.2 A; lane 3). Lack of substantial V5-RBP10 co-purification in the control (Figure 3.5.2 A; lanes 6) indicated that RBP10 pull-down depended on the presence of 4E-IP-myc; this confirms the yeast two-hybrid data using full length RBP10 (Figure 3.4.2 A). Next, to verify whether RBP10/4E-IP interaction in trypanosomes occurs via the C-terminal region of RBP10 as observed in the pairwise yeast two-hybrid (Figure 3.3.4 A), co-immunopreciptation with α-V5 beads

was done using extracts from bloodstream form cells expressing an inducible C-terminus myc tagged RBP10 fragment (RBP10-F3-myc) and in situ V5-4E-IP (Figure 3.5.2 B). Precipitating V5-4E-IP failed to co-precipitate the C-terminus RBP10 fragment, since RBP10-F3-myc was not detected in the eluate fraction from cells extracts having both V5-4E-IP and RBP10-F3 fragment (Figure 3.5.2 B; lanes 1-3) as similarly observed in the control (Figure 3.5.2 B; lanes 4-6). Therefore, this part of two-hybrid data could not be confirmed in trypanosomes.

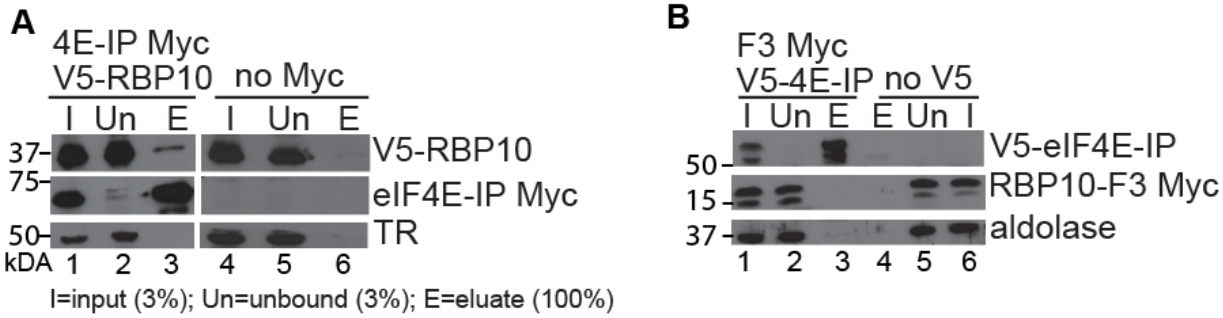

Figure 3.5.2 Interactions of eIF4E-IP and RBP10 in bloodstream form trypanosomes. **A)** With full length RBP10. Co-immunoprecipitation with α-myc beads using extracts from cells expressing V5-RBP10, with (lanes 1-3) or without (lanes 4-6) additional expression of eIF4E-IP-myc. The precipitated proteins were detected by Western blotting using anti-myc, anti-V5 and as control anti-trypanothione reductase (TR). I: input, Un: unbound (both 3% of the lysate) and E: eluate (100% of the boiled beads in sample buffer). **B)** With C-terminus RBP10 fragment (RBP10-F3-myc). Co-immunoprecipitation with α-V5 beads using extracts from cells expressing RBP10-F3-myc, with (lanes 1-3) or without (lanes 4-6) V5-eIF4E-IP. The loading and detection as in (A), but with aldolase as the loading control.

3.6 RBP10 and 4E-IP interdependency in procyclic forms

3.6.1 Independent dual induction system using Tet and IPTG in procyclic forms

RBP10 when over expressed in procyclics is lethal [91]. To investigate whether the effects of RBP10 in procyclic forms is dependent on 4E-IP levels, 4E-IP expression was knocked down prior to RBP10-myc over expression. To do this, the expression of RBP10-myc and depletion of 4E-IP had to be done independently. This was achieved by using a dual induction system where RBP10-myc was IPTG induced while *4E-IP* RNAi was induced using tetracycline.

For the IPTG inducible system [154], transient expression of the *CAT* reporter in the procyclic forms was tested initially since the system is not routinely used in our lab.

The presence of two lac operators downstream of the rRNA promoter allowed IPTG inducible expression of the CAT protein from an episome (Figure 3.6.1 A-B); the cells constitutively expressed the lac repressor (LacR) and were grown in presence of 1 mM IPTG for 15 hours. The uninduced cells, cells lacking the lac repressor and the ones transfected with a reporter without the lac operator served as controls (Figure 3.6.1 A-B). This result was further confirmed using stable procyclic cells lines; unlike in the transient expression (Figure 3.6.1 B), leaky expression was observed with ~ 30% CAT activity detected in absence of IPTG for the most tightly regulated clone (Figure 3.6.1 C).

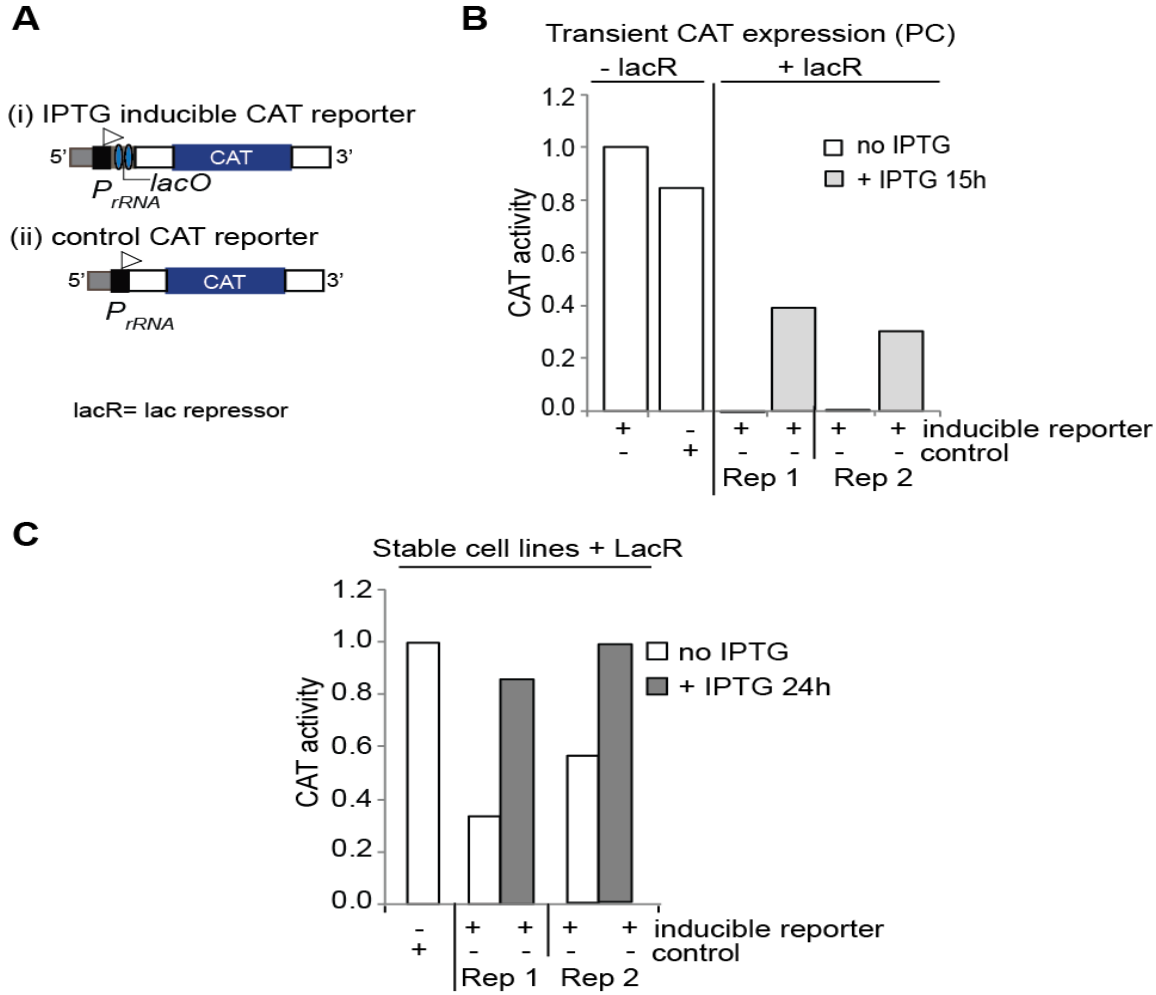

Figure 3.6.1 IPTG inducible system in procyclic trypanosomes. **A)** Schematic representation of the CAT reporters used; i) IPTG inducible construct with 2 lac operators, ii) The control reporter. Both integrate at the silent rRNA locus, bears rRNA promoter and have a puromycin selection marker. **B)** IPTG induced (1 mM IPTG for 15h) transient CAT expression in procyclic cells constitutively expressing the lac repressor (+lacR); cells lacking the lac repressor (-lacR) served as control. CAT protein levels were measured using CAT assay. **C)** CAT expression in stable procyclic cell lines constitutively expressing the lac repressor (+lacR). CAT protein was induced for 24 hours using 1 mM IPTG and then quantified by CAT assay.

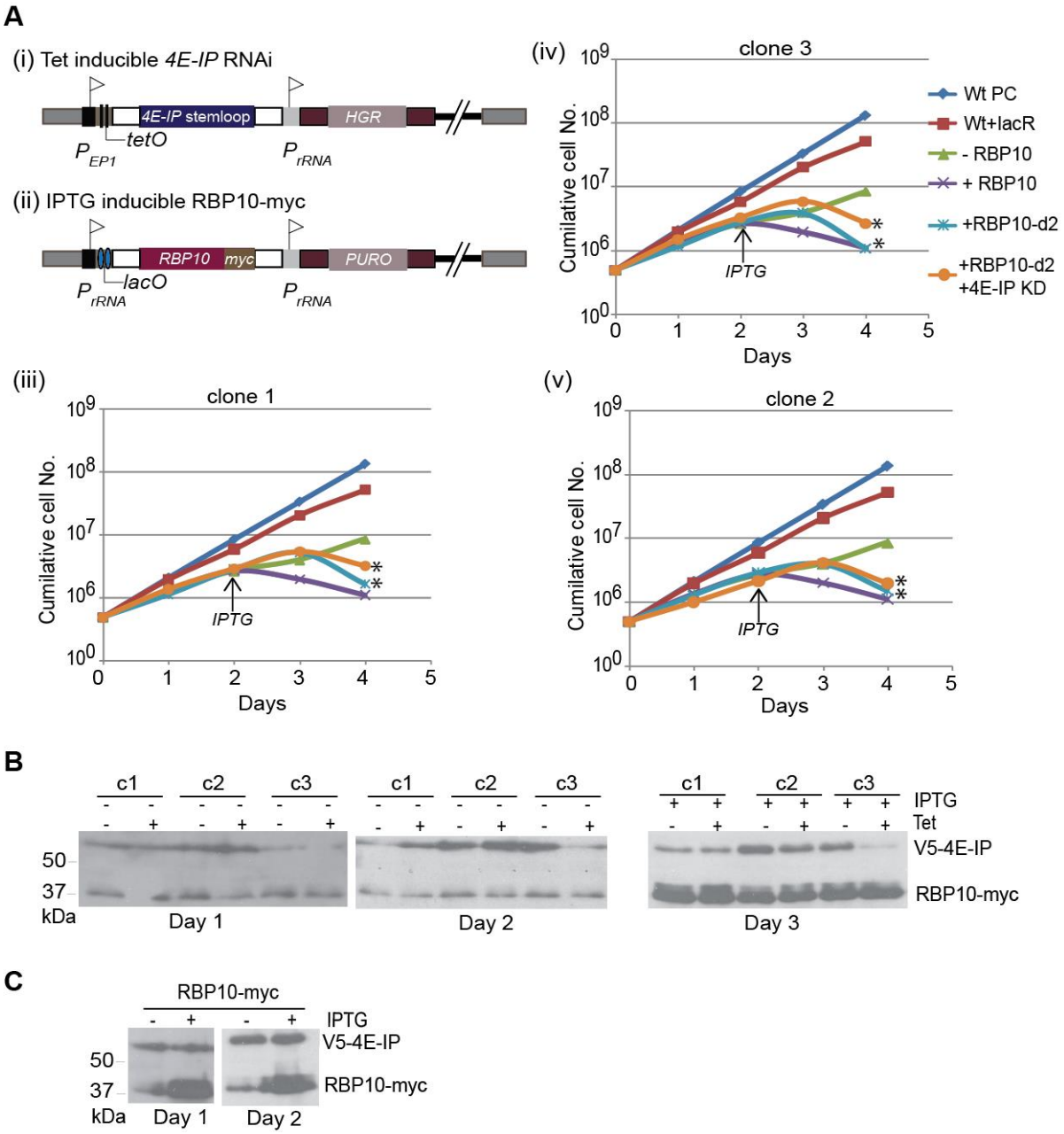

Figure 3.6.2 Independent dual induction of RBP10 expression and eIF4E-IP RNAi in procyclic cells. **A)** From (i-ii) schematic representation of the tet inducible stemloop construct targeting eIF4E-IP, and IPTG inducible RBP10-myc expression construct. (iii-v) the growth effect of over expressing RBP10 in the background of eIF4E-IP knockdown. Induction was through out for *4E-IP* RNAi; from day 2 RBP10-myc was induced. The asterisk (*) highlights the data from the individual clones; the rest of the growth curves (controls) are the same for all the plots. Wild type procyclic cells with (Wt+laR) or without (Wt PC) lac repressor and the uninduced cells served as control. Data from three independent clones is shown. **B)** Western blot showing the expression of V5-eIF4E-IP and RBP10-myc using extracts from cells (clones 1-3) where 4E-IP was knockdown (from day 1-3) and RBP10-myc expression induced from day 2. **C)** Western blot showing expression of RBP10-myc and V5-eIF4E-IP in cells lacking stemloop construct against eIF4E-IP.

Independent induction of *4E-IP* RNAi (Tet) and RBP10-myc (IPTG) expression in procyclic cells was possible as shown on Figure 3.6.2. In comparison to the wild type cells, expression of the lac repressor caused a mild growth defect (Figure 3.6.2 A-i). *4E-IP* RNAi was induced for two days prior to RBP10-myc expression. During the two days when cells were grown in absence of IPTG, leaky expression of RBP10-myc was observed (Figure 3.6.2 B-C) resulting in poor cell growth (Figure 3.6.2. A). After IPTG addition on day two, RBP10-myc was strongly expressed and as expected resulted in a severe growth defect. No difference in terms of cell numbers was observed (Figure 3.6.2 A) for the cell lines where *4E-IP* was knocked down prior to RBP10-myc expression. Because of the leaky expression of RBP10 and inefficient knockdown of 4E-IP, the conclusion from this experiment was not clear.

3.7 Tandem affinity purification of RBP10

To corroborate the yeast two-hybrid data, tandem affinity purification (TAP) was performed. For this purpose one *RBP10* gene was altered so as to encode a protein with TAP tag on the N-terminus (*in situ* tagging). The functionality of the tagged protein was confirmed by knocking out the other copy of *RBP10* (Figure 3.7.1 A; lanes 2-6); the cells grew normally (Figure 3.7.1 B).

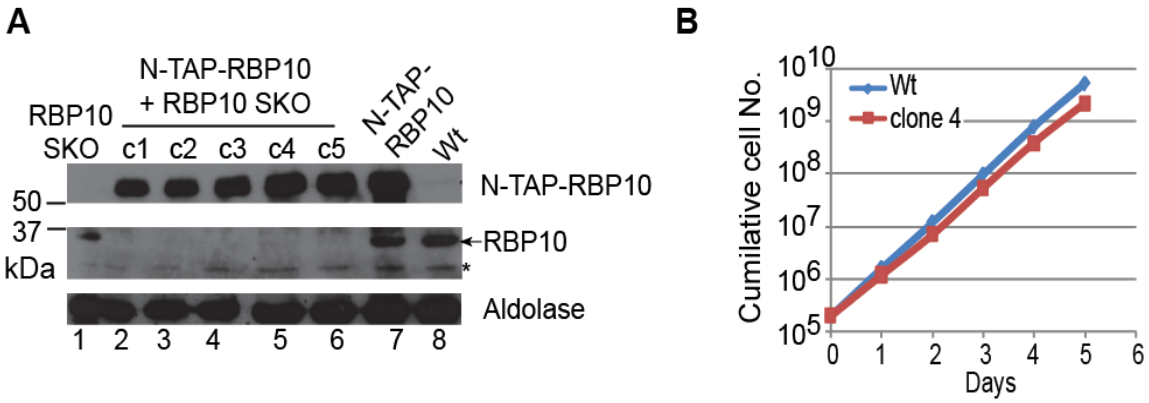

Figure 3.7.1 Functional *in situ* TAP tagging of RBP10 in bloodstream forms. **A)** Western blot confirming the absence of untagged RBP10 (lanes 2-6) in five clones (c1-5) with both in situ N-TAP-RBP10 and RBP10 single knockout (SKO); * is a non specific band. **B)** Cumulative growth curve for one of the clones (c4) in comparison to wild type (Wt) cells.

The purification was done three times with or without RNAse A treatment (Figure 3.7.2 A, C). Cells expressing an inducible GFP with a C-terminal TAP tag served as the control (Figure 3.7.2 B, C). As expected, no signal was detected after the TEV protease cleavage (Figure 3.7.2 A, B; lane 3); the antibody used recognizes the protein A part of the TAP tag that remains bound to the IgG matrix after the 1st purification step.

Figure 3.7.2 C shows a Coomassie stained gel of the eluates recovered after the 2nd purification step using calmodulin beads. In the GFP control, fewer proteins were visible and some of the patterns looked different from the rest. The samples were divided into two gel slices (numbered boxes in Figure 3.7.2 C) per sample and analyzed by mass spectrometry.

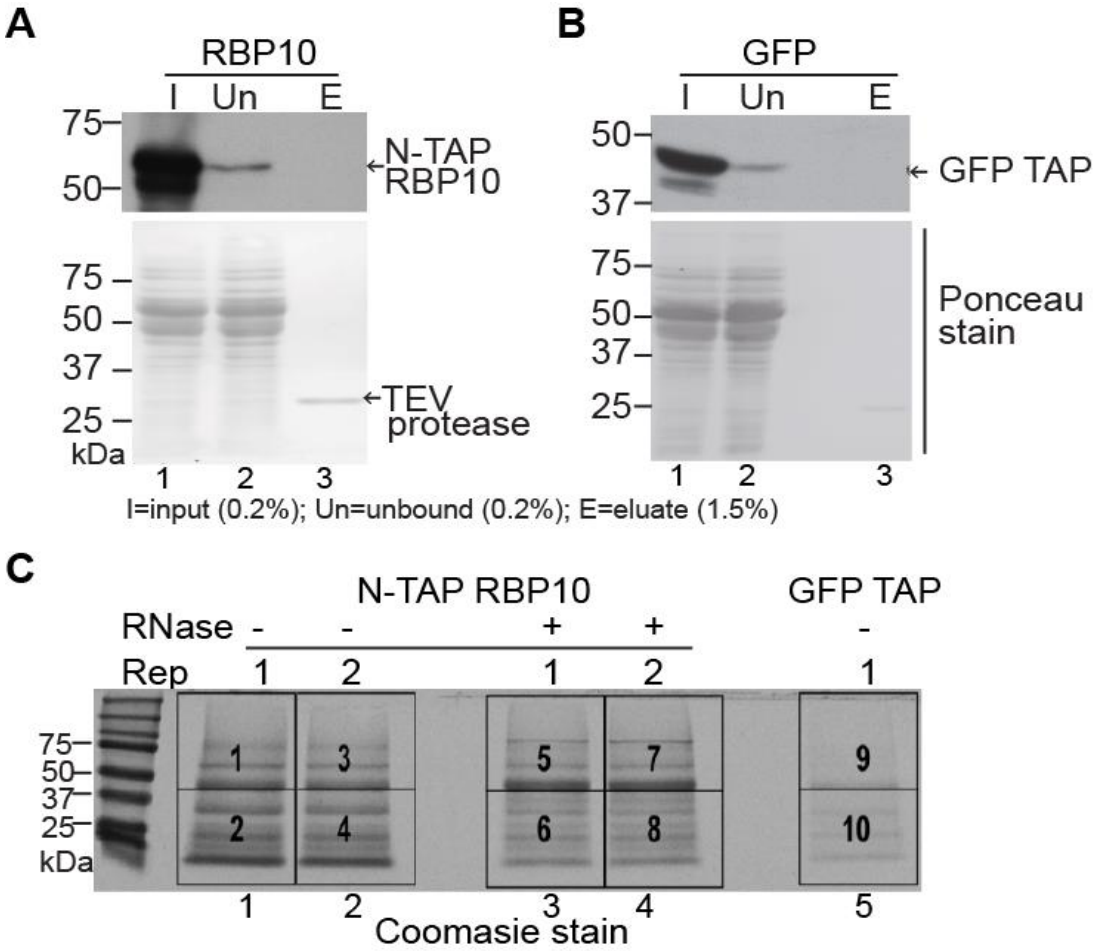

Figure 3.7.2 TAP purification in bloodstream trypanosomes. **A)** First step of the TAP purification with IgG sepharose beads using extracts from cells expressing N-TAP-RBP10 only. Elution was done by TEV protease treatment of the beads. The TEV cleavage efficiency was analysed by western blotting using rabbit anti-PAP antibody. Ponceau staining served as loading control. I: input, Un: unbound (both 0.2% of the lysate) and E: eluate (1.5% of the TEV eluate). A representative blot is shown of the three replicates done with and without RNAse A treatment. **B)** As in (A), but using extracts from cells expressing inducible GFP-TAP that served as background control. **C)** Coomasie stained gel of the TAP elute after the 2nd purification step with calmodulin sepharose beads; samples treated with (lanes 3-4) or without (lanes 1-2, 5) RNAse A. Gel slices shown in numbered boxes were analyzed by mass spectrometry.

120 proteins with at least two peptides in at least two of the replicates were identified. Among those only six were unique to RBP10 purification (Figure 3.7.3, Table 2), the rest (95%) have been identified at least once in other purifications done with different RNA binding proteins. Relative to the GFP control, 21 proteins were 2-fold enriched; their interaction with RBP10 was not dependent on RNA. Common contaminants such as ribosomal proteins and abundant metabolic enzymes were also identified.

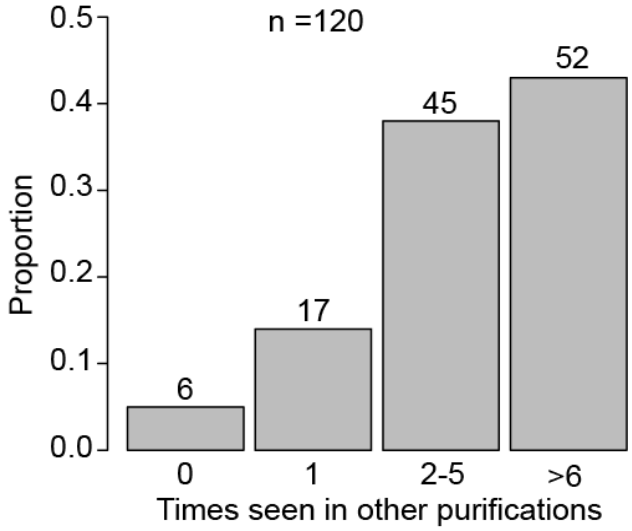

Figure 3.7.3 Proportion of unique proteins identified from RBP10 TAP purification

Since RBP10 promotes mRNA destruction when tethered, proteins involved in mRNA degradation were considered (Table 2). CAF40 was the only one reproducibly detected. It was previously identified in five other purifications, two of them using ZC3H32 and DRBD18; similar to RBP10 both act as negative regulators when tethered to mRNA. Interestingly, DRBD18 was co-purified in RBP10 TAP purification although not unique; ZC3H32 came up in the RBP10 yeast two-hybrid screen (Table 1 B) but only one fragment was detected.

On Figure 3.5.2-A, 4E-IP co-purified RBP10, however, no 4E-IP peptides were detected after RBP10 TAP. Furthermore, none of the RBPs suggested to interact with RBP10 by yeast-two hybrid screen (Figure 3.4.1, Table 1) were detected. One possibility is that interaction with RBP10 was transient; hence the putative interactors of RBP10 failed to withstand the two-step TAP purification. Overall, the RBP10 TAP purification data did not overlap with the genome wide yeast two-hybrid dataset. However, some of these putative partners might be true. An alternative method for example proximity dependent biotin identification (BioID) could be used to confirm interaction in trypanosomes.

Table 2. Peptide numbers of unique proteins plus proteins linked to RNA metabolism that were reproducibly detected after RBP10 TAP. Shown are the 3 replicates (R1-3) with or without RNase A treatment. The GFP-TAP data served as control. A comparison to RBP10 yeast two hybrid is shown plus the fold changes of how much the candidate's fragments repressed a lethal gene in a tethering assay. nd = no data.

		unique peptide counts							average/control		comparison to other studies		
		RBP10 TAP (no RNAse)			RBP10 TAP (+RNAse)			GFP TAP (control)	no Rnase/GFP +1	+ Rnase/GFP+ 1	No. of times previo-usly seen	RBP10 yeast two-hybrid	Tethering effect
Protein ID	Annotation	R1	R2	R3	R1	R2	R3						(Erben.et.al 2014)
Tb927.10.4180	TFIIF-stimulated CTD phosphatase, putative	4	0	2	4	1	1	0	2	2	0	no	1.3
Tb927.4.3400	hypothetical protein, conserved, ARM repeat	3	2	3	2	2	2	0	2.7	2	0	no	1.4
Tb08.27P2.160	ESAG3 pseudogene	0	2	2	2	3	3	0	1.3	2.7	0	no	nd
Tb05.5K5.150	small GTP-binding protein, putative	0	2	2	5	2	3	0	1.3	3.3	0	no	nd
Tb927.7.5480	dihydrofolate reductase-thymidylate synthase	0	2	5	0	0	0	0	2.3	0	0	no	2.1
Tb927.9.12890	hypothetical protein, conserved, Atrophin-1 domain	4	1	2	0	0	0	0	2.3	0	0	no	1.5
Proteins associated with RNA metabolism													
Tb927.8.2780	RBP10	16	26	34	11	36	30	5	4.2	4.3	1	yes	2.6
Tb927.11.14100	DRBD4 (PTB2)	6	5	5	2	1	0	0	5.3	1	3	no	1.0
Tb927.11.4460	ALBA1	3	2	2	3	2	3	0	2.3	2.7	4	no	0.5
Tb927.11.510	UBP2	5	6	9	5	2	1	0	6.7	2.7	6	no	nd
Tb927.11.14090	DRBD18	6	6	6	5	1	0	2	2	0.7	7	no	1.8
Tb927.9.13990	DRBD2	19	3	4	0	0	0	0	8.7	0	7	no	1.0
Tb927.4.410	CAF40	3	3	5	4	5	4	1	1.8	2.2	5	no	1.0

3.8 RBP10 BioID in bloodstream form trypanosomes

As an alternative approach to TAP purification, the proximity-dependent biotin identification method (BioID) was tested in bloodstream forms. To achieve this, RBP10 was in situ tagged on the N-terminus with the bacterial biotin ligase (BirA) and the construct was introduced into RBP10 single knockout cell line. Only a few clones grew; BirA-RBP10 fusion protein was detectable using anti-RBP10 antibody as shown on Figure 3.8-A. However, the untagged RBP10 protein was still present in those cells (slightly less than in the wild type, Figure 3.8-A) indicating either the fusion protein was not fully functional or it was a mixed cell population.

To determine BirA-RBP10 ligase activity, cells were incubated with biotin for 24 hours. A different protein Tb927.8.3850 in situ tagged with BirA, the untreated cells and wild type cells served as controls (Figure 3.8 B). As expected, a band corresponding to BirA-RBP10 was the strongest as detected by Western blotting using anti-streptavidin antibody (Figure 3.8 B, lanes 2,4). Furthermore, BirA-RBP10 resulted to a unique pattern of biotinylated proteins (marked with an asterisk on Figure 3.8 B, lanes 2,4). This highlights potential interacting partners of RBP10 since none was observed in the controls (Figure 3.8 B). Because of time limitation the identity of the proteins by mass spectroscopy was not done in this study.

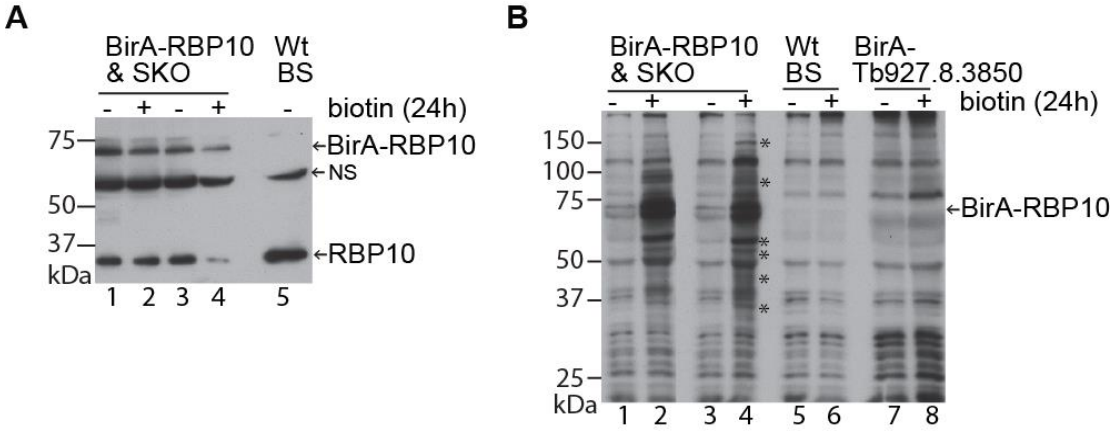

Figure 3.8 N terminus in-situ tagging of RBP10 with BirA in bloodstream form trypanosomes. **A)** Expression of BirA-RBP10 in cells with a single knockout of RBP10. The fusion protein (~69 kDA) and the untagged RBP10 (~32 kDa) were detected using anti-RBP10 antibody. The non-specific (NS) band served as the loading control. Cells were incubated for 24 hours with biotin (lanes 2, 4) to test the ligase activity of the fusion protein. Data from two independent clones is shown. **B)** Detection of biotinylated proteins using anti-streptavidin antibody. Wild type cells (lanes 5, 6) and a different BirA tagged protein (Tb927.8.3850; lanes 7, 8) served as controls. The asterisk (*) highlights RBP10 biotinylated proteins that may be increased with RBP10-BirA.

3.9 Transcriptome wide effects of RBP10 manipulation in *T. brucei*

3.9.1 Time point selection for RNA sequencing

To determine the effects of RBP10 on mRNA abundance and/or translation, I examined transcriptome changes after *RBP10* RNAi in bloodstream forms or RBP10 over expression in procyclic forms. In addition, how such changes affected the association of mRNAs with the polyribosomes was investigated. In both experiments time course studies were done to examine RBP10 changes relative to growth inhibition. Between 14-16 hours post *RBP10* RNAi, no growth inhibition was observed and less than 10% RBP10 protein was detectable (Figure 3.9.1.1).

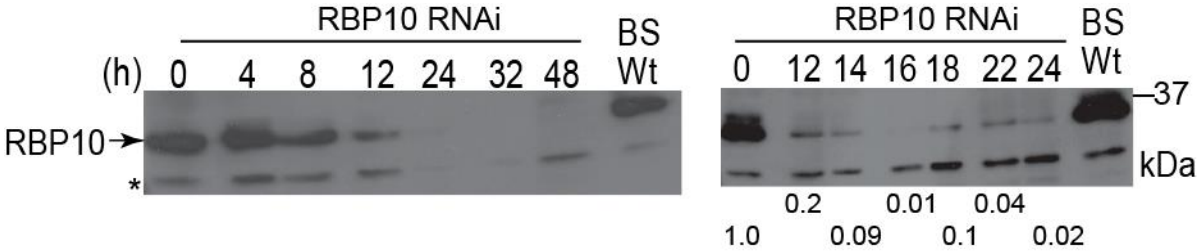

Figure 3.9.1.1 Time course RNAi against *RBP10.* Western blot using cell extracts taken at different time points after *RBP10* RNAi in bloodstream form trypanosomes. RBP10 was detected using anti-RBP10 antibody. The protein quantification relative to the minus RNAi control is shown below the blot. * is non specific band used as the loading control.

In case of RBP10-myc over expression (Figure 3.9.1.2) in procyclics, expression for 24 hours resulted to a moderate decrease in global translation (Figure 3.9.1.2 A, B). Consequently, earlier time points were checked. RBP10-myc was detectable two hours post induction; within six hours it was strongly expressed (Figure 3.9.1.2 C, D). Global translation (Figure 3.9.1.2 B) and growth were unaffected after 4-6 hours of RBP10-myc expression (Figure 3.9.1.2 C, D).

Based on these time courses, 15 hours of *RBP10* RNAi in bloodstream forms (Figure 3.9.1.1) and 6 hours of RBP10-myc expression in procyclic forms (Figure 3.9.1.2) were selected for transcriptome analyses.

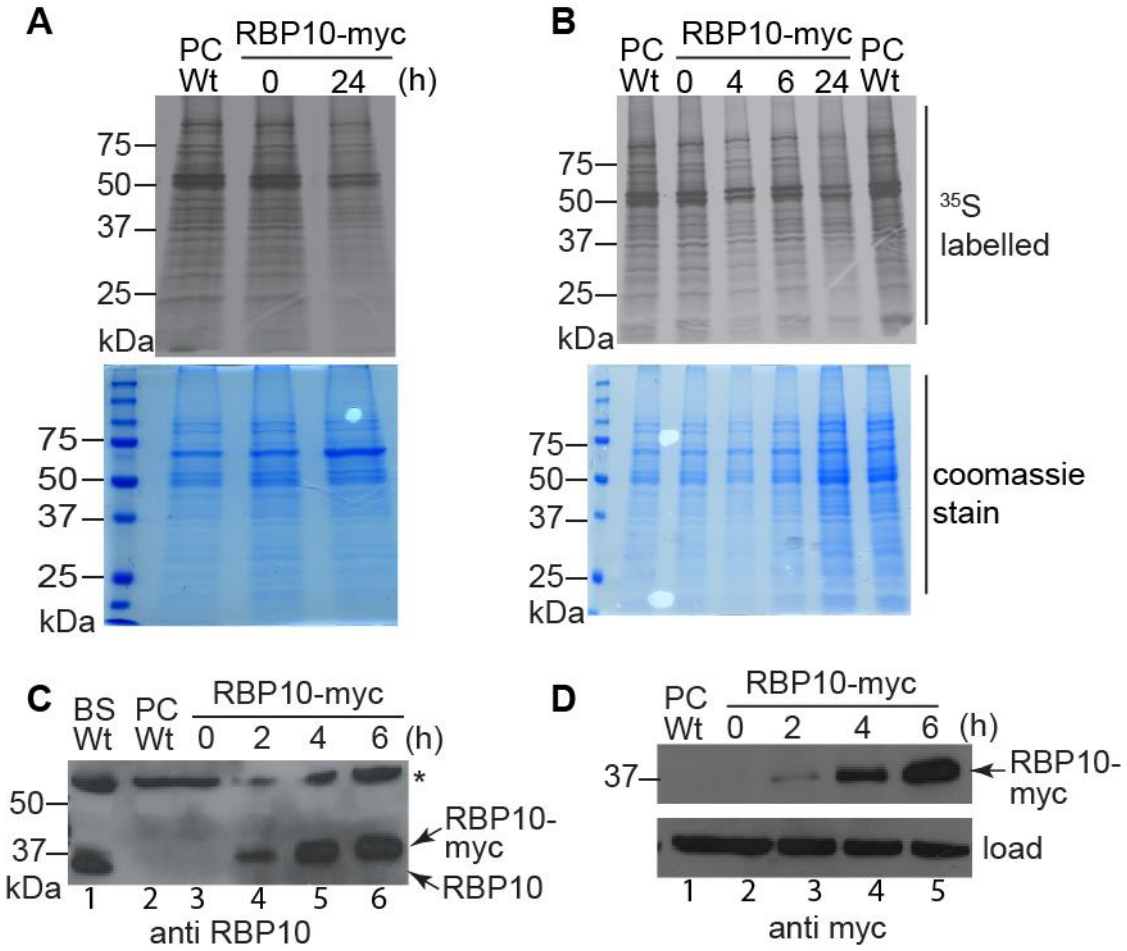

Figure 3.9.1.2 Time course expression of RBP10 in procyclic cells. (**A & B)** Global translation at different time points after RBP10-myc expression as detected using ^{35}S labelling. **(C & D)** Levels of RBP10-myc after 2, 4, and 6 hours of expression as detected by Western blotting using anti RBP10 (C) or anti myc (D) antibody.

After 15 hours of *RBP10* RNAi in bloodstream forms or 6 hours of RBP10-myc expression in procyclic forms, the cell extracts were subjected to polysome fractionation; in both cases uninduced cells served as control. RNAseq was used to quantify the mRNA changes from two pools of the polysome fractions, i) lighter fractions including monosomes, ii) denser fractions with at least two ribosomes; shown on Figure 3.9.1.3-4 (A). Analyses of the input samples (without fractionation) was used to monitor changes in mRNA abundance.

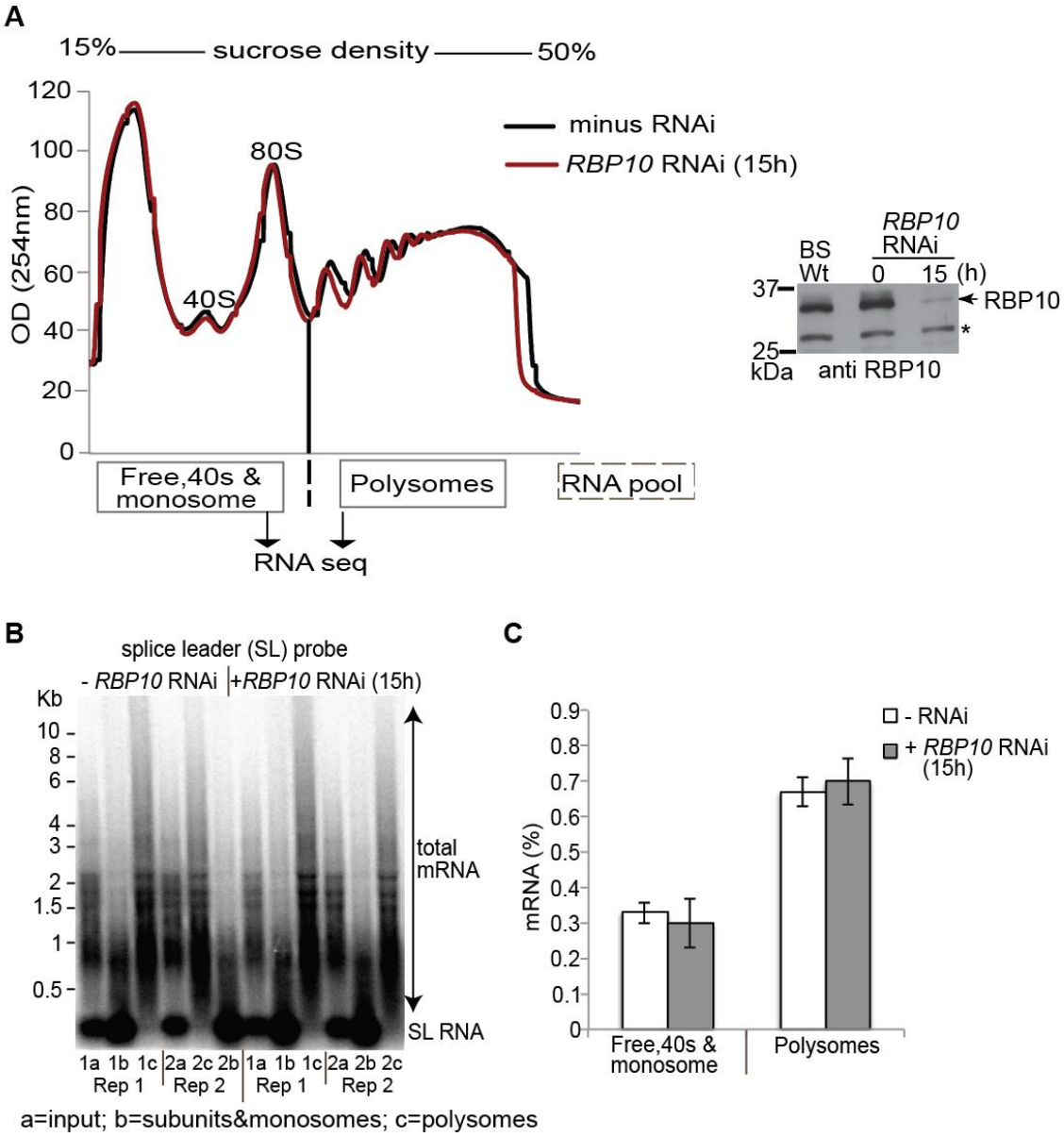

Figure 3.9.1.3 Polysome profiling in bloodstream form trypanosomes depleted of RBP10. **A)** A representative polysome profile obtained using extracts from cells with or without *RBP10* depletion by RNAi for 15 hours. The two mRNA pools analyzed by RNA sequencing after rRNA depletion are shown. The experiment was done in duplicates and cells lacking RNAi served as the control. **B)** Northern blot using 1/20th of the RNA isolated from the pooled (b, c) samples; for the input (a) 2 µg of the RNA was loaded. The total mRNA was detected by hybridization of the blot with a splice leader probe. **C)** Quantifications of the splice leader signal obtained from (B). The Northern blot was repeated twice.* = non specific band.

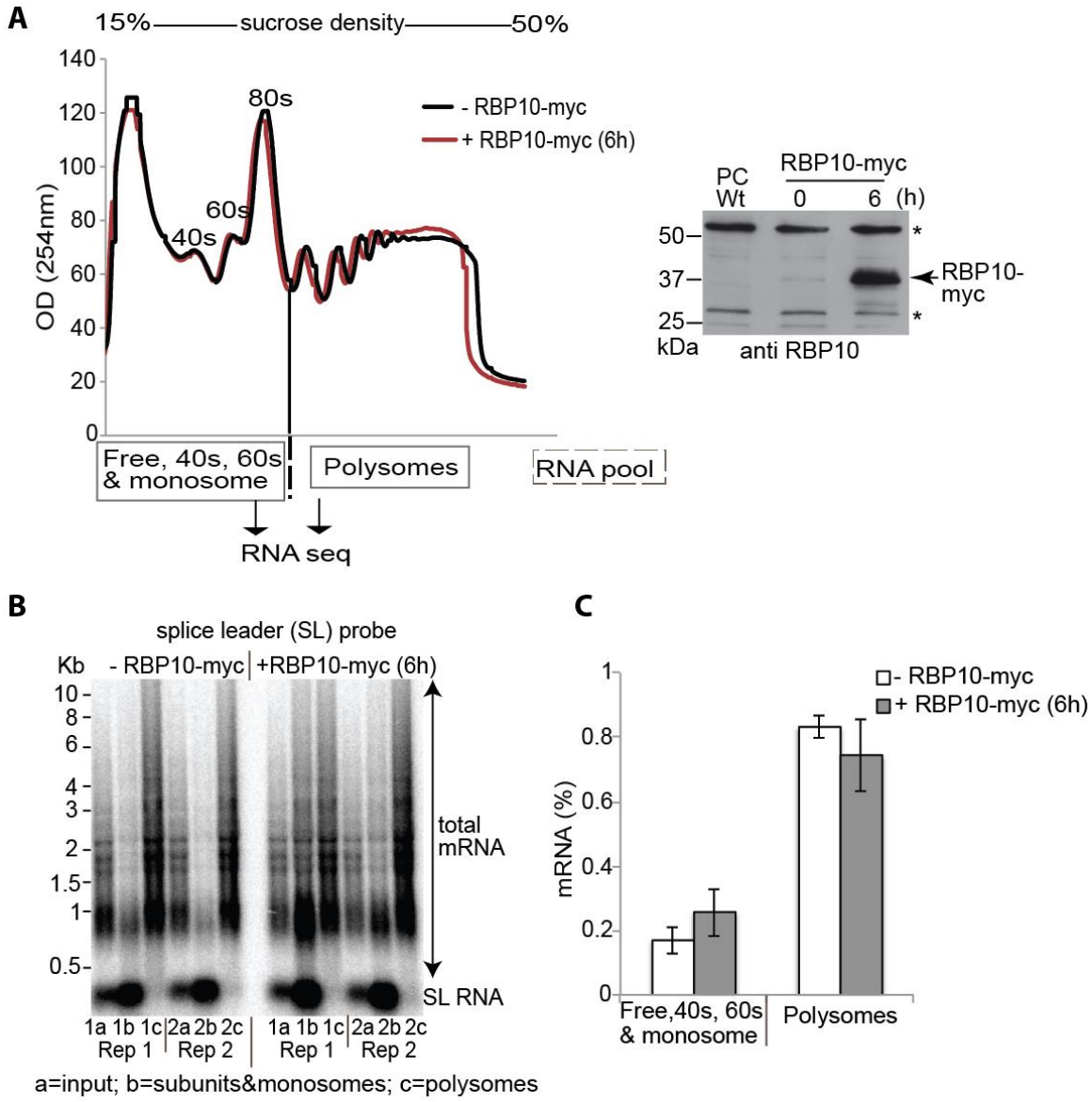

Figure 3.9.1.4 Polysome profiling in procyclic form trypanosomes over expressing RBP10. **A)** A representative polysome profile obtained using extracts from cells with or without RBP10 expression for 6 hours. The two mRNA pools analyzed by RNA sequencing after rRNA depletion are highlighted. For each condition two biological replicates were used. **B)** Northern blot using 1/20th of the RNA isolated from the pooled (b, c) samples; for the input (a) 2 µg of the RNA was loaded. The total mRNA was detected by hybridization of the blot with a splice leader probe. **C)** Quantifications of the splice leader signal obtained from (B). The Northern blot was repeated twice. * = non specific band.

3.9.2 Transcriptome changes after RBP10 depletion for 15 hours

Both the DESeq2 package and RPM (reads per million) estimations were used for the RNAseq data analyses. Figure 3.9.2.1-A shows the reproducibility between replicates of the input samples. Clustering was strong for the samples without *RBP10* RNAi in comparison to the ones where RBP10 was depleted, nonetheless there was 79% variation between the conditions as determined by principal component analysis (PCA). After normalization based on DESeq2 algorithm, similar distributions of the read counts was observed (Figure 3.9.2.1 B), confirming the normalization was successful.

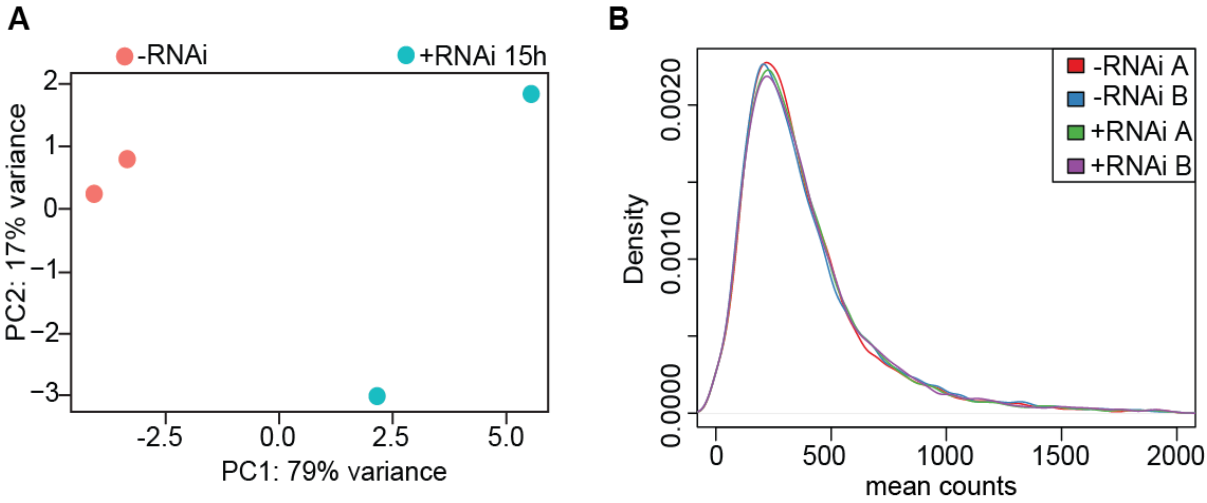

Figure 3.9.2.1 Quality check of the RNAseq data. **A)** PCA plot of the input samples with and without *RBP10* RNAi. **B)** Read counts distribution shown using a density plot.

DESeq2 identified 420 mRNAs as being significantly ($p<0.05$) affected after RBP10 depletion (highlighted in red in Figure 3.9.2.2). Considering those with at least ~1.5 fold change (those above/below the dotted line in Figure 3.9.2.2), 103 mRNAs were up regulated while 103 mRNAs got down regulated.

Most of the mRNAs down regulated are more abundant in bloodstream forms (Figure 3.9.2.3); ~10% of the mRNAs that showed reduced expression encode proteins involved in glucose metabolism; the main source of ATP in bloodstream form trypanosomes. Apart from RBP10, five other regulatory proteins PUF11, ZC3H31, ZC3H46, DRBD5 and RBP9 were significantly reduced possibly leading to combinatorial effects.

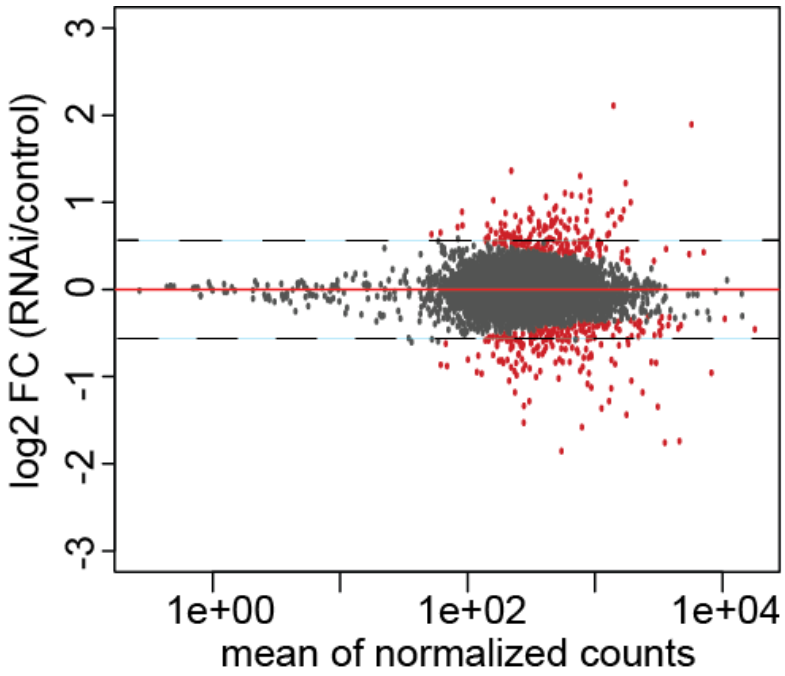

Figure 3.9.2.2 mRNA changes (input samples) after *RBP10* RNAi for 15 hours. The MA plot highlights the fold changes relative to the cells without *RBP10* RNAi (uninduced). mRNAs that showed differential expression ($p<0.05$) are highlighted in red; the two dashed lines show the mRNAs subsets with at least 1.5 fold change.

The majority of the mRNAs up-regulated were procyclic form specific (Figure 3.9.2.3). The transition from bloodstream to procyclic form involves changing of the surface coat proteins, mitochondrion activation and switching of the energy metabolism from glycolysis to oxidative phosphorylation. 34% of the mRNAs that showed increased expression encode proteins involved in different mitochondrion-specific pathways. Among these ~14% code for proteins involved in the citric acid cycle. The mRNA coding for the surface coat protein EP1 was 4x up-regulated. ZC3H22 and ZC3H20 were the only RNA binding proteins significantly up-regulated after RBP10 depletion for 15 hours.

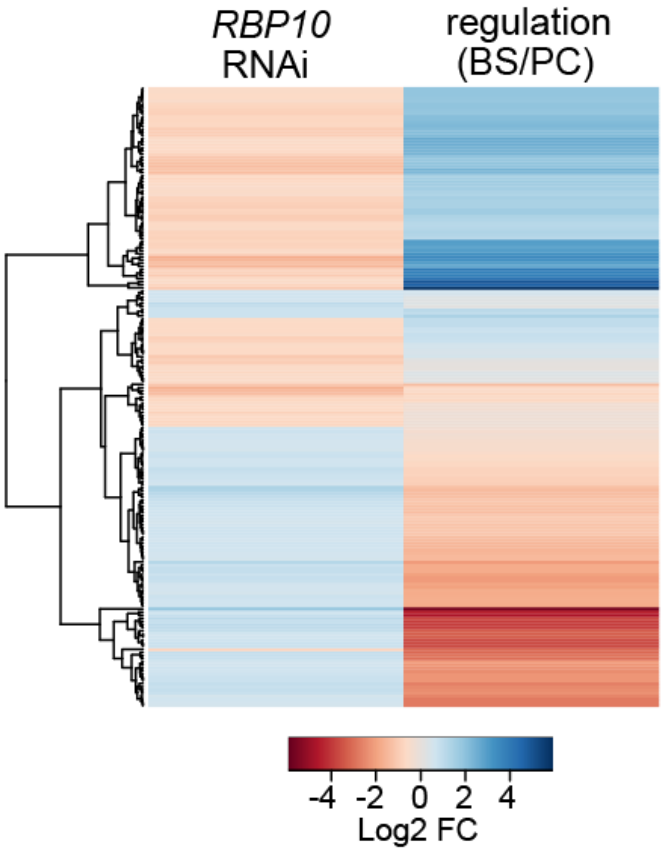

Figure 3.9.2.3 mRNA changes after *RBP10* RNAi. **A)** A clustered heatmap showing the relationship between *RBP10* RNAi effects and development regulation; the dataset for developmental regulation is from [144], comparing bloodstream (BS) versus the procyclic forms (PC). Uninduced cells served as control.

Next, RBP10 effect on translation was investigated by comparing the changes seen on total mRNA (input samples) versus polysome loading (proportion of mRNAs associating with polyribosomes). In this case, RPM (reads per million) for each gene were calculated. For the polysome fractions pools, RPM values were calculated after normalizing for the amounts of starting RNA using the signal from the splice leader (Figure 3.9.1.3 B-C). To estimate the proportion of the mRNAs associating with the polyribosomes, individual gene RPM was divided with the sum of the amount found in the two pools of mRNA sequenced after sucrose density fractionation (Figure 3.6.1.3 A); comparison were made between samples with or without *RBP10* RNAi. As expected, RBP10 mRNA showed reduced abundance as well as decreased polyribosome association (Figure 3.9.2.4). PUF11 is the only other RNA binding protein that had a similar pattern. 53 mRNAs showed increased (at least 1.5x) association (violet-red and cyan coloured on Figure 3.9.2.4) with the polyribosomes

when RBP10 was depleted. Among those 8 had both increased (at least 1.5x) polysome loading and abundance (coloured in cyan on Figure 3.9.2.4). Overall, there was only a very slight correlation (R=0.38) between the effects of RBP10 depletion on mRNA abundance and on polysome loading.

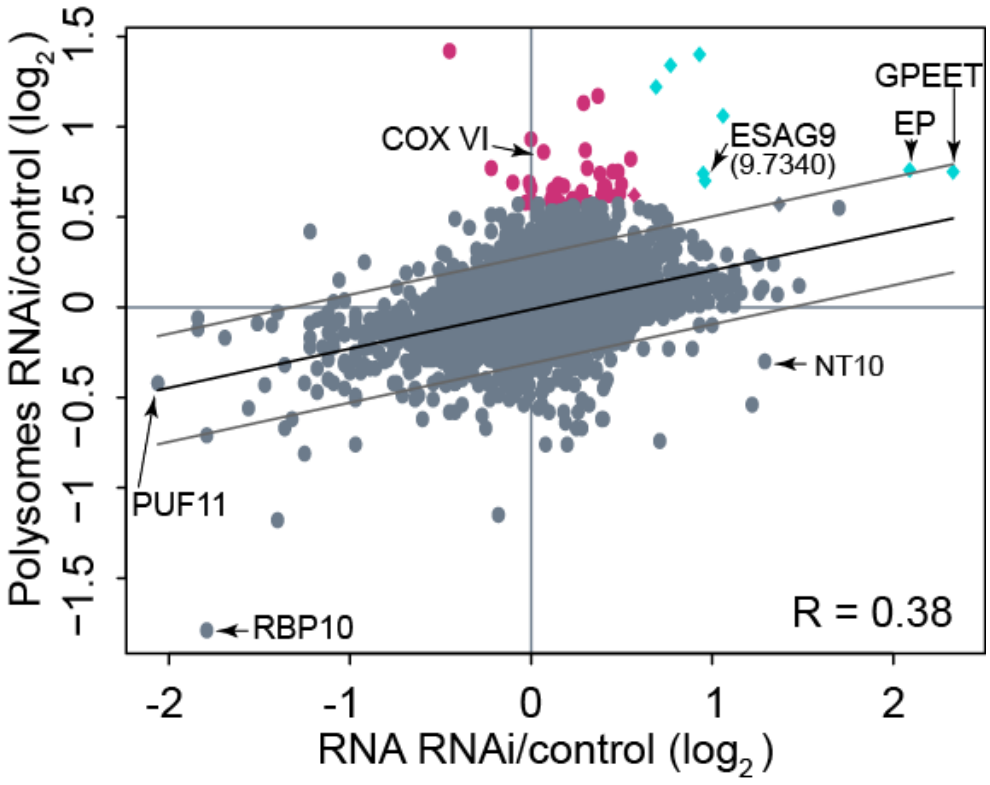

Figure 3.9.2.4 The relationship between the effects of RBP10 depletion on the mRNA abundance and polysome loading (proportion of mRNAs associating with polyribosomes). mRNAs that moved (at least 1.5 fold) to the polysomes are highlighted in two colours, i) violet-red are those that showed increased polysome association but less changes in abundance. ii) In cyan had both changes; the names of a few selected mRNAs that are developmentally regulated are shown. The linear regression line (in black) plus the 95% confidence limits (in grey) of the scatter plot are included.

3.9.3 Transcriptome changes after RBP10 expression for 6 hours in procyclic cells

The polysome RNA seq data was analyzed as described in section 3.9.2. PCA plot (Figure 3.9.3.1 A) was made to check variation between the replicates. Figure 3.9.3.1 B highlights the success of the normalization procedure using DESeq2 algorithm.

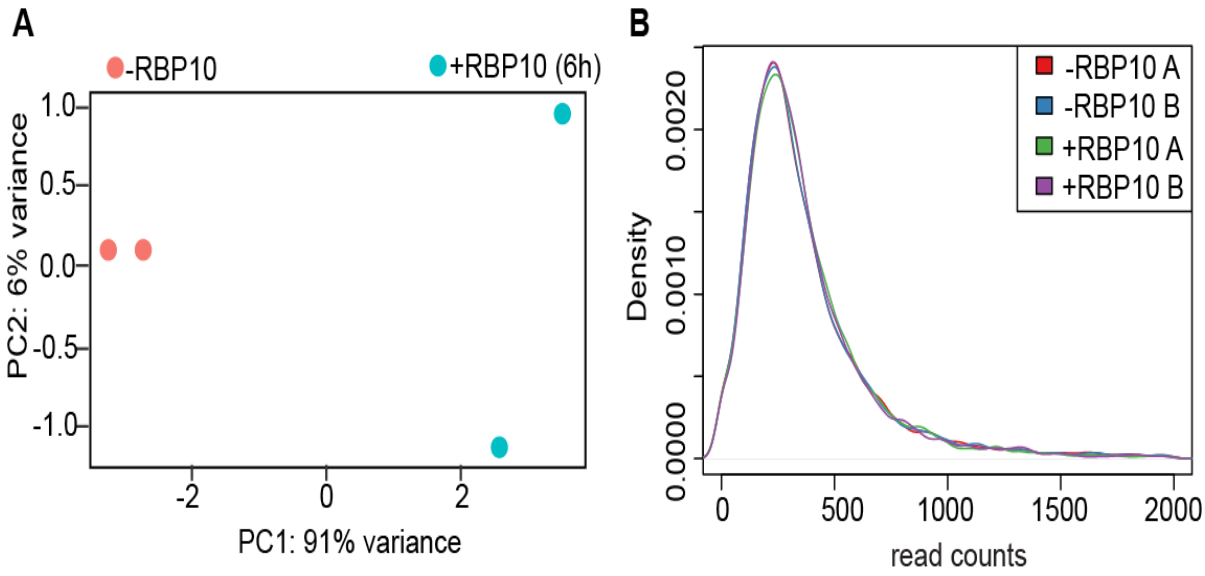

Figure 3.9.3.1 Quality check of the RNAseq data. **A)** PCA plot of the input samples with and without RBP10 expression. **B)** Read counts distribution shown using a density plot.

Upon RBP10 expression for 6 hours in procyclic cells, 453 mRNAs were significantly affected ($p<0.05$) as determined using DESeq2. Considering those with at least ~1.5 fold change (those above/below the dotted line in Figure 3.9.3.2), 29 mRNAs were up regulated while 48 mRNAs got down regulated.

RBP10 mRNA was 18x up regulated (Figure 3.9.3.2); within those up regulated, 16 are more abundant in the bloodstream forms (Figure 3.9.3.3). Although it was not a perfect inverse of what was observed after *RBP10* RNAi in bloodstream form, 5 mRNAs up regulated after RBP10 expression were at least 1.5 fold down regulated after *RBP10* RNAi. Three encodes enzymes involved in the glycolytic pathway, one encodes RNA binding protein DRBD5 and the other one codes for a phosphatase that is bloodstream form specific.

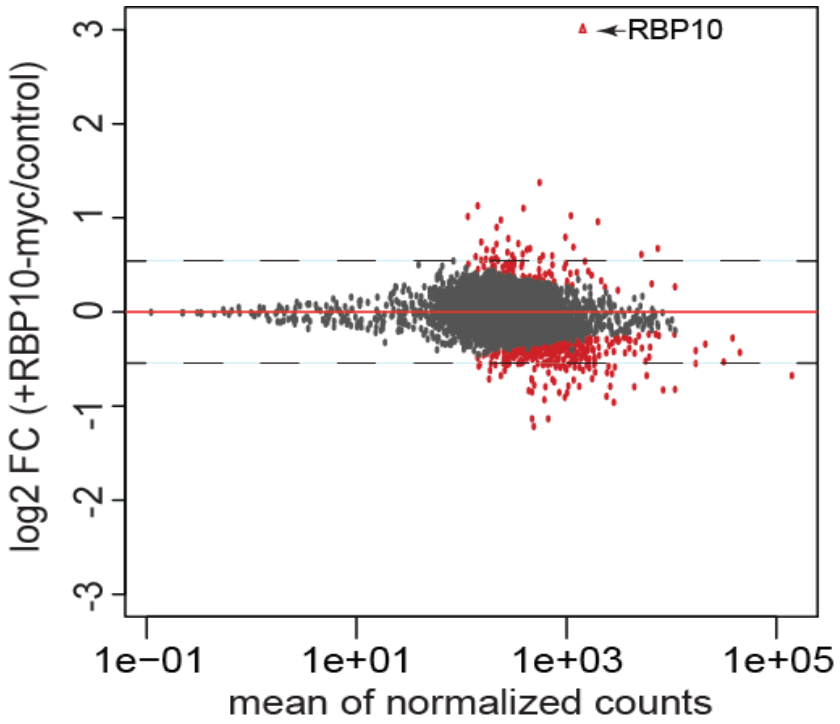

Figure 3.9.3.2 mRNA changes (input samples) after RBP10 expression for 6 hours in procyclic cells. The MA plot highlights the fold changes relative to the cells without RBP10-myc (uninduced). mRNAs that showed differential expression ($p<0.05$) are highlighted in red; the two dashed lines show the mRNAs subsets with at least 1.5 fold change.

RBP10 expression for 6 hours resulted to a moderate effect on the transcriptome; mRNAs significantly down regulated formed the majority (Figure 3.9.2-3). Within the 48 mRNAs that met the cut off criteria ($p<0.05$ and 1.5x down), 26 are more abundant in the procyclic trypanosomes. Interestingly, 11 of those showed opposite effects when RBP10 was depleted in bloodstream forms. EP1 mRNA was 4 times up regulated after RBP10 RNAi in bloodstream form; upon expression of RBP10 in procyclics EP mRNA was 1.6 fold decreased. In addition trans sialidase mRNA which encodes another membrane protein showed similar pattern. In terms of regulatory factors, ZC3H22 mRNA was affected after RBP10 depletion however expression of RBP10 in procyclics resulted to only 1.4 fold decrease of the ZC3H22 mRNA.

Seven chaperone mRNAs more abundant in procyclic forms were significantly down regulated. The procyclic form specific phoshoglycerate kinase B (PGK-B) was 1.6 fold down regulated after RBP10 expression in procyclic cells.

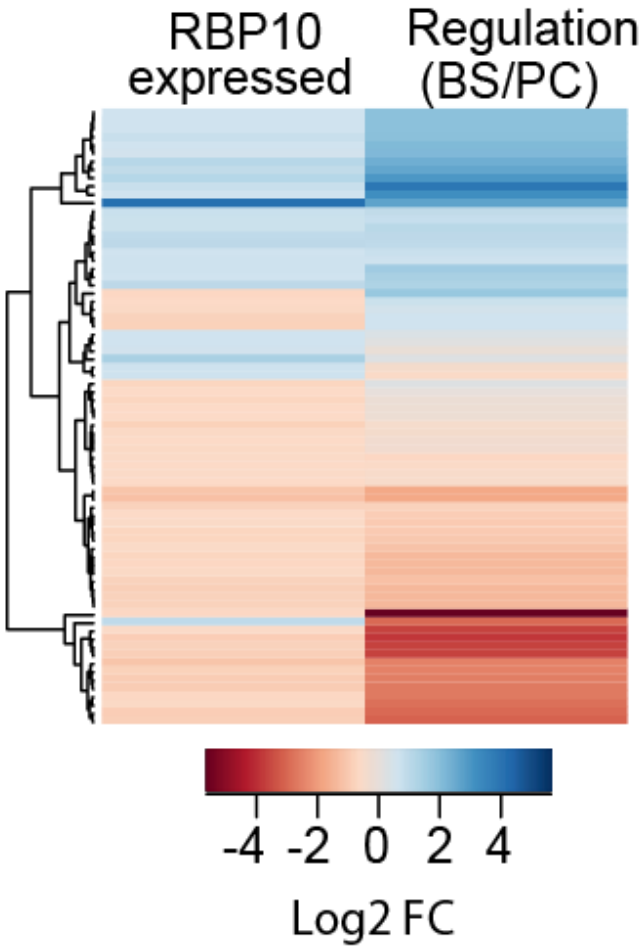

Figure 3.9.3.3 mRNA changes after RBP10-myc expression in procyclic cells. **A)** A clustered heatmap showing the relationship between effects of RBP10-myc expression and development regulation; the dataset for developmental regulation is from [144], comparing bloodstream (BS) versus the procyclic forms.

3.10 RBP10 mRNA targets

Changes in RBP10 levels significantly up-regulate or down-regulate many mRNAs (sections 3.9.2-3.9.3). But, does RBP10 bind all of these mRNAs? To identify mRNAs directly bound by RBP10, a cell line with functional N-terminally TAP tagged RBP10 was used. The mRNP complexes were preserved by UV cross-linking. After precipitation using IgG-agarose beads, elution was done using TEV cleavage and the proteins digested with proteinase K. The co-purified (bound) RNA was isolated and identified by RNA seq. To determine the enrichment, the unbound fraction served as the control. The replicates from both eluate and unbound fractions were highly reproducible as determined by PCA plot (Figure 3.10.1 A).

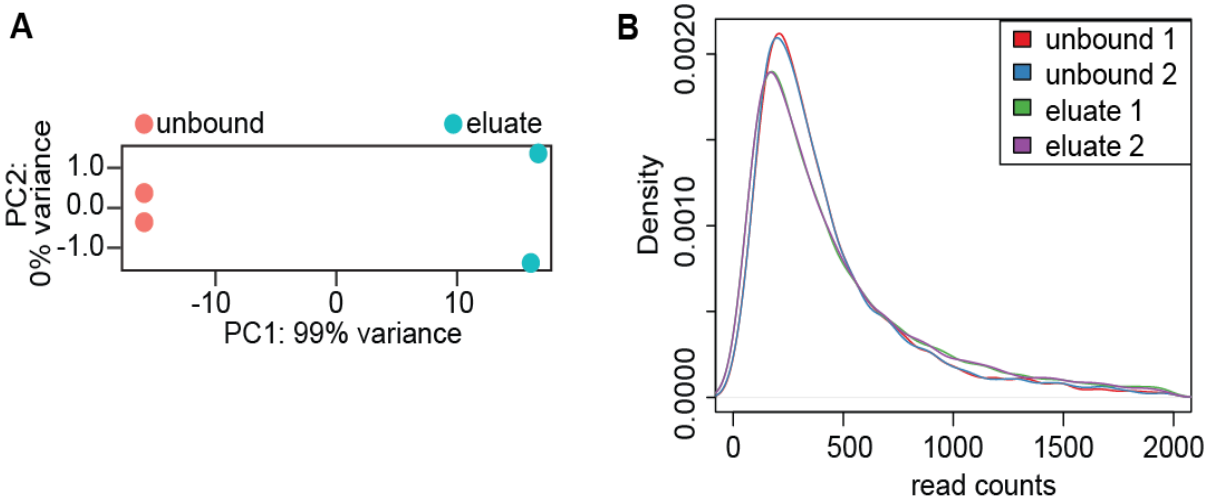

Figure 3.10.1 Quality check of the RNAseq data. **A)** PCA plot showing replicates of the eluate and unbound samples after TAP-RBP10 RNA pull downs. **B)** Read counts distribution shown using a density plot.

About 300 mRNAs (Figure 3.10.2 A) were at least 3 fold enriched on average when comparing bound versus the unbound. mRNAs down-regulated (>1.5x) after RBP10 RNAi were strongly underrepresented in the bound fraction. Interestingly, 39% (40 out of 103) of the mRNAs up-regulated (>1.5x) after RBP10 RNAi (15h) were significantly enriched in the RNA IP data set (Fig. 3.10.2 A). This data set highlights mRNAs targets bound and regulated by RBP10 in bloodstream form. On the contrary, <1% (3 out of 103) were both bound and down regulated (1.5x) after RBP10 RNAi; one of them being the mRNA encoding RBP10.

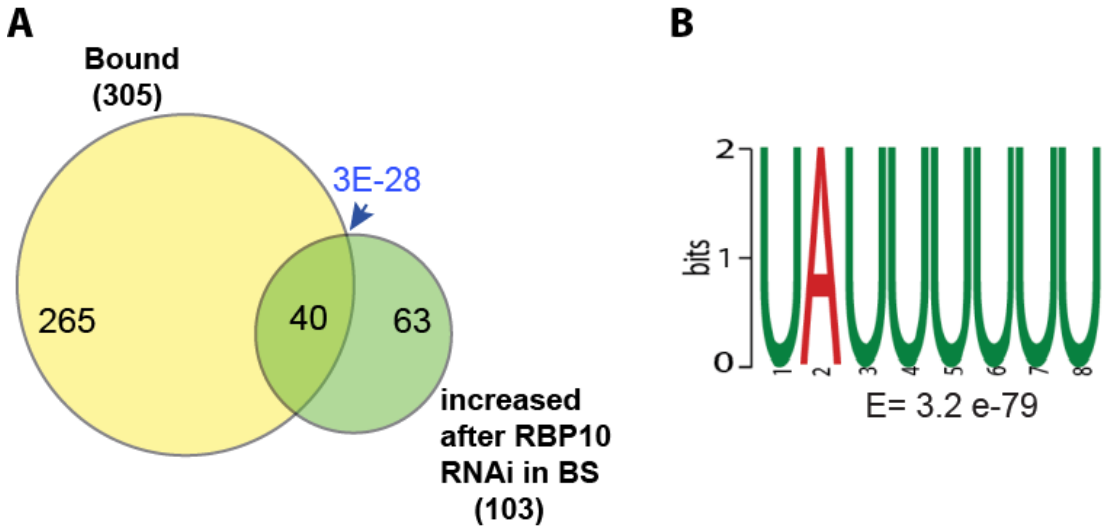

Figure 3.10.2 mRNAs bound and regulated by RBP10. A) A Venn diagram highlighting mRNAs bound and regulated by RBP10. The RNAi data set is the same as in Figure 3.9.2.3. mRNA was considered bound if it had at least 3 fold enrichment (bound/unbound). Fischer exact test values, calculated for the overlap in the datasets, are in blue. **C**) The motif enriched in the 3' UTRs of mRNAs strongly bound by RBP10; the DREME tool (http://meme-suite.org/tools/meme) was used for the motif search.

To narrow down to the potential direct targets of RBP10, I considered the fact that RBP10 acts as a negative regulator when attached to a reporter mRNA (Figure 3.1). Therefore, I looked for the mRNAs that are i) bound by RBP10 ii) up-regulated after RBP10 RNAi iii) developmentally regulated (highly unstable in bloodstream forms). Using these filters, 31 mRNAs (Figure 3.10.2 and Table 3) were identified as direct target of RBP10 and possibly are repressed by RBP10. Interesting candidates within this list included mRNAs encoding surface protein EP procyclin, two RNA binding proteins ZC3H20 & ZC3H22, three different cytochrome complex proteins, four enzymes required for the procyclic form energy metabolism, two protein kinase, one phosphatase and several other proteins of unknown function (full list is shown on Table 3). Next, the 3' UTRs of the enriched transcripts were analysed for motifs using the motif-based sequence analysis tool (DREME). We used 3' UTRs of the <0.7x bound mRNAs as background control. The motif UAUUUUUU was highly enriched (E=3.2 e-79) in the bound fraction (Figure 3.10.2 B). Convincingly, all the 31 high confidence targets of RBP10 contained at least one of the identified motif (Table 3).

Table 3 High confidence mRNA targets of RBP10. BS=bloodstream form, PC=procyclic form, KD=knockdown

Gene ID	Annotation	BS KD	PC+ RBP10	PC/ BS	Repressor effect	UA(A_6) motif count
Membrane						
Tb927.10.10250	EP2 procyclin (EP2)	3.7	0.6	50	n/a	3
Tb927.7.6850	Trans-sialidase (TbTS)	2.2	0.6	2.6	n/a	4
Tb927.8.7340	Trans-sialidase	1.6	0.4	3.3	n/a	3
Tb927.4.3500	Amastin-like protein	1.6	0.9	10	n/a	1
Mitochondrion						
Tb927.10.2350	Pyruvate dehydrogenase complex	1.7	0.8	4.8	n/a	1
Tb927.9.5900	Glutamate dehydrogenase	1.8	1.1	7.1	n/a	4
Tb927.10.4280	Complex III cyctochrome bc_1	1.8	0.7	25	n/a	2
Tb927.11.15550	NADH-cytochrome b5 reductase	1.5	0.9	2.9	n/a	1
Tb927.5.3040	Cytochrome c Oxidase complex	1.6	0.9	2.5	n/a	1
Tb927.7.210	Proline dehydrogenase	1.8	1.0	7.7	n/a	3
Tb927.9.4310	Tricarboxylate carrier	1.5	1.3	4.3	n/a	1
Regulation						
Tb927.2.4200	Protein kinase	1.5	0.8	2.3	n/a	4
Tb927.11.15010	NEK21 protein kinase	1.9	0.9	1.8	n/a	4
Tb927.10.8050	Protein phosphatase	1.9	0.8	3.5	n/a	2
Tb927.7.2680	ZC3H22	1.6	0.7	2.4	4.9	7
Tb927.7.2660	ZC3H20	1.5	1.1	3.5	n/a	2
Other						
Tb927.9.7470	Purine nucleoside transporter NT10	2.2	0.7	12.5	n/a	1
Tb927.10.7700	ABC transporter	1.6	0.7	2.9	n/a	2
Tb927.8.7730	Dihydroceramide synthase	1.7	0.5	2.3	n/a	2
Tb927.10.10770	Generative cell specific 1	1.6	0.6	2.1	n/a	2
Tb927.4.2410	Haloacid dehalogenase-like hydrolase	1.5	1.4	1.9	n/a	1
Tb927.10.3300	Unknown function	1.5	0.8	3.3	n/a	1
Tb927.4.4940	Unknown function	1.8	0.8	4.4	n/a	3
Tb927.7.5550	Unknown function	1.5	0.7	2.8	n/a	2
Tb927.6.3880	Unknown function	2.5	0.6	2.6	n/a	6
Tb927.10.11630	Unknown function	1.6	0.9	2.2	0.6	2
Tb927.1.1470	Unknown function	1.5	0.9	2.6	1.5	5
Tb927.11.1830	Unknown function	1.8	0.9	2.2	n/a	2
Tb927.9.1520	Unknown function	1.7	0.8	3.5	n/a	1
Tb927.9.13200	Unknown function	1.7	0.6	12.5	n/a	2
Tb927.11.16300	Unknown function	1.5	0.9	6.7	n/a	1

3.11 RBP10 binding to *EP 3' UTR*

To validate the RNA IP data, *EP 3' UTR* was selected. It is already known [155-158] that in the bloodstream forms the *EP* procyclin mRNA is not translated and is highly unstable (half life <5 mins). Interestingly, two repeats of the UAUUUUUU sequence are present in the 26mer instability element on the *EP 3' UTR*. The 26mer element is responsible for the extreme instability and translation repression of *EP mRNA* in bloodstream form [155, 158], and is known to be single stranded *in vivo* [156]. To test whether RBP10 binds *EP* mRNA via this element, RNA pull-down was performed using anti-RBP10 antibody in cells constitutively expressing *CAT* reporter mRNA containing either full length *EP 3' UTR* (*CAT-EP*) or a deletion version lacking the 26mer (*CAT-EPΔ26*) instability element (Figure 3.11.1 A, B). The cDNA prepared from the co-purified RNA or the unbound fraction was used in a semi-quantitative RT-PCR (Figure 3.11.1 C), and a region on the ORF of the *CAT* reporter and tubulin mRNAs was amplified. The *CAT* PCR product was detected only in the sample with the full length *EP 3' UTR*. On the other hand, no product was detectable by PCR when the motif was deleted (Figure 3.11.1 C), confirming the regulatory 26mer element is required for RBP10 binding the *EP 3' UTR*.

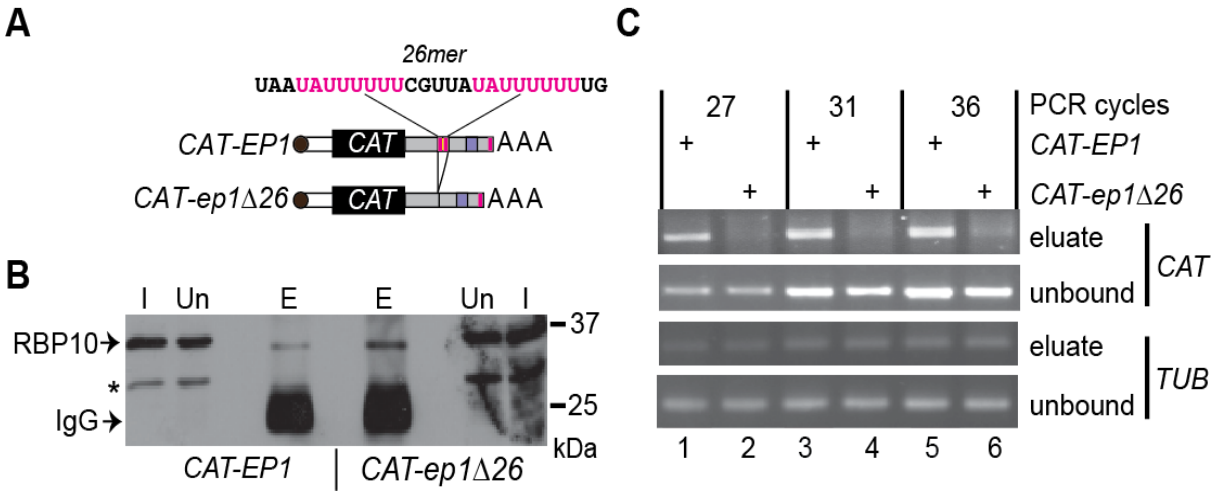

Figure 3.11.1 The *26mer* instability element in the *EP1 3' UTR* is required for RBP10 binding. **A)** Schematic representation of the *CAT* reporter used. **B)** Western blot showing RBP10 protein was equally pulled down after immunoprecipitation using agarose beads coupled with anti-RBP10 antibody; I = input (3%); Un = unbound (3%); E = eluate (1%). **C)** RT-PCR to detect *CAT* and *tubulin* (*TUB*) mRNAs after RBP10 immunoprecipitation. The cells used expressed *CAT* reporter with either full length *EP* or *EPΔ26* 3' UTR. For tubulin only a very faint band was detected in the eluate. * = non specific band used as the loading control.

Depletion of RBP10 by RNAi for 17 or 24 hours successfully rescued the translation repression conferred by the presence of the 26mer instability element (Figure 3.11.2). Increased expression was only seen for the *CAT-EP1* reporter but not for mutant (*CAT-EPΔ26*) version. However, a decline in the CAT protein after 24 hours of RBP10 RNAi was observed in the cells with *CAT-EPΔ26 3' UTR*; possibly due to growth inhibition.

These results demonstrate that RBP10 binds the *EP* 3'-UTR via the 26mer sequence, and the element is required for regulation by RBP10 of a reporter mRNA bearing the *EP1* 3'-UTR. The data strongly support $UA(U)_6$ as the binding motif for RBP10, however direct binding to the motif will need to be confirmed using other methods such as electrophoretic mobility shift assay (EMSA).

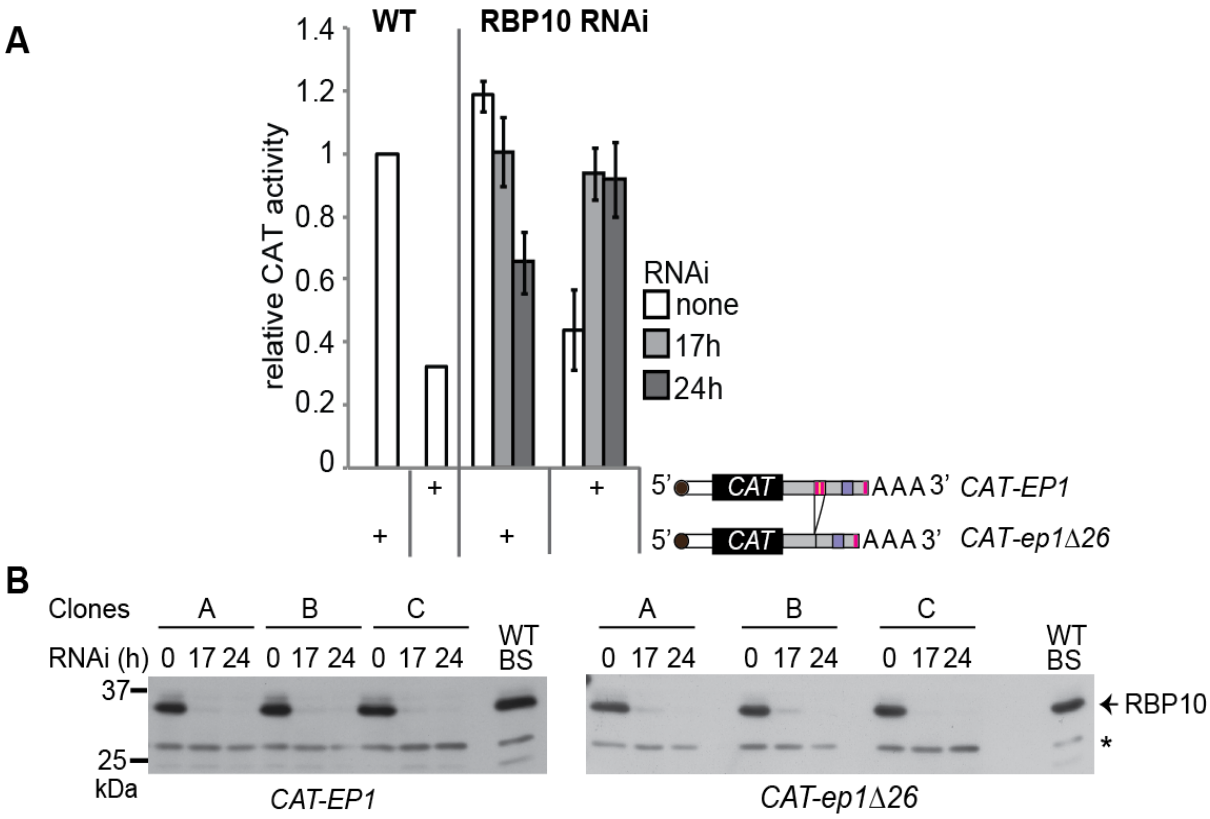

Figure 3.11.2 Regulation by the *EP1 3' UTR* is dependent on the presence of RBP10**. A)** CAT activity was measured in the same cell lines as in Figure 3.11.1 but with an inducible RNAi construct targeting *RBP10*. Depletion of RBP10 for 17h or 24h increased the expression of the CAT reporter (*CAT-EP*1) containing the two UAUUUUUU repeat sequence. **B)** RBP10 levels after 17h and 24h knockdown as detected by Western blot using anti-RBP10 antibody; data from three independent clones is shown.

3.12 RBP10 and trypanosome differentiation

3.12.1 Bloodstream to procyclic form conversion

Bloodstream-to-procyclic form differentiation involves an intermediate stage known as the stumpy form. In culture, synchronised stumpy form cells differentiate efficiently to procyclic forms after incubation with 6mM cis-aconitate at 27^0C [123, 159].

Given that RBP10 binds to and inhibits expression of procyclic form specific mRNAs in bloodstream forms, I wanted to determine whether depletion of RBP10 accelerates differentiation towards procyclic forms in absence of external stimuli such as cis-aconitate. To test this, a tetracycline inducible stem-loop construct targeting RBP10 was introduced in pleomorphic bloodstream form trypanosomes. After 17 hours of RNAi induction RBP10 protein was strongly reduced, the cell density was unaffected and the differentiation marker PIP39 was strongly up regulated while EP procyclin was undetectable (Figure 3.12.1.1 A). Prolonged induction of RBP10 RNAi inhibited growth (Figure 3.12.1 A) with PIP39 levels increasing over time (Figure 3.12.1.1 B).

In vitro differentiation to procyclic forms requires a change in media and temperature reduction. Therefore, after 17 hours of *RBP10* RNAi, cells were transferred into procyclic form media and incubated at 27^oC (Figure 3.12.1.1 C). As positive control for differentiation, the growth arrested stumpy cells couldn't be used since the *RNAi* was done using log phase bloodstream form cell. Consequently, high-density wild type EATRO1125 cells were treated for 17 hours with 6 mM cis-aconitate at 27^0C before transfer to procyclic medium without cis-aconitate. The untreated and the minus RNAi cells died after 3 days (Figure 3.12.1.1 C). Surprisingly, the RNAi cells survived and showed similar differentiation kinetics to the cis-aconitate treated control (Figure 3.12.1.1 C); with expression of the procyclic form surface coat protein EP within 6 hours after the temperature shift and both PIP39 and EP proteins reaching wild type levels after 3 days when the cell number started to increase (Figure 3.12.1.1 C, D). The cell population treated with cis-aconitate, however, retained some RBP10 expression until day two (Figure 3.12.1.1 C, D); this could be a result of a few bloodstream form cells persisting, then dying after three days.

To confirm that the cells were *bona fide* procyclic forms, the morphology of the cells was further analysed using indirect immunofluorescence (a blind analysis was done; where Prof. Christine Clayton did the quantification for all the images that I recorded). We looked for morphological changes consistent with differentiation from bloodstream to the procyclic forms; such as the surface coat exchange from VSG to EP/GPEET, expression of PAD1 (stumpy form marker), and kinetoplast repositioning from posterior end in bloodstream to a midway position between the nucleus and the posterior end in procyclic forms. After *RBP10* RNAi for 17 hours, cells retained the long slender morphology, and PAD1, EP1 and GPEET could not be detected. However, upon incubation in MEM medium at 27^{0}C for three days, more than 80% of the cells became GPEET positive and their kinetoplast was repositioned at a midway position between the nucleus and the posterior end; characteristic of procyclic cells (Figure 3.12.1.2 A, C). Similar results were obtained for the positive control cells induced to differentiate using cis-aconitate (Figure 3.12.1.2 A, D).

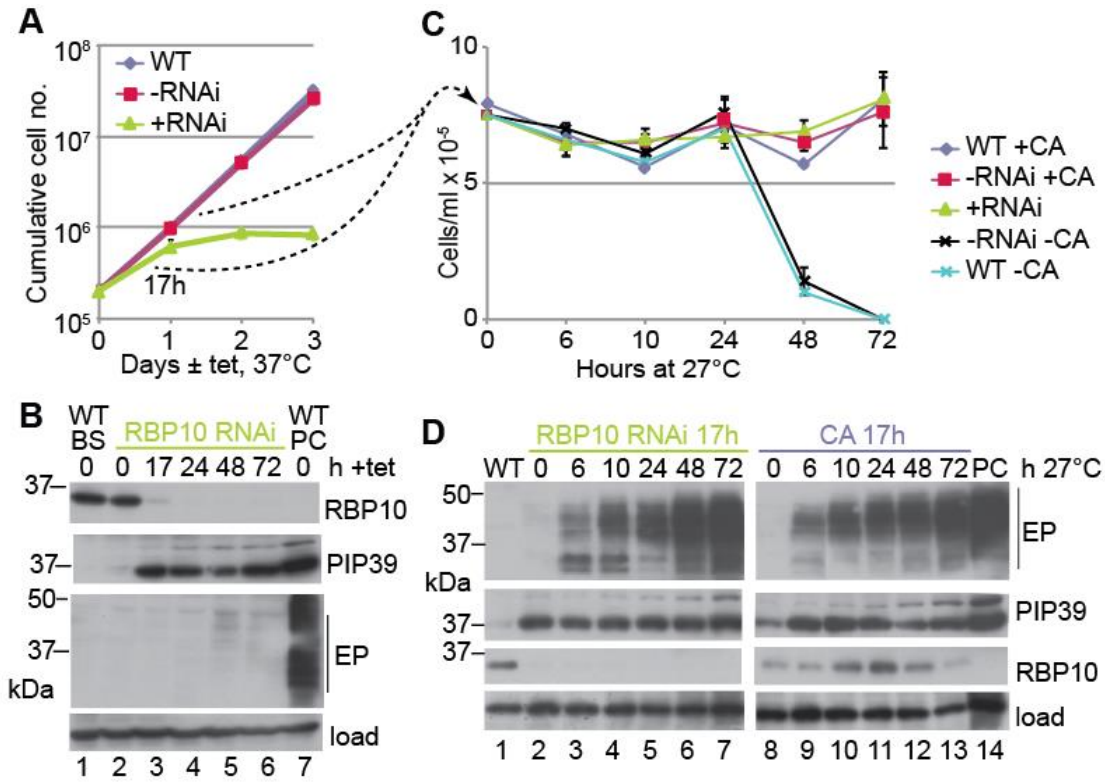

Figure 3.12.1.1 RBP10 depletion primed bloodstream form cells to differentiate to procyclic forms. **A)** Growth curve after RBP10 RNAi. **B)** Expression of RBP10, PIP39 and EP1 for samples from (A). **C)** Cell densities after transfer to procyclic medium at 27ºC for cells transformed by 17 hours RBP10 RNAi (+RNAi) or 6 mM cis-aconitate (+CA): the untreated ones (WT-CA & -RNAi-CA) served as negative control. **D)** Protein expression from (C). (Figure modified by C. Clayton)

In bloodstream forms the kinetoplast is very near the posterior cell tip, so that the ratio of distance from the posterior end to the kinetoplast (P-K) and the distance from kinetoplast to the nucleus (K-N) is much smaller than in procyclic forms. We used this ratio (P-K/K-N) to compare the kinetoplast repositioning after differentiation mediated by *RBP10* RNAi or cis-aconitate. The cells generated by *RBP10* RNAi or cis-aconitate showed a similar pattern of kinetoplast positioning; this was much closer to that of established procyclic form culture, and it differed significantly from the starting bloodstream form culture where the ratio was much smaller (Figure 3.12.1.2 B). In both *RBP10* RNAi and cis-aconitate converted cells, ~10% of the cells retained the slender morphology and some were GPEET negative (Figure 3.12.1.2 C, D).

Based on these results, we suggest RBP10 is a major negative regulator of differentiation in bloodstream forms. It acts by repressing mRNAs normally enriched in the procyclic forms. Loss of this regulation is lethal and gives rise to cells primed to differentiate from bloodstream to procyclic forms; this occurs in absence of cis-aconitate but is dependent on temperature reduction to 27^0C.

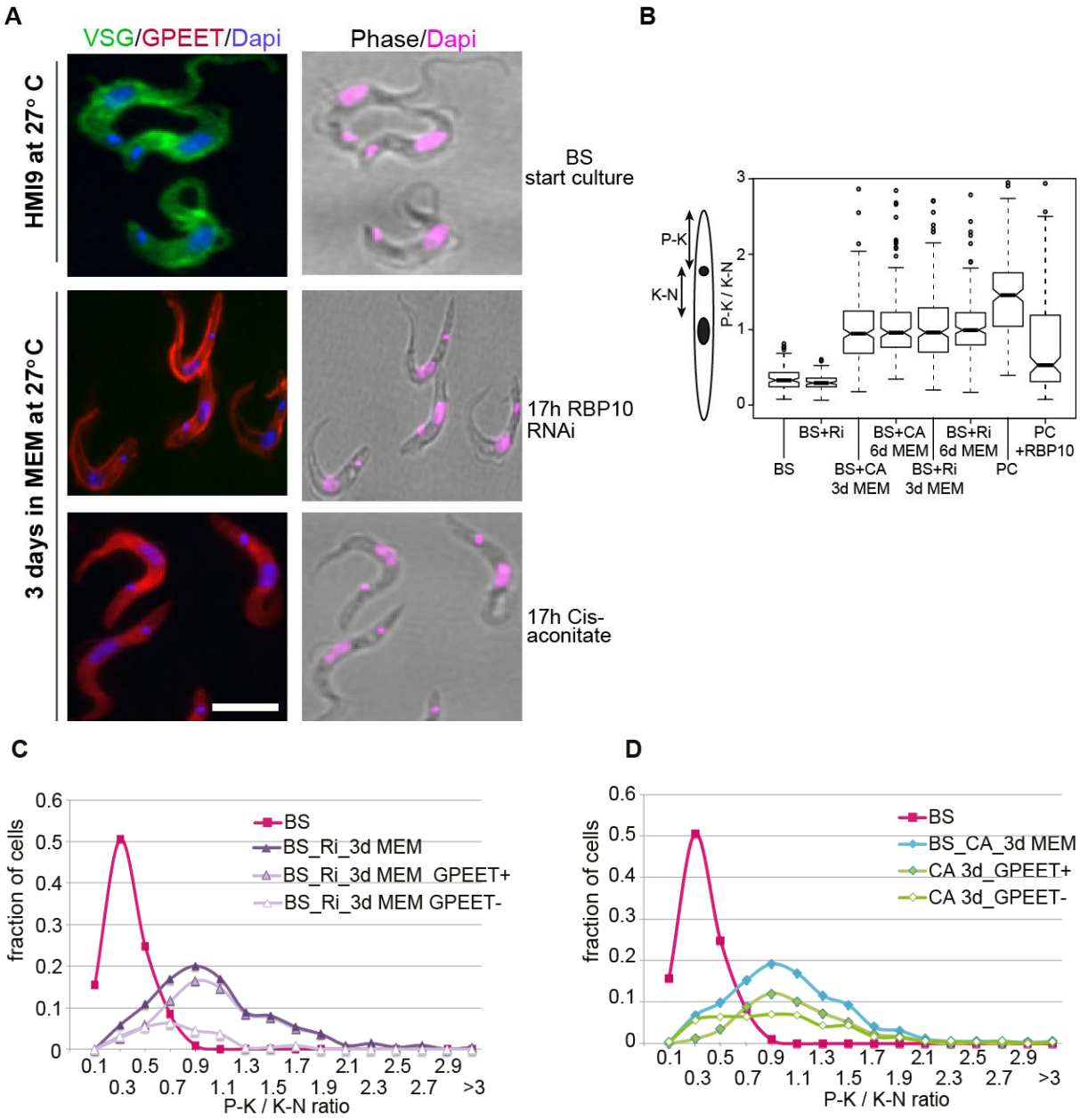

Figure 3.12.1.2 RBP10 depleted bloodstream form cells survive after transfer to MEM medium at 27ºC. **A**) Surface coat exchange from VSG to GPEET after *RBP10* RNAi for 17h or cis-aconitate treatment then transfer to MEM medium at 27ºC for 3 days. **B**) Kinetoplast repositioning estimated as a ratio of the distance from the posterior end to the kinetoplast (P-K) and the distance from kinetoplast to the nucleus (K-N); analysis was done for cells depleted for RBP10 for 17h or cis-aconitate treated, also for cells taken after day 3 and 6 in MEM medium at 27ºC. **C-D**) The distribution of the estimated ratio (P-K/K-N) with the percentage of the GPEET positive cells included. BS=bloodstream; BS+Ri=BS after 17h of RBP10 RNAi; BS+CA= BS after 17h cis-aconitate treatment. (Figure B,C,D were made by C. Clayton)

3.12.2 Procyclic to bloodstream form conversion

Depletion of RBP10 primes the bloodstream forms cells to differentiate to procyclic forms. I then asked, could over expression of RBP10 in procyclic forms convert the cells to bloodstream forms? To test this, RBP10-myc was ectopically expressed in a procyclic cell line derived from EATRO1125 strain. As expected, prolonged expression of RBP10 significantly inhibited cell growth, and caused reduction of EP procyclin (Figure 3.12.2.1 A, B). After 2 days the cells had increased expression of bloodstream form markers hnRNPH, TAO and native RBP10 (Figure 3.12.2.1 B, C). VSG mRNA was detected within 24 hours (Figure 3.12.2.1 C) and VSG protein after 2 days (Figure 3.12.2.2).

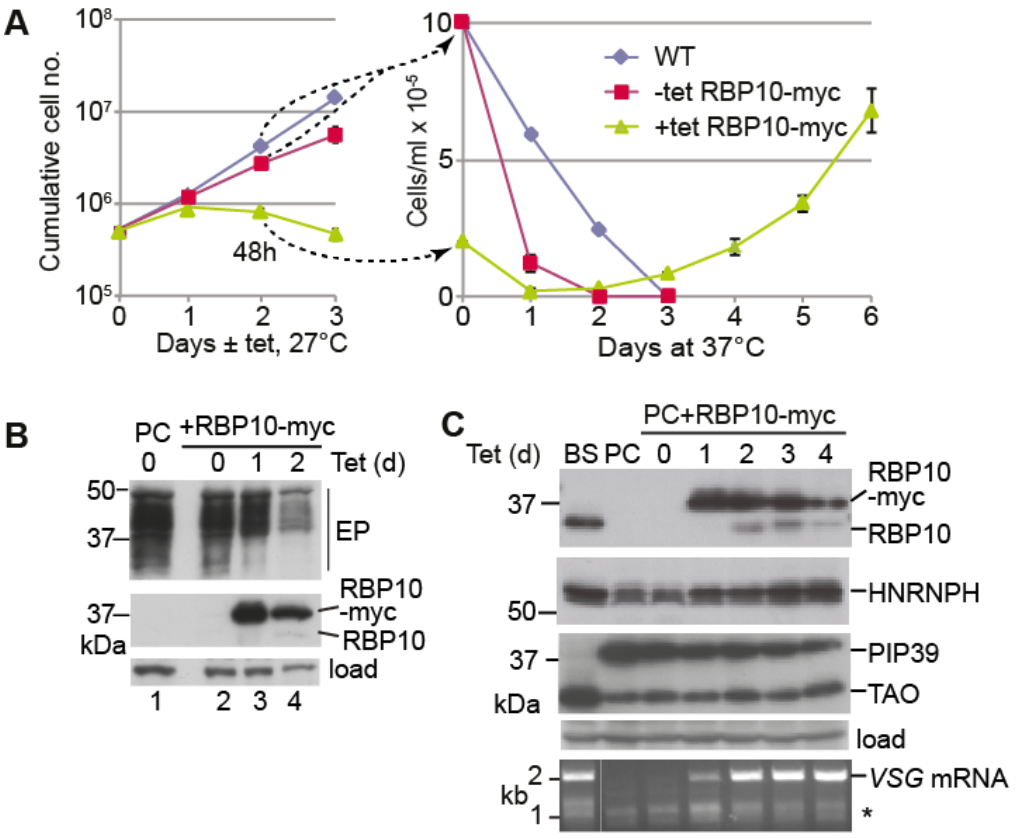

Figure 3.12.2.1 RBP10 expression triggers procyclic forms to differentiate to bloodstream forms. **A**) Left chart: Growth curve after RBP10-myc induction; Right chart: Cell densities after transfer to bloodstream medium at 37ºC; the samples taken are highlighted on the left chart; no tetracycline was added once the transfer was made. Uninduced (-tetRBP10-myc) and wild type (WT) cells served as controls. **B-C**: Expression of EP1, TAO, hnRNPH, PIP39 and native RBP10 after over expression of RBP10 for 4 days in procyclic cells; detection of *VSG* mRNAs by RT-PCR is also shown(C). (Figure modified by C. Clayton)

FACs analysis confirmed half of the cells had reduced or no EP procyclin and 16% cross-reacted with anti-VSG117 (Figure 3.12.2.2 A, C); some of the cells had already acquired bloodstream form morphology (Figure 3.12.2.2 B; Figure 3.12.2.3 A). Interestingly, after 2 days of RBP10 expression and transfer to 37ºC in HMI-9 medium some cells survived and started to divide as bloodstream forms (Figure 3.12.2.1 A right chart); with >60% of the cells being VSG positive after six days as determined by immunofluorescence using anti-VSG117 (Figure 3.12.2.3 B). The cell lacking RBP10 as well as the wild type procyclic forms died after 3 days confirming there was no prior bloodstream form cells in the population.

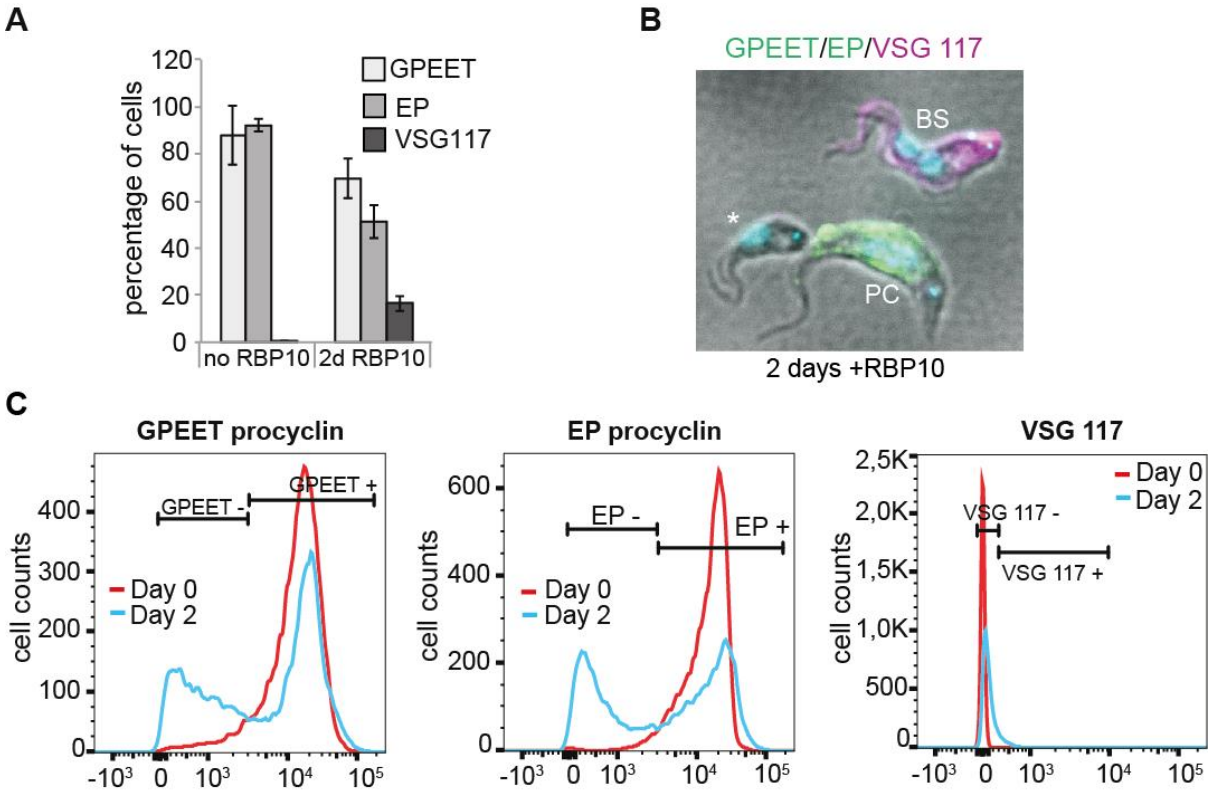

Figure 3.12.2.2 Over expression of RBP10 in procyclics cells leads to loss of surface coat procyclins. **A)** Percentage of cells expressing EP1, GPEET and VSG117 at the surface as determined by FACs analysis. Values represent means from four biological replicates. **B**) Image showing a mixed cell population after RBP10 expression for two days; PC=procyclic form cell with reduced EP procyclin, BS= bloodstream form cell positive for VSG117, * unusual cell with bloodstream form like morphology. **C**) Representative FACs plots showing expression levels of EP, GPEET and VSG117. (Figure modified by C. Clayton)

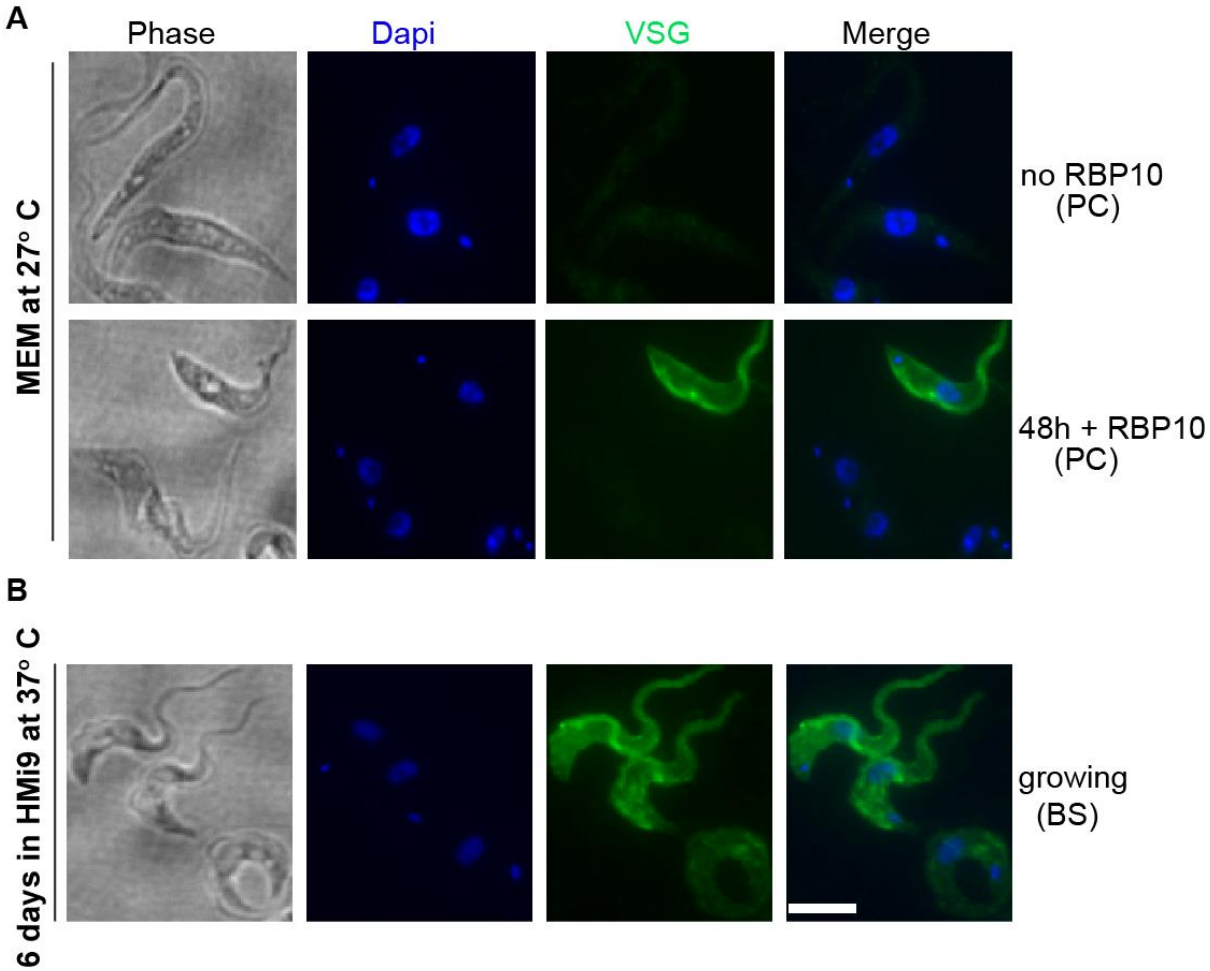

Figure 3.12.2.3 Some procyclic cells expressing RBP10 converts to bloodstream forms. **A**) Procyclic cells with or without RBP10 expression single stained with anti-VSG117. **B**) Proliferating bloodstream form cells generated by over expression of RBP10 then transfer to HMI-9 medium at 37^0C for six days.

Since anti-VSG117 failed to stain some of the cells that had acquired bloodstream form morphology, transmission electron microscopy (TEM) was used to further confirm surface coat exchange from the thin EP procyclin coat to a very dense VSG coat. Similar to IFA and FACS data, TEM analysis detected VSG coat on a subset (~15%) of the differentiating cells (not shown).

3.12.3 Transcriptome analysis of differentiating procyclic cells

To determine the wider consequences of RBP10 expression in the differentiating procyclic cells, I analysed the transcriptome changes of the cells expressing RBP10 for 24 or 48 hours using RNA-seq; the cells lacking RBP10 served as control. Both experiments were done in duplicate (Figure 3.12.3.1). In addition, the transcriptome for the newly converted bloodstream form cells was examined to confirm their similarity to the wild type EATRO1125 cells; in this case only a single sample from each cell line was analyzed.

The PCA plot on Figure 3.12.3.1 shows the general overview of the RNA-seq data, and the reproducibility between the replicates. For the procyclic samples, the control clustered separately from the samples where RBP10 was expressed for 24 or 48 hours; clearly indicating significant differences in the transcriptomes. Interestingly, samples from newly converted bloodstream form were closer to wild type bloodstream form than to the differentiating cells or the control procyclics cells (Figure 3.12.3.1). Their transcriptomes also showed a strong correlation in comparison to the rest.

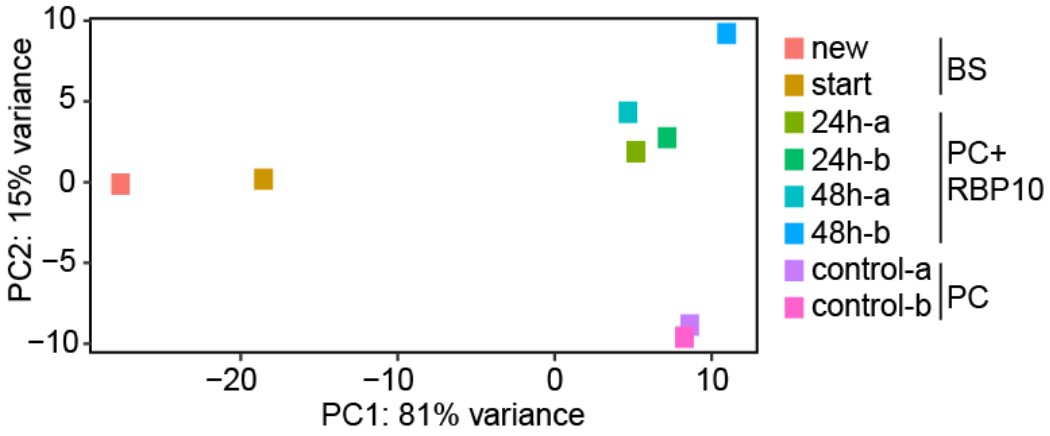

Figure 3.12.3.1 PCA plot showing transcriptomes variability. PC=procyclic form; BS=bloodstream form; new=newly converted BS; start=parental AnTat1.1 BS with an ectopic copy of RBP10.

As reported for the monomorphic cells, expression of RBP10 in procyclic forms caused widespread effects on the transcriptome (Figure 3.12.3.2-4). RBP10 expression for 24 hours led to 357 mRNAs being significantly differentially regulated (at least 2x with $p<0.05$). 242 were up-regulated while 115 mRNAs got down-regulated. 70% of those up-regulated are more abundant in bloodstream form

whereas 61% of those down-regulated are more abundant in the insect stage (Figure 3.12.3.2-4). The effect got stronger with prolonged expression of RBP10 for 48 hours (Figure 3.12.3.2-4). In this case, 1211 transcripts were significantly differentially regulated. 760 were up regulated (25% are more abundant in bloodstream form) while 451 were down regulated (22% are more abundant in procyclic forms). There was a strong positive correlation (R=0.87) between the transcriptome changes after 24 and 48 hours of RBP10 expression.

mRNAs encoding procyclic specific membrane proteins were significantly enriched within the down regulated transcripts after 24 or 48 hours of RBP10 expression (Table 4, Figure 3.12.3.3-4). In agreement with the Western blot result (Figure 3.12.2.1 A), EP mRNA was more than 5 fold decreased. Other gene classes within that list included those involved in the citric acid cycle and mitochondrion respiratory pathway (Table 4, Figure 3.12.3.3-4).

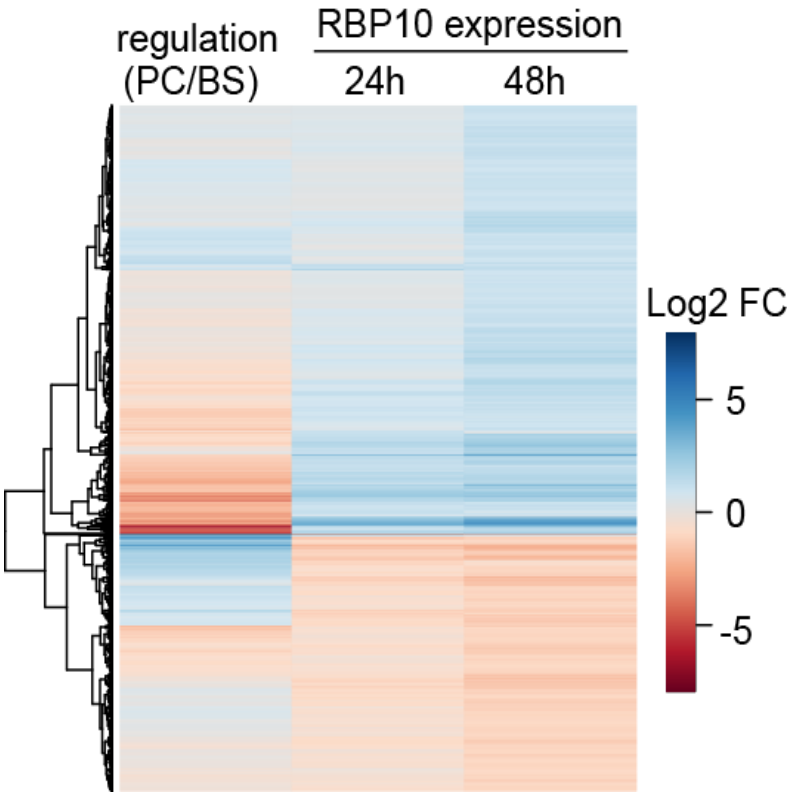

Figure 3.12.3.2 mRNA changes in differentiating procyclic cells. **A)** A clustered heatmap showing the relationship between effects of RBP10 expression and developmental regulation. Only mRNAs that were differentially regulated by at least 2 fold are shown. The developmental regulation data is a ratio of comparison between procyclic (PC) versus bloodstream (BS).

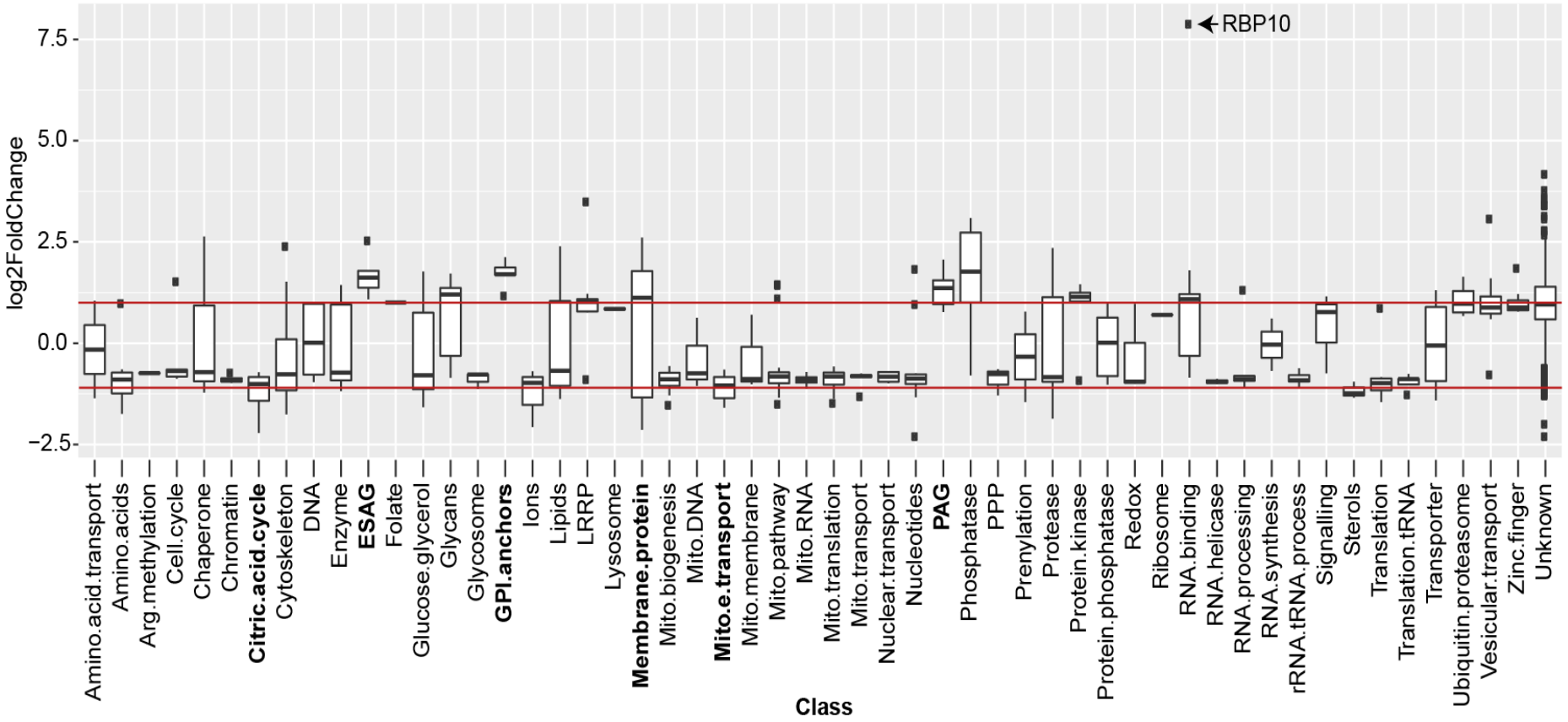

Figure 3.12.3.3 mRNA changes after RBP10 expression for 24h in procyclic forms. The effect of RBP10 expression on different trypanosome gene classes is shown using a box plot. The red line highlight 2 fold cut-off relative to the sample without RBP10. In bold are the gene classes found to be significantly over-represented in the subset of mRNAs which were either up or down regulated after RBP10 expression for 24h; the pvalue is shown on Table 5. The figure was generated using a custom tool for trypanosome transcriptomes, which was designed by Kevin Leiss.

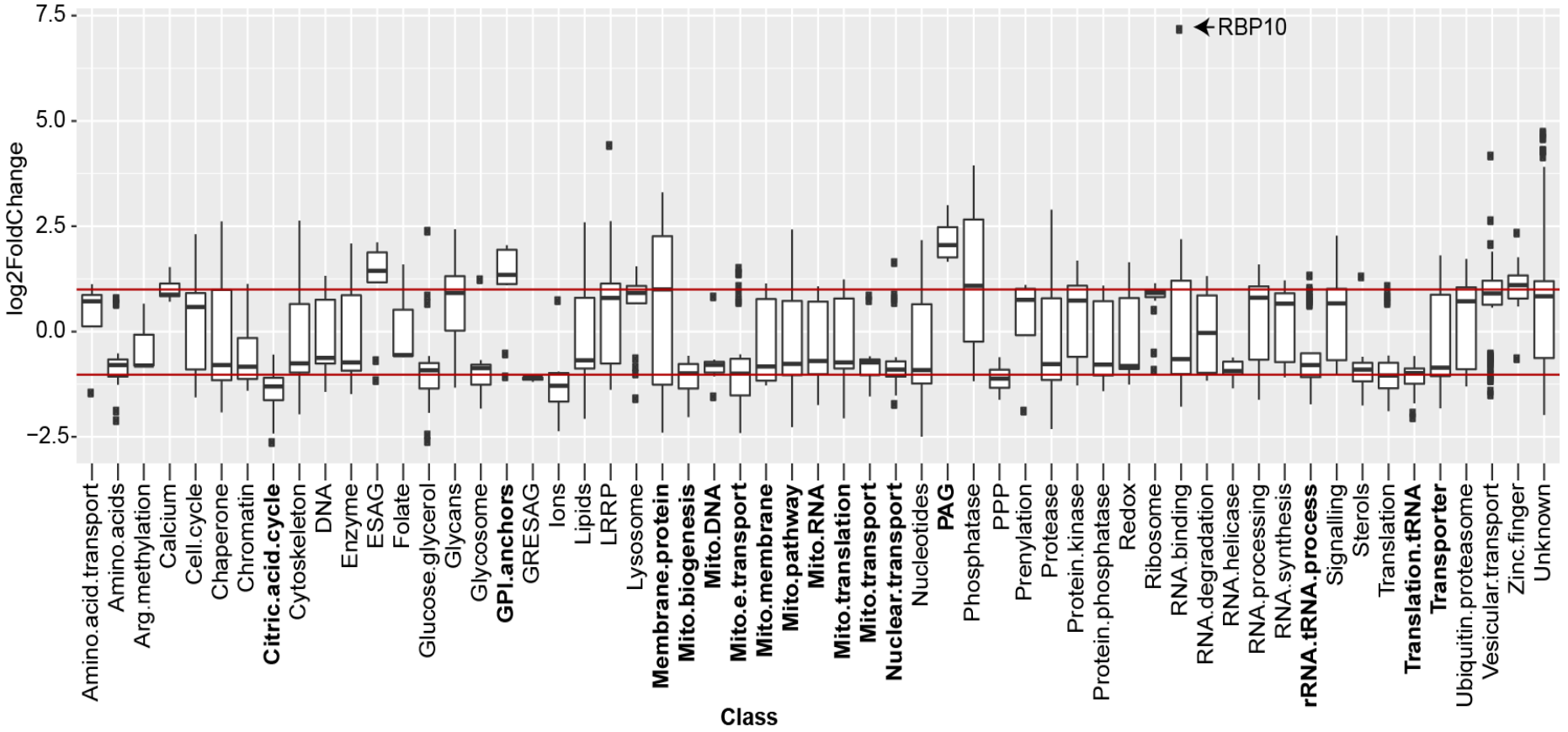

Figure 3.12.3.4 mRNA changes after RBP10 expression for 48h in procyclic forms. The analysis was done as in Figure 3.12.3.3.

Table 4. Gene class enrichment after RBP10 expression for one or two days. The number of the transcripts that met the cut-off (at least 2x, p<0.05) is shown. Gene enrichment was determined by Fishers exact test and p-value adjustment using Benjamini Hochberg method. The total number of the genes present in the genome for each class are shown; --- means the number is indicated in the 24h sample. Mito=mitochondrion, e=electron, PAG=procyclin-associated gene, ESAG=expression site associated gene

Gene Class	Adjusted p value	No. of Genes affected	No. of genes in the genome	Sample
up-regulated				24h +RBP10
ESAG	9.67E-05	8	29	
GPI.anchors	9.89E-03	5	17	
Membrane.protein	9.67E-05	9	29	
PAG	6.66E-03	4	7	
Down-regulated				
Citric.acid.cycle	8.28E-05	6	20	
Membrane.protein	3.46E-04	6	---	
Mito.biogenesis	2.76E-02	4	29	
Mito.e.transport	3.07E-08	14	86	
up-regulated				48h +RBP10
GPI.anchors	0.01	9	---	
Membrane.protein	0.01	12	---	
Down-regulated				
Citric.acid.cycle	1.38E-06	13	---	
Membrane.protein	1.03E-02	8	---	
Mito.biogenesis	3.07E-02	7	---	
Mito.e.transport	2.83E-04	19	86	
Mito.pathway	4.16E-02	29	272	
Nuclear.transport	3.26E-02	9	47	
Nucleotides	2.81E-03	12	51	
Protease	1.23E-02	13	73	
rRNA.tRNA.process	4.40E-04	20	100	
Translation	2.83E-04	18	78	
Translation.tRNA	9.25E-04	10	29	

The RNA coding for terminal alternative oxidase (TAO) which was used as a differentiation marker in the Western blot on Figure 3.12.2.1, increased by two fold after 24h of RBP10 expression. RBP10 mRNA was >30x increased (Figure 3.12.3.3-4), ten other bloodstream form specific transcripts encoding RBPs were also up-regulated by ~2-5 fold. Consistent with acquisition of bloodstream form morphology, the differentiating cells showed increased expression of mRNA coding for membrane proteins which are normally abundant in bloodstream forms. For example invariant surface glycoprotein ISG65 mRNA was 10x upregulated, major surface protease MSP A and C mRNAs were increased by >3x while the procylic form MSP C was significantly decreased. Moreover, 5 mRNAs encoding protein required for GPI anchor biosynthesis were up regulated as well as 8 ESAG genes (Table 4). Surprisingly, mRNAs coding for glycolytic enzyme were mainly up-regulated after 24h of RBP10 expression but not on day two when bloodstream form cells were detected (Figure 3.12.3.3-4); possibly due to growth defect seen on day two (Figure 3.12.2.1 A**)**. mRNAs encoding proteins involved in translation appeared not to be significantly affected after 24 hours of RBP10 expression, however longer expression for two days clearly affected translation with mRNAs encoding different factors needed for general translation being enriched in the down regulated set of mRNAs (Figure 3.12.3.3-4, Table 4). Others that were down regulated and enriched only in the 48 hours dataset included chaperone, protease, nuclear transport and nucleotides gene classes (Table 4).

Despite only 16% of the cells having morphological signatures of bloodstream form cells after 48 hours of RBP10 expression, the transcriptome points out an on going procyclic to bloodstream differentiation process. Comparison between the transcriptomes of the differentiating cells and the transcriptomes of the trypanosomes found in the tsetse fly showed a weak positive correlation to those parasites found in the proventriculus as they differentiate towards the salivary gland. BARP mRNA which is strongly up regulated (12x) in one such life stage (epimastigote) was ~2 fold up regulated after RBP10 expression for 2 days, however BARP protein was not detected by Western blotting. This suggests that RBP10 meditated conversion does not follow the classical life cycle progression seen in the fly, rather the differentiating cells may be jumping straight to the metacyclic forms. Alternatively the cells could be differentiating backwards from procyclic to bloodstream forms.

3.12.4 Metacyclic *VSGs* identification

To find out whether the procyclic cells were going forward in the life cycle or switching backwards, I looked for the *VSG* transcripts expressed in the metacyclic forms. One of the unique features of metacyclic trypanosomes is the monocistronic transcription of their VSGs. The promoter driving the metacyclic VSG expression is located ~3 kb upstream of the active VSG, this is in contrast to the normal bloodstream forms where the promoter is ~ 60 kb upstream of the active VSG gene; normally separated by ESAG genes. To find out whether there were metacyclic *VSGs* expressed in the cell populations, I did a *de novo* assembly of the transcriptomes shown in Figure 3.12.3.1. The obtained contigs were then searched for a conserved motif present in the 3' UTR of all *VSG* mRNAs. 47 contigs ranging 709-4788 bp were found. Majority of them (41 out of the 47) were identified from procyclic cells expressing RBP10, the other six contigs were from the newly converted bloodstream and the starting bloodstream forms cultures. Consistent with the RT-PCR (Figure 3.12.2.1) result, no contigs containing the conserved VSG motif were found in uninduced procyclic cells. To determine whether the contigs from the 48h RBP10 cultures contained partial or full length VSG CDS, the sequences were subjected to a BLASTn analysis against the *T. brucei* genome at NCBI. The results were comparable between the replicates with multiple contigs having the same VSG hit. Out of the 47 contigs, 20 unique VSG hits were present; among the unique set 50% had complete VSG CDS while 45% had a near complete VSG (78-98% identical) CDS (Table 5). Interestingly, the metacyclic VSG MVSG5 [160] was present among the eight unique contigs obtained from procyclic cells expressing RBP10 for 24h or 48h (Table 5). To estimate the transcript levels for the unique set of VSGs, I re-analysed the transcriptome data and included the CDS sequences from the 20 VSGs (Table 5). The starting bloodstream form cells mainly expressed AnTat1.1 VSG transcript which accounted for ~5% of the total mRNA (Table 5). On the contrary, procyclic cells expressing RBP10 for two days showed >100x increased expression of eight different VSG transcripts (Table 5 highlighted in blue); which as well accounted for 4-5% of the total mRNA. Out of eight, only MVSG5 expression site was previously known to be present in the ETRO1125 genome [161], the other seven might be novel metacyclic *VSGs*.

To find out whether the other seven VSGs are putative metacyclic VSGs I looked for the presence of the metacyclic promoter. This was not possible using the transriptome assembly data because the promoter is not present in the primary transcript. Therefore, I assembled shotgun reads from the EATRO1125 genome (~30x coverage [148]). 18 contigs with at least 4 kb (the longest was ~17 kb) included sequence with resemblance to *VSG* mRNAs. 3 out of the 18 were detected after RBP10 expression in procyclic cells and had a complete *VSG* CDS (Table 5 highlighted in light green). Interestingly, in all three contigs a metacyclic promoter was present on the sequence upstream of the CDS. Multiple sequence alignment showed a high conservation with known metacyclic promoter sequences (Figure 3.12.4). A putative ESAG1 could be identified in two of the contigs. One sequence (~7.4 kb) was 99.4% identical to the previously reported metacyclic MVSG5 expression site [160, 161]; the other (~17 kb) had a putative ESAG9, and showed resemblance to ILTat 1.22 metacyclic expression site [162] but with a different VSG CDS.

Therefore, RBP10 expression for 2 days causes the procyclic cells to switch towards the metacyclic state.

Table 5 The raw read counts for 20 unique *VSG* CDS identified from the transcriptomes of EATRO1125 and procyclic cells expressing RBP10. Highlighted in light green are the three putative metacyclic *VSGs*. The eight *VSGs* detected after RBP10 expression are shown in blue. The RPMs of total VSG counts estimated based on the total RNA is shown.

Nearest VSG Accession No.	CDS_ size (NCBI)	CDS % identity	BS new	BS start	24h PC (A)	24h PC (B)	48h PC (A)	48h PC (B)	ctrl PC (A)	ctrl PC (B)	Metacyclic promoter?
KX700731.1 (AnTat1.1)	1521	89	38	1530446	2	5	737	280	4	46	na
KX699100.1	1527	100	333	559434	51	79	3898	1270	5	25	na
KX701110.1	1329	100	763	536	33780	126840	2011130	965390	266	125	yes
KX701187.1	1266	100	84	56	2200	7146	183965	105932	54	11	na
KX699860.1	1551	100	52	22	3389	11911	89263	38447	136	34	na
KX699226.1	1431	100	9	6	1042	3469	45387	17981	52	25	yes
AF259553.1 (MVSG5)	1554	100	9	6	709	2324	45237	31741	53	10	yes
KX698697.1	1473	100	145	83	99	132	13327	2820	3	1	na
KX699246.1	1392	100	164	116	88	134	12814	2763	3	2	na
KX700057.1	1428	78	1	4	211	916	10622	5951	5	3	na
KX700095.1	1512	87	51	10	19	54	3149	1029	2	0	na
KX699008.1	1605	98	144	14	14	22	1433	402	0	0	na
KX699224.1	1434	92	24	27	330	1016	792	300	192	116	na
KX699312.1	1350	100	231	193	33	100	714	250	3	1	na
KX700970.1	1416	87	50	21	3	4	502	75	2	0	na
KX698690.1	1488	100	107	55	14	9	435	20	2	0	na
KX700619.1	1620	95	37	523	3	4	116	46	3	0	na
KX699109.1	1518	83	653170	32	2	517	60	19	1	1	na
KX698657.1	1521	47	189	17	0	1	1	2	0	0	na
		Total	655601	2091601	41989	154683	2423582	1174718	786	400	
		RPM	19565	75401	1582	1943	42683	51171	21	14	
		%VSG	1.96	7.54	0.16	0.19	4.27	5.12	0.00	0.00	

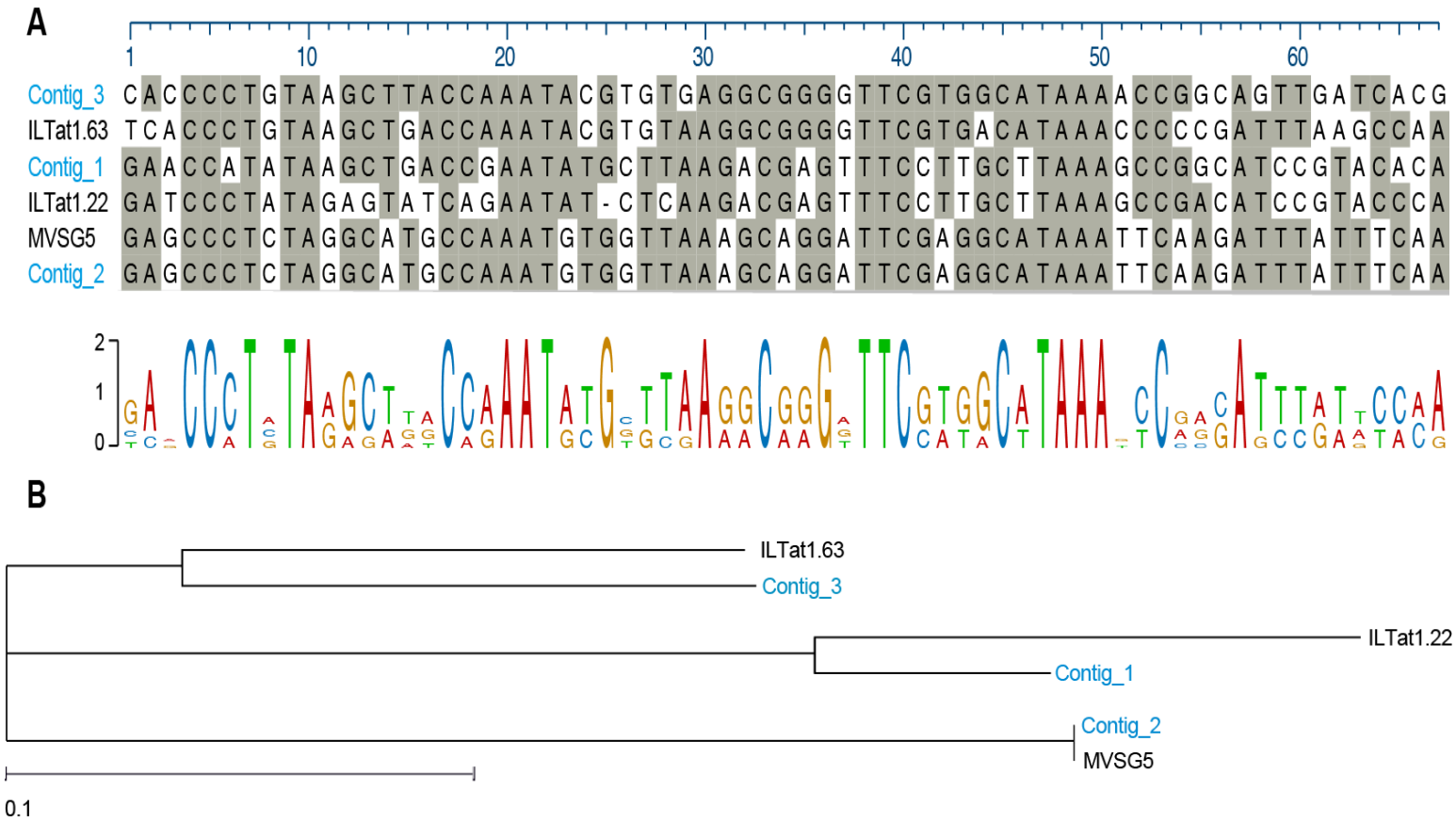

Figure 3.12.4 Multiple sequence alignment of the identified promoters. **A**) Sequence conservation with known metacyclic promoters for IlTat1.63, ILTat1.22 and MVSG5 metacyclic VSGs. **B**) A phylogenetic tree showing distances of the alignment.

4 Discussion

Previous study showed RB10 is a cytoplasmic RNA binding protein that is essential in bloodstream form trypanosome [91]. RBP10 expression is tightly regulated during life cycle transitions. In the stumpy and the insect stage procyclic forms RBP10 protein is undetectable [91, 105, 131], RBP10 mRNA is however up regulated in the life stage forms present in the salivary gland of the tsetse fly [163]. Alteration of RBP10 expression by RNAi in bloodstream form or over expression in procyclics cells caused profound transcriptome changes [91] mimicking early differentiation events between the two life stages. During my PhD, I investigated how RBP10 acting as a repressor mediates such changes. In addition, the role of RBP10 in differentiation commitment was examined.

4.1 RBP10 is a repressor protein

When the C-terminus of RBP10 is attached to a reporter mRNA, the mRNA is removed from actively translating polyribosomes resulting to mRNA degradation and translation repression. We speculated RBP10 acts by associating with the mRNA degradation machinery, however, we obtained no convincing evidence for this hypothesis. Yeast two hybrid screening and TAP purification using RBP10 failed to identify such a link. The only possible candidate from RBP10 TAP was CAF40, but RBP10 was not detected in the mass spec analysis after V5-CAF40 pull-down. Alternatively, the degradation machinery may be recruited indirectly via RBPs associating with RBP10. Interesting candidates found by yeast two hybrid included DRBD11, DRBD7, ZC3H13, ZC3H15, and three hypothetical proteins (Tb927.11.14220, Tb927.6.5010, Tb927.8.910). All were unique to RBP10 two hybrid screen, and they act as repressor of gene expression when tethered to a reporter mRNA. Six of the interacting partners bind to mRNAs in bloodstream form [82] and in case of ZC3H15, interaction with deadenylase CAF1 in a yeast two hybrid has been reported [82].

For the non RBPs, 4E-IP, a major repressor protein when tethered, [82, 83] was also present. Convincingly, the 4E-IP yeast two hybrid screen identified RBP10 as one of its interacting partners. The interaction between 4E-IP and RBP10 was confirmed in

trypanosomes as well, and only a very small proportion of the proteins associated with each other. Knock out of 4E-IP is however possible in bloodstream forms (unpublished result from our lab) which suggests that it is not a key contributor to the biological function of RBP10 in bloodstream forms. Follow up studies on the identified putative partners of RBP10 will help to understand its mechanism of action.

The repressive C-terminal (88 residues) fragment of RBP10 is rich in polar (31%) and hydrophobic (34%) amino acids. Deletion of this region prevents aggregation of RBP10 *in vitro* (Bin Liu and Igor Minia, unpublished). So far there is no evidence showing aggregation of RBP10 in bloodstream form trypanosomes where it is normally expressed. However, ectopic expression of eYFP tagged RBP10 in procyclic forms [164] showed RBP10 colocalize with processing body (P-body) markers DHH1 and SCD6, which further support the hypothesis that RBP10 acts as a repressor of gene expression. In mammalian cells and yeast, P-bodies contain proteins involved in mRNA degradation and translation repression as well as the silenced mRNAs [67].

4.2 RBP10 affects developmental regulation of many mRNAs bound by it

Regulation of trypanosome gene expression is important especially during life cycle transitions. One well studied differentiation step is from bloodstream to procyclic form. Evidence from transcriptome and proteome analyses during differentiation [105, 131, 165] point out a well coordinated series of events including change of the surface coat proteins, mitochondrion activation and switching of the energy metabolism from glycolysis to oxidative phosphorylation. Transcripts encoding proteins involved in these pathways are normally repressed in long slender bloodstream form. We speculated the negative regulation in bloodstream form is facilitated by RBP10. This is strongly supported by several facts: first, RBP10 acts as a repressor when tethered to mRNA. Second, RBP10 is strongly down regulated during differentiation and is absent in procyclic forms. Third, RNAi targeting RBP10 in bloodstream form mimic changes seen early during differentiation to procyclic forms; such as increase in expression of mRNAs encoding procyclic surface coat EP procyclin, regulatory proteins ZC3H22, ZC3H20 and several proteins involved in different mitochondrion specific pathways. Lastly, RBP10 co-purifies many mRNAs

that are normally up regulated in the procyclic forms and/or show increased expression during differentiation; this includes more than one third of the transcripts up-regulated after RBP10 RNAi.

The motif $UA(U)_6$ found to be enriched in the *3' UTRs* of transcripts bound by RBP10 is present in a wide selection of procyclic specific mRNAs. It was suggested more than ten years ago that the U-rich motif could be required for widespread developmental regulation of procyclic specific mRNAs [166]; similar to the destabilising AU-rich elements found in many mammalian mRNAs. In *T. brucei* bloodstream forms, the U-rich motif is already implicated in the negative regulation of mRNAs encoding EP procyclin [155], the B isoform of phoshoglycerate kinase [167], pyruvate phosphate dikinase [167] and cytochrome oxidase subunit COX V [166]. This is the first study to identify a bloodstream form *trans* acting factor that binds the motif and negatively regulates a significant number of procyclic specific mRNAs including some of those previously studied. Hence, RBP10 is a key regulator which is required for the maintenance of the bloodstream form life stage. We suggest it acts by recognizing the $UA(U)_6$ motif present in many procyclic specific mRNAs targeting them for translation repression and mRNA degradation.

In other eukaryotes, the deadenylation complex CCR4-CAF1-NOT triggers mRNA degradation; this is directed by RNA binding proteins (RBPs) which interact with the mRNA degradation machinery. For example yeast Mpt5p (RBP of the PUF family) recruits Pop2p (CAF1), the RNA helicase Dhh1p and decapping activator Dcp1 to heme oxygenase (HO) transcripts, which promote *HO* mRNA degradation [168]. The zinc finger protein tristetrapolin (TTP) mediates degradation of many AU-rich element (ARE) containing mRNA in mammalian cells by interacting with different enzymes involved in mRNA degradation pathways [169].

The mode of action of RBP10 appears not to involve the recruitment of the CAF1-NOT complex since tandem affinity purification using TAP tagged RBP10 failed to co-purify the machinery involved in mRNA degradation. However, the interaction with RBP10 may be transient hence not detectable using the TAP approach. This can be improved by adding an *in vivo* cross-linking step to freeze both weak and transient interactions taking place in intact cells before lysis.

Not all mRNAs that changed in abundance after RBP10 RNAi were bound by RBP10. This could be as a result of indirect effects from other regulatory proteins; the up regulated RBPs include ZC3H22 and ZC3H20. ZC3H22 is essential in procyclic forms [105] and acts as a negative regulator [82] when tethered to a reporter mRNA. Increased ZC3H22 expression therefore will result in a decrease in its target mRNAs. ZC3H20 is more abundant in procyclic forms where it is known to bind and stabilize two mRNAs [96]. Interestingly, one of ZC3H20 targets increased after RBP10 RNAi and was not bound by RBP10; this could be due to the increased levels of ZC3H20 after RBP10 RNAi.

For the transcripts that were bound by RBP10 but did not increase after RBP10 RNAi, we speculate other proteins dominate the regulation. Trypanosome mRNAs have a median 3'UTR length of 300nt. There is therefore space to bind multiple RNA binding proteins (RBPs), which can further interact with other proteins. As a result, the fate of the mRNA is determined by the ensemble of all associated RBPs.

RBP10 contain a single RRM domain, which probably is not sufficient to bind the eight nucleotide long UAUUUUUU motif present in many of its target mRNAs. The RRM is a sequence of about 90 amino acids, and consists of a four-stranded antiparallel β-sheet packed against two α-helices [170]; this forms a platform that predominantly binds single-stranded RNA. Each RRM domain can bind 4-6 nucleotides [171]. In the well studied mammalian RRM containing proteins, the presence of more than one RRM domain confers increased affinity and sequence specificity [172]. I have showed RBP10 self-interact in a yeast two hybrid assay, it is conceivable that RBP10 dimerizes as a homodimer; this possibly improves binding specificity of RBP10 since it has a single RRM domain. Furthermore, the presence of multiple copies of the motif within a single transcript raises the probability of binding and could allow more than two RBP10 to bind.

4.3 Role of RBP10 in differentiation commitment

A PTP1/PIP39 phosphatase cascade [152] controls the differentiation from bloodstream to procyclic forms via an intermediate stage known as the stumpy form. Up regulation of phosphorylated PIP39 and its relocation to the glycosome is required for the differentiation; this is promoted by temperature reduction plus incubation with the differentiation trigger cis-aconitate.

The PTP1 and PIP39 mRNAs were neither bound by RBP10 nor affected by depletion of RBP10 in bloodstream forms. However, the PIP39 protein level was strongly increased after RBP10 RNAi; which is a key signature for cells differentiating to procyclic forms. This finding plus the fact that alteration of RBP10 expression in procyclic forms causes reverse developmental regulation of many mRNAs led us to speculate a biological function in differentiation commitment. Indeed, pleomorphic bloodstream forms cells depleted of RBP10 differentiated to procyclic forms after transfer into procyclic growth media and incubation at 27ºC. Therefore, depletion of RBP10 primes bloodstream form cells to differentiate to procyclic forms in absence of differentiation trigger cis-aconitate. This could be as a result of metabolic changes after RBP10 RNAi that triggers downstream signaling pathways involved in differentiation; one example is increased expression of PIP39 protein. Other regulatory factors implicated in bloodstream to procyclic form differentiation include two transcripts encoding protein kinases, RDK1 and NRKA, that were bound but not regulated by RBP10. Similar to PIP39, their protein levels might have changed as part of a regulatory cascade caused by the loss of RBP10. RDK1 is essential in bloodstream form [122] hence it should not be regulated by RBP10 that acts as a repressor. NRKA [105] is up regulated after the onset of the stumpy form differentiation. Similar to RBP10, depletion of RDK1 primes bloodstream forms to differentiate to procyclic forms [122]. More than 20 mRNAs show similar changes after depletion of either RDK1 or RBP10, 16 of these are both bound and regulated by RBP10. Surprisingly, RBP10 transcript was not affected after RDK1 RNAi, however, the protein level and phosphorylation status of RBP10 were not determined.

The connection between differentiation signaling pathways and trypanosome gene expression is not well known. In the absence of transcription initiation control in *T. brucei*, RBP10 could be one of the effector molecules targeted by the differentiation signaling pathways. Understanding how RBP10 is regulated is the next major question. Regulation could be at the level of mRNA or protein. The very long (~7.5 kb) 3' UTR of RBP10 possibly harbours multiple regulatory elements as well as secondary structures. Recruitment of different mRNP complexes via such elements could influence stability and/or translation of RBP10 mRNA in a life stage specific manner; it won't be surprising that RBP10 3' UTR acts as a sensor during differentiation. On going studies (being done by Larissa Melo Do Nascimento) aim to map the elements required for the negative regulation of RBP10 mRNA in procyclic forms. The long-term goal is to identify the responsible *trans* acting factors. In terms of protein, RBP10 is known to be phosphorylated (four sites) in bloodstream forms [173]; the kinase involved is not known. Also, it remains to be determined whether those modifications regulate the biological function of RBP10. This could be tested using conditional knock out in the presence of either wild type RBP10 or mutants lacking the phosphorylation sites.

The conversion from procyclic form back to the mammalian infective metacyclic forms occurs naturally in the tsetse fly. It involves complex life cycle transitions with more than five different life stage forms being generated [37]; most of these life stages cannot be cultured *in vitro*. Strikingly, over expression of RBP10 in EATRO1125 derived procyclic cells for two days converted a subset of the cells to bloodstream forms. As expected, mRNAs enriched in the procyclic forms were strongly down regulated after RBP10 expression; including most of RBP10 mRNA targets. Secondary effects cannot be ruled out since growth was also affected. The transcriptome of the differentiating cells showed a weak positive correlation to the transcriptome of the parasites found in the tsetse fly, specifically those migrating from the proventriculus towards the salivary gland. One epimastigote stage marker BARP [42] which is twelve fold up regulated (at the mRNA level) relative to the procyclic forms [163] was only increased by about two fold after RBP10 expression. BARP protein was undetectable as were cells with epimastigote morphology; perhaps they constituted less than one percent of the cell population making it hard to find them. This is in contrast to what was observed in case of over expression of

RNA binding protein RBP6 [92]; where RBP6 promoted *in vitro* generation of epimastigotes within a day and mature infective metacyclic forms in five days [92]. It is likely that RBP10 mediated conversion jumps straight to the metacyclic forms hence does not follow the natural life cycle. This is supported by the detection of the transcripts encoding metacyclic VSGs in the differentiating cells expressing RBP10 for 24 and 48 hours. Interestingly, depletion of RBP10 in procyclic cells over expressing RBP6 leads to epimastigote cells but generation of metacyclic forms is blocked (unpublished result from C. Tschudi lab). This demonstrates that RBP10 is downstream of RBP6, and helps to define the long slender bloodstream form life stage.

When RBP10 is absent in bloodstream forms, the cells can only survive as procyclic forms. Conversely, expression of RBP10 in procyclic forms converts some of the cells to bloodstream form. We suggest RBP10 acts as a regulatory switch controlling *Trypanosoma brucei* differentiation.

5 References

1. Bruce D: Preliminary report on the tsetse fly disease or nagana, in Zululand. Durban: Bennett & Davis; 1895.
2. Dutton JE: Preliminary note upon a trypanosome occurring in the blood of man. Thompson Yates Lab Rep 1902, 4:455–468.
3. Castellani A: On the Discovery of a Species of Trypanosoma in the Cerebro-Spinal Fluid of Cases of Sleeping Sickness. Proceedings of the Royal Society of London 1902, 71(467-476):501-508.
4. Kleine: Positive Infektionsversuche mit *Trypanosoma brucei* durch Glossina palpalis. Dtsch med Wochenschr 1909, 35(11):469-470.
5. Franco JR, Simarro PP, Diarra A, Ruiz-Postigo JA, Jannin JG: The journey towards elimination of gambiense human African trypanosomiasis: not far, nor easy. 2014.
6. Franco JR, Simarro PP, Diarra A, Jannin JG: Epidemiology of human African trypanosomiasis. Clinical epidemiology 2014, 6:257-275.
7. MacLean L, Reiber H, Kennedy PGE, Sternberg JM: Stage Progression and Neurological Symptoms in *Trypanosoma brucei rhodesiense* Sleeping Sickness: Role of the CNS Inflammatory Response. Plos Neglect Trop Dis 2012, 6(10):e1857.
8. Kennedy PGE: Diagnosing central nervous system trypanosomiasis: two stage or not to stage? Transactions of The Royal Society of Tropical Medicine and Hygiene 2008, 102(4):306-307.
9. Kuepfer I, Hhary EP, Allan M, Edielu A, Burri C, Blum JA: Clinical Presentation of *T.b. rhodesiense* Sleeping Sickness in Second Stage Patients from Tanzania and Uganda. Plos Neglect Trop Dis 2011, 5(3):e968.
10. Kennedy PGE: Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). The Lancet Neurology 2013, 12(2):186-194.
11. Nagle AS, Khare S, Kumar AB, Supek F, Buchynskyy A, Mathison CJN, Chennamaneni NK, Pendem N, Buckner FS, Gelb MH et al: Recent developments in drug discovery for leishmaniasis and human African trypanosomiasis. Chemical reviews 2014, 114:11305-11347.
12. Priotto G, Kasparian S, Mutombo W, Ngouama D, Ghorashian S, Arnold U, Ghabri S, Baudin E, Buard V, Kazadi-Kyanza S et al: Nifurtimox-eflornithine combination therapy for second-stage African *Trypanosoma brucei gambiense* trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. The Lancet 2009, 374(9683):56-64.
13. Brun R, Don R, Jacobs RT, Wang MZ, Barrett MP: Development of novel drugs for human African trypanosomiasis. Future Microbiology 2011, 6(6):677-691.
14. Field MC, Horn D, Fairlamb AH, Ferguson MAJ, Gray DW, Read KD, De Rycker M, Torrie LS, Wyatt PG, Wyllie S et al: Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. Nature Reviews Microbiology 2017.
15. Isaac C, Ciosi M, Hamilton A, Scullion KM, Dede P, Igbinosa IB, Nmorsi OPG, Masiga D, Turner CMR: Molecular identification of different trypanosome species and subspecies in tsetse flies of northern Nigeria. Parasit Vectors 2016, 9:301.

16. Morrison LJ, Vezza L, Rowan T, Hope JC: Animal African Trypanosomiasis: Time to Increase Focus on Clinically Relevant Parasite and Host Species. Trends in Parasitology 2016, 32(8):599-607.
17. Kristjanson PM, Swallow BM, Rowlands GJ, Kruska RL, de Leeuw PN: Measuring the costs of African animal trypanosomosis, the potential benefits of control and returns to research. Agricultural Systems 1999, 59(1):79-98.
18. Desquesnes M, Dia ML: Mechanical transmission of *Trypanosoma vivax* in cattle by the African tabanid Atylotus fuscipes. Vet Parasitol 2004, 119(1):9-19.
19. Desquesnes M, Dia ML: Mechanical transmission of *Trypanosoma congolense* in cattle by the African tabanid Atylotus agrestis. Exp Parasitol 2003, 105(3-4):226-231.
20. Mihok S, Maramba O, Munyoki E, Kagoiya J: Mechanical transmission of *Trypanosoma spp*. by African Stomoxyinae (Diptera: Muscidae). Tropical medicine and parasitology : official organ of Deutsche Tropenmedizinische Gesellschaft and of Deutsche Gesellschaft fur Technische Zusammenarbeit (GTZ) 1995, 46(2):103-105.
21. Stephens NA, Kieft R, MacLeod A, Hajduk SL: Trypanosome resistance to human innate immunity: targeting Achilles' heel. Trends in parasitology 2012, 28(12):539-545.
22. Xong HV, Vanhamme L, Chamekh M, Chimfwembe CE, Van Den Abbeele J, Pays A, Van Meirvenne N, Hamers R, De Baetselier P, Pays E: A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. Cell 1998, 95(6):839-846.
23. Uzureau P, Uzureau S, Lecordier L, Fontaine F, Tebabi P, Homble F, Grelard A, Zhendre V, Nolan DP, Lins L et al: Mechanism of *Trypanosoma brucei gambiense* resistance to human serum. Nature 2013, 501(7467):430-434.
24. Kieft R, Capewell P, Turner CMR, Veitch NJ, MacLeod A, Hajduk S: Mechanism of *Trypanosoma brucei gambiense* (group 1) resistance to human trypanosome lytic factor. Proc Natl Acad Sci U S A 2010, 107(37):16137-16141.
25. Simpson L: The Mitochondrial Genome of Kinetoplastid Protozoa: Genomic Organization, Transcription, Replication, and Evolution. Annual Review of Microbiology 1987, 41(1):363-380.
26. Opperdoes FR, Borst P: Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: The glycosome. FEBS Letters 1977, 80(2):360-364.
27. Michels PAM, Bringaud F, Herman M, Hannaert V: Metabolic functions of glycosomes in trypanosomatids. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 2006, 1763(12):1463-1477.
28. Rudenko G, Bishop D, Gottesdiener K, Van der Ploeg LH: Alpha-amanitin resistant transcription of protein coding genes in insect and bloodstream form *Trypanosoma brucei*. The EMBO Journal 1989, 8(13):4259-4263.
29. Vaughan S, Gull K: The trypanosome flagellum. Journal of Cell Science 2003, 116(5):757.
30. Langousis G, Hill KL: Motility and more: the flagellum of *Trypanosoma brucei*. Nature reviews Microbiology 2014, 12(7):505-518.
31. Kohl L, Sherwin T, Gull K: Assembly of the Paraflagellar Rod and the Flagellum Attachment Zone Complex During the *Trypanosoma brucei* Cell Cycle. Journal of Eukaryotic Microbiology 1999, 46(2):105-109.

32. Simpson L, Thiemann OH, Savill NJ, Alfonzo JD, Maslov DA: Evolution of RNA editing in trypanosome mitochondria. Proc Natl Acad Sci U S A 2000, 97(13):6986-6993.
33. Stuart K: The RNA editing process in *Trypanosoma brucei*. Seminars in Cell Biology 1993, 4(4):251-260.
34. Robinson DR, Gull K: Basal body movements as a mechanism for mitochondrial genome segregation in the trypanosome cell cycle. Nature 1991, 352(6337):731-733.
35. Field MC, Carrington M: The trypanosome flagellar pocket. Nat Rev Micro 2009, 7(11):775-786.
36. Engstler M, Boshart M: Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in *Trypanosoma brucei*. Genes & development 2004, 18:2798-2811.
37. Ooi C-P, Bastin P, Hammarton TC, Ginger M: More than meets the eye: understanding *Trypanosoma brucei* morphology in the tsetse. 2013.
38. Vassella E, Probst M, Schneider A, Studer E, Renggli CK, Roditi I: 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):3986-3993.
39. Vassella E, Den Abbeele JV, Bütikofer P, Renggli CK, Furger A, Brun R, Roditi I: 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):615-626.
40. Urwyler S, Vassella E, Abbeele JVD, Renggli CK, Blundell P, Barry JD, Roditi I: Expression of Procyclin mRNAs during Cyclical Transmission of *Trypanosoma brucei*. PLoS Pathog 2005, 1(3):e22.
41. Rotureau B, Van Den Abbeele J: Through the dark continent: African trypanosome development in the tsetse fly. Frontiers in Cellular and Infection Microbiology 2013, 3:53.
42. Urwyler S, Studer E, Renggli CK, Roditi I: A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of *Trypanosoma brucei*. Mol Microbiol 2007, 63(1):218-228.
43. Van Den Abbeele J, Claes Y Fau - van Bockstaele D, van Bockstaele D Fau - Le Ray D, Le Ray D Fau - Coosemans M, Coosemans M: *Trypanosoma brucei spp*. development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis. 1999(0031-1820 (Print)).
44. Sharma R, Peacock L, Gluenz E, Gull K, Gibson W, Carrington M: Asymmetric Cell Division as a Route to Reduction in Cell Length and Change in Cell Morphology in Trypanosomes. Protist 2008, 159(1):137-151.
45. Tetley L, Vickerman K: Differentiation in *Trypanosoma brucei*: host-parasite cell junctions and their persistence during acquisition of the variable antigen coat. Journal of Cell Science 1985, 74(1):1.
46. Cross GA: Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. Parasitology 1975, 71(3):393-417.
47. Ferguson MA, Homans SW, Dwek RA, Rademacher TW: Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. Science 1988, 239(4841):753.
48. Cross GAM: Antigenic Variation in Trypanosomes. Proceedings of the Royal Society of London Series B Biological Sciences 1978, 202(1146):55.

49. Lythgoe KA, Morrison LJ, Read AF, Barry JD: Parasite-intrinsic factors can explain ordered progression of trypanosome antigenic variation. Proc Natl Acad Sci U S A 2007, 104(19):8095-8100.
50. Horn D: Antigenic variation in African trypanosomes. Molecular and Biochemical Parasitology 2014, 195(2):123-129.
51. Marcello L, Barry JD: 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):1344-1352.
52. Cross GAM, Kim H-S, Wickstead B: Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. Molecular and Biochemical Parasitology 2014, 195(1):59-73.
53. Fenn K, Matthews KR: The cell biology of *Trypanosoma brucei* differentiation. Curr Opin Microbiol 2007, 10(6):539-546.
54. Matthews KR: The developmental cell biology of *Trypanosoma brucei*. Journal of cell science 2005, 118(Pt 2):283-290.
55. Clayton C: The Regulation of Trypanosome Gene Expression by RNA-Binding Proteins. PLoS Pathog 2013, 9(11):e1003680.
56. Daniels J-P, Gull K, Wickstead B: Cell Biology of the Trypanosome Genome. Microbiology and Molecular Biology Reviews 2010, 74(4):552-569.
57. Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, Wang X, Dewell S, Cross GAM: Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes & Development 2009, 23(9):1063-1076.
58. Wright JR, Siegel TN, Cross GAM: Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*. Molecular and biochemical parasitology 2010, 172(2):141-144.
59. Lee JH, Nguyen TN, Schimanski B, Günzl A: Spliced Leader RNA Gene Transcription in *Trypanosoma brucei* Requires Transcription Factor TFIIH. Eukaryotic Cell 2007, 6(4):641-649.
60. Das A, Zhang Q, Palenchar JB, Chatterjee B, Cross GAM, Bellofatto V: Trypanosomal TBP Functions with the Multisubunit Transcription Factor tSNAP To Direct Spliced-Leader RNA Gene Expression. Molecular and Cellular Biology 2005, 25(16):7314-7322.
61. Gilinger G, Bellofatto V: Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. Nucleic Acids Research 2001, 29(7):1556-1564.
62. Jensen BC, Ramasamy G, Vasconcelos EJR, Ingolia NT, Myler PJ, Parsons M: Extensive stage-regulation of translation revealed by ribosome profiling of *Trypanosoma brucei*. BMC Genomics 2014, 15(1):911.
63. Clayton CE: NEW EMBO MEMBER'S REVIEW Life without transcriptional control? From fly to man and back again. EMBO J 2002, 21.
64. Jaé N, Wang P, Gu T, Hühn M, Palfi Z, Urlaub H, Bindereif A: Essential Role of a Trypanosome U4-Specific Sm Core Protein in Small Nuclear Ribonucleoprotein Assembly and Splicing. Eukaryotic Cell 2010, 9(3):379-386.
65. Mair G, Shi H, Li H, Djikeng A, Aviles HO, Bishop JR, Falcone FH, Gavrilescu C, Montgomery JL, Santori MI et al: A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. RNA 2000, 6(2):163-169.

66. Freire ER, Vashisht AA, Malvezzi AM, Zuberek J, Langousis G, Saada EA, Nascimento JDF, Stepinski J, Darzynkiewicz E, Hill K et al: eIF4F-like complexes formed by cap-binding homolog TbEIF4E5 with TbEIF4G1 or TbEIF4G2 are implicated in post-transcriptional regulation in *Trypanosoma brucei*. RNA 2014, 20(8):1272-1286.
67. Decker CJ, Parker R: P-Bodies and Stress Granules: Possible Roles in the Control of Translation and mRNA Degradation. Cold Spring Harbor Perspectives in Biology 2012, 4(9):a012286.
68. Kramer S: RNA in development: how ribonucleoprotein granules regulate the life cycles of pathogenic protozoa. Wiley Interdisciplinary Reviews: RNA 2014, 5(2):263-284.
69. Cassola A, De Gaudenzi JG, Frasch AC: Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. Mol Microbiol 2007, 65(3):655-670.
70. Fritz M, Vanselow J, Sauer N, Lamer S, Goos C, Siegel T N, Subota I, Schlosser A, Carrington M, Kramer S: Novel insights into RNP granules by employing the trypanosome's microtubule skeleton as a molecular sieve. Nucleic Acids Research 2015, 43(16):8013-8032.
71. Tucker M, Valencia-Sanchez MA, Staples RR, Chen J, Denis CL, Parker R: The Transcription Factor Associated Ccr4 and Caf1 Proteins Are Components of the Major Cytoplasmic mRNA Deadenylase in Saccharomyces cerevisiae. Cell 2001, 104(3):377-386.
72. Anderson JS, Parker RP: The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. The EMBO Journal 1998, 17(5):1497-1506.
73. Wang Z, Kiledjian M: Functional Link between the Mammalian Exosome and mRNA Decapping. Cell 2001, 107(6):751-762.
74. Hsu CL, Stevens A: Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. Molecular and Cellular Biology 1993, 13(8):4826-4835.
75. Milone J, Wilusz J, Bellofatto V: Characterization of deadenylation in trypanosome extracts and its inhibition by poly(A)-binding protein Pab1p. RNA 2004, 10(3):448-457.
76. Erben E, Chakraborty C, Clayton C: The CAF1-NOT complex of trypanosomes. Frontiers in Genetics 2013, 4:299.
77. Schwede A, Ellis L, Luther J, Carrington M, Stoecklin G, Clayton C: A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. Nucleic Acids Research 2008, 36(10):3374-3388.
78. Fadda A, Färber V, Droll D, Clayton C: The roles of 3′-exoribonucleases and the exosome in trypanosome mRNA degradation. RNA 2013, 19(7):937-947.
79. Schwede A, Manful T, Jha BA, Helbig C, Bercovich N, Stewart M, Clayton C: The role of deadenylation in the degradation of unstable mRNAs in trypanosomes. Nucleic Acids Research 2009, 37(16):5511-5528.
80. Gherzi R, Lee K-Y, Briata P, Wegmüller D, Moroni C, Karin M, Chen C-Y: A KH Domain RNA Binding Protein, KSRP, Promotes ARE-Directed mRNA Turnover by Recruiting the Degradation Machinery. Molecular Cell 2004, 14(5):571-583.

81. Chen C-Y, Gherzi R, Ong S-E, Chan EL, Raijmakers R, Pruijn GJM, Stoecklin G, Moroni C, Mann M, Karin M: AU Binding Proteins Recruit the Exosome to Degrade ARE-Containing mRNAs. Cell 2001, 107(4):451-464.
82. Lueong S, Merce C, Fischer B, Hoheisel JD, Erben ED: Gene expression regulatory networks in *Trypanosoma brucei*: insights into the role of the mRNA-binding proteome. Mol Microbiol 2016, 100(3):457-471.
83. Erben ED, Fadda A, Lueong S, Hoheisel JD, Clayton C: A Genome-Wide Tethering Screen Reveals Novel Potential Post-Transcriptional Regulators in *Trypanosoma brucei*. PLoS Pathog 2014, 10(6):e1004178.
84. Kolev NG, Ullu E, Tschudi C: The emerging role of RNA-binding proteins in the life cycle of *Trypanosoma brucei*. Cellular microbiology 2014, 16(4):482-489.
85. Wurst M, Robles A, Po J, Luu V-D, Brems S, Marentije M, Stoitsova S, Quijada L, Hoheisel J, Stewart M et al: An RNAi screen of the RRM-domain proteins of *Trypanosoma brucei*. Molecular and Biochemical Parasitology 2009, 163(1):61-65.
86. De Gaudenzi J, Frasch AC, Clayton C: RNA-Binding Domain Proteins in Kinetoplastids: a Comparative Analysis. Eukaryotic Cell 2005, 4(12):2106-2114.
87. Alsford S, Turner DJ, Obado SO, Sanchez-Flores A, Glover L, Berriman M, Hertz-Fowler C, Horn D: High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. Genome Research 2011, 21(6):915-924.
88. Hartmann C, Benz C, Brems S, Ellis L, Luu V-D, Stewart M, D'Orso I, Busold C, Fellenberg K, Frasch ACC et al: Small Trypanosome RNA-Binding Proteins TbUBP1 and TbUBP2 Influence Expression of F-Box Protein mRNAs in Bloodstream Trypanosomes. Eukaryotic Cell 2007, 6(11):1964-1978.
89. Das A, Morales R, Banday M, Garcia S, Hao L, Cross GAM, Estevez AM, Bellofatto V: The essential polysome-associated RNA-binding protein RBP42 targets mRNAs involved in *Trypanosoma brucei* energy metabolism. RNA 2012, 18(11):1968-1983.
90. Gupta SK, Kosti I, Plaut G, Pivko A, Tkacz ID, Cohen-Chalamish S, Biswas DK, Wachtel C, Waldman Ben-Asher H, Carmi S et al: The hnRNP F/H homologue of *Trypanosoma brucei* is differentially expressed in the two life cycle stages of the parasite and regulates splicing and mRNA stability. Nucleic Acids Research 2013, 41(13):6577-6594.
91. Wurst M, Seliger B, Jha BA, Klein C, Queiroz R, Clayton C: Expression of the RNA recognition motif protein RBP10 promotes a bloodstream-form transcript pattern in *Trypanosoma brucei*. Mol Microbiol 2012, 83:1048-1063.
92. Kolev NG, Ramey-Butler K, Cross GAM, Ullu E, Tschudi C: Developmental Progression to Infectivity in *Trypanosoma brucei* Triggered by an RNA-Binding Protein. Science (New York, NY) 2012, 338(6112):1352-1353.
93. Hendriks EF, Matthews KR: Disruption of the developmental programme of *Trypanosoma brucei* by genetic ablation of TbZFP1, a differentiation-enriched CCCH protein. Mol Microbiol 2005, 57.
94. Hendriks EF, Robinson DR, Hinkins M, Matthews KR: A novel CCCH protein which modulates differentiation of *Trypanosoma brucei* to its procyclic form. EMBO J 2001, 20.
95. Walrad P, Paterou A, Acosta-Serrano A, Matthews KR: Differential Trypanosome Surface Coat Regulation by a CCCH Protein That Co-

Associates with procyclin mRNA cis-Elements. PLoS Pathog 2009, 5(2):e1000317.
96. Ling AS, Trotter JR, Hendriks EF: A Zinc Finger Protein, TbZC3H20, Stabilizes Two Developmentally Regulated mRNAs in Trypanosomes. The Journal of Biological Chemistry 2011, 286(23):20152-20162.
97. Droll D, Minia I, Fadda A, Singh A, Stewart M, Queiroz R, Clayton C: Post-Transcriptional Regulation of the Trypanosome Heat Shock Response by a Zinc Finger Protein. PLoS Pathog 2013, 9(4):e1003286.
98. Singh A, Minia I, Droll D, Fadda A, Clayton C, Erben E: Trypanosome MKT1 and the RNA-binding protein ZC3H11: interactions and potential roles in post-transcriptional regulatory networks. Nucleic Acids Research 2014, 42(7):4652-4668.
99. Archer SK, Luu V-D, de Queiroz RA, Brems S, Clayton C: *Trypanosoma brucei* PUF9 Regulates mRNAs for Proteins Involved in Replicative Processes over the Cell Cycle. PLoS Pathog 2009, 5(8):e1000565.
100. Droll D, Archer S, Fenn K, Delhi P, Matthews K, Clayton C: The trypanosome Pumilio-domain protein PUF7 associates with a nuclear cyclophilin and is involved in ribosomal RNA maturation. Febs Letters 2010, 584(6):1156-1162.
101. Schumann Burkard G, Käser S, de Araújo PR, Schimanski B, Naguleswaran A, Knüsel S, Heller M, Roditi I: Nucleolar proteins regulate stage-specific gene expression and ribosomal RNA maturation in *Trypanosoma brucei*. Mol Microbiol 2013, 88(4):827-840.
102. Subota I, Rotureau B, Blisnick T, Ngwabyt S, Durand-Dubief M, Engstler M, Bastin P: ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation. Molecular Biology of the Cell 2011, 22(22):4205-4219.
103. Mani J, Güttinger A, Schimanski B, Heller M, Acosta-Serrano A, Pescher P, Späth G, Roditi I: Alba-Domain Proteins of *Trypanosoma brucei* Are Cytoplasmic RNA-Binding Proteins That Interact with the Translation Machinery. PLoS ONE 2011, 6(7):e22463.
104. Kramer S: Developmental regulation of gene expression in the absence of transcriptional control: The case of kinetoplastids. Molecular and Biochemical Parasitology 2012, 181:61-72.
105. Domingo-Sananes MR, Szöőr B, Ferguson MAJ, Urbaniak MD, Matthews KR: Molecular control of irreversible bistability during trypanosome developmental commitment. The Journal of Cell Biology 2015, 211(2):455-468.
106. Vassella E, Reuner B, Yutzy B, Boshart M: Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. Journal of Cell Science 1997, 110:2661-2671.
107. Waters CM, Bassler BL: QUORUM SENSING: Cell-to-Cell Communication in Bacteria. Annual Review of Cell and Developmental Biology 2005, 21(1):319-346.
108. Breidbach T, Ngazoa E, Steverding D: *Trypanosoma brucei*: in vitro slender-to-stumpy differentiation of culture-adapted, monomorphic bloodstream forms. Experimental Parasitology 2002, 101(4):223-230.
109. Tagoe DNA, Kalejaiye TD, de Koning HP: The ever unfolding story of cAMP signaling in trypanosomatids: vive la difference! Frontiers in pharmacology 2015, 6:185.

110. Mancini PE, Patton CL: Cyclic 3′,5′-adenosine monophosphate levels during the developmental cycle of *Trypanosoma brucei brucei* in the rat. Molecular and Biochemical Parasitology 1981, 3(1):19-31.
111. Laxman S, Riechers A, Sadilek M, Schwede F, Beavo JA: Hydrolysis products of cAMP analogs cause transformation of *Trypanosoma brucei* from slender to stumpy-like forms. Proc Natl Acad Sci U S A 2006, 103(50):19194-19199.
112. Shalaby T, Liniger M, Seebeck T: The regulatory subunit of a cGMP-regulated protein kinase A of *Trypanosoma brucei*. European Journal of Biochemistry 2001, 268(23):6197-6206.
113. Saada EA, Kabututu ZP, Lopez M, Shimogawa MM, Langousis G, Oberholzer M, Riestra A, Jonsson ZO, Wohlschlegel JA, Hill KL: Insect Stage-Specific Receptor Adenylate Cyclases Are Localized to Distinct Subdomains of the *Trypanosoma brucei* Flagellar Membrane. Eukaryotic Cell 2014, 13(8):1064-1076.
114. Paindavoine P, Rolin S, Van Assel S, Geuskens M, Jauniaux JC, Dinsart C, Huet G, Pays E: A gene from the variant surface glycoprotein expression site encodes one of several transmembrane adenylate cyclases located on the flagellum of *Trypanosoma brucei*. Molecular and Cellular Biology 1992, 12(3):1218-1225.
115. Lopez MA, Saada EA, Hill KL: Insect Stage-Specific Adenylate Cyclases Regulate Social Motility in African Trypanosomes. Eukaryotic Cell 2015, 14(1):104-112.
116. Salmon D, Bachmaier S, Krumbholz C, Kador M, Gossmann JA, Uzureau P, Pays E, Boshart M: Cytokinesis of *Trypanosoma brucei* bloodstream forms depends on expression of adenylyl cyclases of the ESAG4 or ESAG4-like subfamily. Mol Microbiol 2012, 84(2):225-242.
117. Mony BM, MacGregor P, Ivens A, Rojas F, Cowton A, Young J, Horn D, Matthews K: Genome wide dissection of the quorum sensing signaling pathway in *Trypanosoma brucei*. Nature 2014, 505(7485):681-685.
118. Czichos J, Nonnengaesser C, Overath P: Trypanosoma brucei: cis-aconitate and temperature reduction as triggers of synchronous transformation of bloodstream to procyclic trypomastigotes in vitro. Exp Parasitol 1986, 62.
119. Dean SD, Marchetti R, Kirk K, Matthews K: A surface transporter family conveys the differentiation signal in African trypanosomes. Nature 2009, 459.
120. Szöor B, Ruberto I, Burchmore R, Matthews KR: A novel phosphatase cascade regulates differentiation in *Trypanosoma brucei* via a glycosomal signaling pathway. Genes & development 2010, 24:1306-1316.
121. Szöőr B, Wilson J, McElhinney H, Tabernero L, Matthews KR: Protein tyrosine phosphatase TbPTP1: a molecular switch controlling life cycle differentiation in trypanosomes. The Journal of Cell Biology 2006, 175(2):293-303.
122. Jones NG, Thomas EB, Brown E, Dickens NJ, Hammarton TC, Mottram JC: Regulators of *Trypanosoma brucei* cell cycle progression and differentiation identified using a kinome-wide RNAi screen. PLoS Pathog 2014, 10:e1003886.
123. Vassella E, Krämer R, Turner CMR, Wankell M, Modes C, Van Den Bogaard M, Boshart M: Deletion of a novel protein kinase with PX and FYVE-related domains increases the rate of differentiation of *Trypanosoma brucei*. Mol Microbiol 2001, 41(1):33-46.

124. Domenicali Pfister D, Burkard G, Morand S, Renggli CK, Roditi I, Vassella E: A Mitogen-Activated Protein Kinase Controls Differentiation of Bloodstream Forms of *Trypanosoma brucei*. Eukaryotic Cell 2006, 5(7):1126-1135.
125. Barquilla A, Saldivia M, Diaz R, Bart J-M, Vidal I, Calvo E, Hall MN, Navarro M: Third target of rapamycin complex negatively regulates development of quiescence in *Trypanosoma brucei*. Proc Natl Acad Sci U S A 2012, 109(36):14399-14404.
126. Szöőr B, Dyer NA, Ruberto I, Acosta-Serrano A, Matthews KR: Independent Pathways Can Transduce the Life-Cycle Differentiation Signal in *Trypanosoma brucei*. PLoS Pathog 2013, 9(10):e1003689.
127. Kabani S, Fenn K, Ross A, Ivens A, Smith TK, Ghazal P, Matthews K: Genome-wide expression profiling of in vivo- derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in *Trypanosoma brucei*. BMC Genomics 2009, 10(1):427.
128. Rico E, Ivens A, Glover L, Horn D, Matthews KR: Genome-wide RNAi selection identifies a regulator of transmission stage-enriched gene families and cell-type differentiation in *Trypanosoma brucei*. PLoS Pathog 2017, 13(3):e1006279.
129. Walrad P, Paterou A, Acosta Serrano A, K M: Differential trypanosome surface coat regulation by a CCCH protein that co-associates with procyclin mRNA cis-elements. PLoS Pathog 2009, 5.
130. Zimmermann H, Subota I, Batram C, Kramer S, Janzen CJ, Jones NG, Engstler M: A quorum sensing-independent path to stumpy development in *Trypanosoma brucei*. PLoS Pathog 2017, 13(4):e1006324.
131. Dejung M, Subota I, Bucerius F, Dindar G, Freiwald A, Engstler M, Boshart M, Butter F, Janzen CJ: Quantitative Proteomics Uncovers Novel Factors Involved in Developmental Differentiation of *Trypanosoma brucei*. PLoS Pathog 2016, 12(2):e1005439.
132. Alibu VP, Storm L, Haile S, Clayton C, Horn D: A doubly inducible system for RNA interference and rapid RNAi plasmid construction in *Trypanosoma brucei*. Molecular and Biochemical Parasitology 2005, 139(1):75-82.
133. Benz C, Mulindwa J, Ouna B, Clayton C: The *Trypanosoma brucei* zinc finger protein ZC3H18 is involved in differentiation. Molecular and Biochemical Parasitology 2011, 177(2):148-151.
134. Hirumi H, Hirumi K: Continuous cultivation of *Trypanosoma brucei* blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. J Parasitol 1989, 75.
135. Estévez AM, Kempf T, Clayton C: The exosome of *Trypanosoma brucei*. The EMBO Journal 2001, 20(14):3831-3839.
136. Gietz RD, Schiestl RH, Willems AR, Woods RA: Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast 1995, 11(4):355-360.
137. Langmead B, Trapnell C, Pop M, Salzberg SL: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009, 10.
138. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S: The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009, 25(16):2078-2079.

139. Archer SK, Luu D, Queiroz R, Brems S, Clayton CE: *Trypanosoma brucei* PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle. PLoS Pathog 2009, 5.
140. Adiconis X, Borges-Rivera D, Satija R, DeLuca DS, Busby MA, Berlin AM, Sivachenko A, Thompson DA, Wysoker A, Fennell T et al: Comprehensive comparative analysis of RNA sequencing methods for degraded or low input samples. Nature methods 2013, 10(7):623-629.
141. Minia I, Merce C, Terrao M, Clayton C: 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 Neglect Trop Dis 2016, 10(9):e0004982.
142. Martin M: Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal; Vol 17, No 1: Next Generation Sequencing Data Analysis 2011.
143. Love MI, Huber W, Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 2014, 15(12):550.
144. Fadda A, Ryten M, Droll D, Rojas F, Färber V, Haanstra JR, Merce C, Bakker BM, Matthews K, Clayton C: Transcriptome-wide analysis of trypanosome mRNA decay reveals complex degradation kinetics and suggests a role for co-transcriptional degradation in determining mRNA levels. Mol Microbiol 2014, 94(2):307-326.
145. Siegel TN, Hekstra DR, Wang X, Dewell S, Cross GAM: Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. Nucleic Acids Research 2010, 38(15):4946-4957.
146. Bailey TL: DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics 2011, 27(12):1653-1659.
147. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD et al: SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology 2012, 19(5):455-477.
148. Mulindwa J, Mercé C, Matovu E, Enyaru J, Clayton C: Transcriptomes of newly-isolated *Trypanosoma brucei rhodesiense* reveal hundreds of mRNAs that are co-regulated with stumpy-form markers. BMC Genomics 2015, 16(1):1118.
149. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B et al: Fiji: an open-source platform for biological-image analysis. Nat Meth 2012, 9(7):676-682.
150. Zomer AWM, Allert S, Chevalier N, Callens M, Opperdoes FR, Michels PAM: Purification and characterisation of the phosphoglycerate kinase isoenzymes of *Trypanosoma brucei* expressed in Escherichia coli. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology 1998, 1386(1):179-188.
151. Gupta SK, Kosti I, Plaut G, Pivko A, Tkacz ID, Cohen-Chalamish S: The hnRNP F/H homologue of *Trypanosoma brucei* is differentially expressed in the two life cycle stages of the parasite and regulates splicing and mRNA stability. Nucleic Acids Res 2013, 41.

152. Szoor B, Ruberto I, Burchmore R, Matthews KR: A novel phosphatase cascade regulates differentiation in *Trypanosoma brucei* via a glycosomal signaling pathway. Genes Dev 2010, 24.
153. Dean S, Marchetti R, Kirk K, Matthews K: A surface transporter family conveys the trypanosome differentiation signal. Nature 2009, 459.
154. Tschopp F, Charrière F, Schneider A: In vivo study in *Trypanosoma brucei* links mitochondrial transfer RNA import to mitochondrial protein import. EMBO reports 2011, 12:825-832.
155. Hotz HR, Hartmann C, Huober K, Hug M, Clayton C: Mechanisms of developmental regulation in *Trypanosoma brucei*: a polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation. Nucleic Acids Research 1997, 25(15):3017-3026.
156. Drozdz M, Clayton C: Structure of a regulatory 3' untranslated region from *Trypanosoma brucei*. RNA 1999, 5(12):1632-1644.
157. Furger A, Schürch N, Kurath U, Roditi I: Elements in the 3' untranslated region of procyclin mRNA regulate expression in insect forms of *Trypanosoma brucei* by modulating RNA stability and translation. Molecular and Cellular Biology 1997, 17(8):4372-4380.
158. Irmer H, Clayton C: Degradation of the unstable EP1 mRNA in *Trypanosoma brucei* involves initial destruction of the 3′-untranslated region. Nucleic Acids Research 2001, 29(22):4707-4715.
159. Overath P, Czichos J, Haas C: The effect of citrate/cis-aconitate on oxidative metabolism during transformation of *Trypanosoma brucei*. Eur J Biochem 1986, 160.
160. Nagoshi YL, Alarcon CM, Donelson JE: The putative promoter for a metacyclic VSG gene in African trypanosomes. Molecular and Biochemical Parasitology 1995, 72(1):33-45.
161. Bringaud F, Biteau N, Donelson JE, Baltz T: Conservation of metacyclic variant surface glycoprotein expression sites among different trypanosome isolates. Molecular and Biochemical Parasitology 2001, 113(1):67-78.
162. Matthews KR, Shiels PG, Graham SV, Cowan C, Barry JD: Duplicative activation mechanisms of two trypanosome telomeric VSG genes with structurally simple 5' flanks. Nucleic Acids Research 1990, 18(24):7219-7227.
163. Savage AF, Kolev NG, Franklin JB, Vigneron A, Aksoy S, Tschudi C: Transcriptome Profiling of *Trypanosoma brucei* Development in the Tsetse Fly Vector Glossina morsitans. PLoS ONE 2016, 11(12):e0168877.
164. De Pablos LM, Kelly S, de Freitas Nascimento J, Sunter J, Carrington M: Characterization of RBP9 and RBP10, two developmentally regulated RNA-binding proteins in *Trypanosoma brucei*. Open Biology 2017, 7(4):160159.
165. Queiroz R, Benz C, Fellenberg K, Hoheisel JD, Clayton C: Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. BMC Genomics 2009, 10(1):495.
166. Mayho M, Fenn K, Craddy P, Crosthwaite S, Matthews K: Post-transcriptional control of nuclear-encoded cytochrome oxidase subunits in *Trypanosoma brucei*: evidence for genome-wide conservation of life-cycle stage-specific regulatory elements. Nucleic Acids Res 2006, 34.
167. Quijada L, Guerra-Giraldez C, Drozdz M, Hartmann C, Irmer H, Ben-Dov C, Cristodero M, Ding M, Clayton C: Expression of the human RNA-binding protein HuR in *Trypanosoma brucei* increases the abundance of mRNAs

containing AU-rich regulatory elements. Nucleic Acids Research 2002, 30(20):4414-4424.
168. Goldstrohm AC, Hook BA, Seay DJ, Wickens M: PUF proteins bind Pop2p to regulate messenger RNAs. Nat Struct Mol Biol 2006, 13(6):533-539.
169. Lykke-Andersen J, Wagner E: Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. Genes & Development 2005, 19(3):351-361.
170. Nagai K, Oubridge C, Jessen TH, Li J, Evans PR: Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. Nature 1990, 348(6301):515-520.
171. Maris C, Dominguez C, Allain FHT: The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. FEBS Journal 2005, 272(9):2118-2131.
172. Park S, Myszka DG, Yu M, Littler SJ, Laird-Offringa IA: HuD RNA Recognition Motifs Play Distinct Roles in the Formation of a Stable Complex with AU-Rich RNA. Molecular and Cellular Biology 2000, 20(13):4765-4772.
173. Urbaniak MD, Martin DMA, Ferguson MAJ: Global Quantitative SILAC Phosphoproteomics Reveals Differential Phosphorylation Is Widespread between the Procyclic and Bloodstream Form Lifecycle Stages of *Trypanosoma brucei*. Journal of Proteome Research 2013, 12(5):2233-2244.