The biological role of the RBP7910 Z-binding protein in the mitochondrial mRNA processing of *Trypanosoma brucei*

Nisha Ramamurthy

Department of Parasitology
McGill University
Montréal, Canada
August 2020

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science

Table of contents

1. List of Figures and Tables

2. List of Abbreviations

ADAR	Adenosine deaminase acting on RNA
ATP	Adenosine triphosphate
BF	Bloodstream form
CCA	Citrate Cis-Aconitate
CoA	Co-enzyme A
COIII	Cytochrome oxidase III
Cyb	Cytochrome b
DNA	Deoxyribonucleic acid
EP	Glu-Pro
GG	Glycerol gradient
GPEET	Gly-Pro-Glu-Glu-Thr
GRBC	Guide RNA binding complex
hADAR	Human ADAR
HAT	Human African Trypanosomiasis
IEX	Ion exchange chromatography
IFA	Immuno-fluorescence assay
K	Kinetoplastid
k_d	Dissociation constant
KREL1	Kinetoplastid RNA editing ligase 1
KREL2	Kinetoplastid RNA editing ligase 2
KREPA1	Kinetoplastid RNA editing protein A 1
KREPA2	Kinetoplastid RNA editing protein A 2
KREPA3	Kinetoplastid RNA editing protein A 3
LC-MS	Liquid Chromatography coupled with Mass Spectrometry
LS	Long slender
MERS1	Mitochondrial edited mRNA stability factor 1
MERS2	Mitochondrial edited mRNA stability factor 2
MERS3	Mitochondrial edited mRNA stability factor 3
Mito-extract	Mitochondrial extract
MMP	Mitochondrial membrane potential
MS	Mass Spectrometry
MST	Microscale Thermophoresis
MURF	Maxi-circle unassigned reading frame
N	Nucleus
PAMC	Polyadenylation mediator complex
PARP	Procyclic acidic repetitive protein
PBS-G	Phosphate Buffered Saline with Glucose
PF	Procyclic form
RBD	RNA binding domain
RBP7910	RNA binding protein 7910
RECC	RNA editing core complex
REMC	RNA editing mediating complex
RESC	RNA editing substrate binding complex
RNA	Ribonucleic acid
RNAi	RNA interference
RRM	RNA recognition motif
SM-427	Single marker 427
SS	Short stumpy
WHTH	Wings helix turn helix
WT-29.13	Wild type 29.13
ZBD	Z-binding DNA
ZBP	Z binding protein
Z-α	Z-alpha
Z-β	Z-Beta

3. Abstract

Trypanosoma brucei, *Trypanosoma cruzi,* and the *Leishmania* species are the major pathogenic trypanosomatid organisms that cause significant diseases in humans and animals, with high incidence and mortality rates. The target organism in this research was *T. brucei,* a eukaryote where the energy regulation varies in the organism according to the life stages. RNA editing is a form of post-transcriptional RNA processing, which occurs in mitochondrial mRNAs, and is essential for parasite energy generation and viability in *T. brucei*. Our lab recently identified RBP7910 as a protein associated with RNA editing. RBP7910 is essential for the viability of the parasite and potentially contains a Z-DNA binding domain with a regulatory role in RNA editing. The focus of this thesis was to test if RBP7910 is indeed a Z-binding protein and its regulatory role in RNA editing. We used a tagged RBP7910 in *T. brucei* to identify its protein partners and explored its role in the differentiation of the bloodstream stage to the procyclic/insect stage. Using glycerol gradient fractionation of the tagged RBP7910 cell line, we validated RBP7910 as an RNA binding protein since its migration was affected by RNase treatment. Treatment of cell extract with salt before fractionation, however, did not affect the migration of RBP7910, supporting the hypothesis that RBP7910 is a Z-binding protein, as they are known to interact with RNA substrates in a sequence-specific manner rather than based on electrostatic charge. The proteins pulled down by RBP7910 showed that RBP7910 associates with the RNA editing substrate-binding complex (RESC) known to play a role in the RNA editing progression. Our preliminary data also suggest that overexpression of RBP7910 in the bloodstream forms can significantly alter the parasite morphology, including features resembling the insect stage of the parasite.

Résumé

Trypanosoma brucei, Trypanosoma cruzi et les espèces de Leishmania sont les principaux organismes trypanosomatides pathogènes qui provoquent des maladies importantes chez les humains et les animaux, avec une incidence et des taux de mortalité élevés. L'organisme cible de cette recherche était *T. brucei*, un eucaryote dont la régulation énergétique varie dans l'organisme en fonction du cycle de vie du parasite. L'édition de l'ARN est une forme de modification post-transcriptionnelle de l'ARN, qui se produit dans les ARNm mitochondriaux, et est essentielle pour la production d'énergie et la viabilité du parasite chez *T. brucei*. Notre laboratoire a récemment identifié RBP7910 comme une protéine associée à l'édition de l'ARN. RBP7910 est essentielle pour la viabilité du parasite et contient potentiellement un domaine de liaison à l'ADN-Z ayant un rôle régulateur dans l'édition de l'ARN. L'objectif de cette thèse était de tester si RBP7910 est effectivement une protéine de liaison Z et son rôle régulateur dans l'édition de l'ARN. Nous avons utilisé une RBP7910 marquée chez *T. brucei* pour identifier ses partenaires protéiques et avons exploré son rôle dans la différenciation du stade sanguine au stade procyclique/insecte. En utilisant le fractionnement du gradient de glycérol de la lignée cellulaire RBP7910 marquée, nous avons validé la RBP7910 comme protéine de liaison de l'ARN puisque sa migration était affectée par un traitement par la RNase. Le traitement de l'extrait cellulaire avec du sel avant le fractionnement n'a cependant pas affecté la migration de RBP7910, ce qui soutient l'hypothèse selon laquelle RBP7910 est une protéine de liaison Z, car on sait qu'elle interagit avec les substrats d'ARN de manière spécifique à la séquence plutôt que sur la base d'une charge électrostatique. Les protéines entraînées par RBP7910 ont montré que RBP7910 s'associe au complexe de liaison aux substrats d'édition de l'ARN connu pour jouer un rôle dans la progression de l'édition de l'ARN. Nos données préliminaires suggèrent également que la surexpression de RBP7910 dans les formes sanguines peut modifier de manière significative la morphologie du parasite, y compris des caractéristiques ressemblant a l’étape d'insecte du parasite.

4. Acknowledgments

I would first like to express my sincerest gratitude to my supervisor, Dr. Reza Salavati who has guided me throughout my masters with his knowledge and patience. His encouraging and motivating behaviour made me confident and made me enjoy the research I did over the two years. His optimism and encouragement made me work with confidence. I am also thankful for the financial support he gave me throughout. I am so thankful to be a part of this journey in the Salavati lab. I am extremely grateful to my advisory committee member, Dr. Igor Cestari who has guided me through my masters with his knowledge and resources.

I am extremely thankful to Dr. Vaibhav Mehta for training me in the lab. He has been a great support to me during the past two years. His lab skills and knowledge motivated me throughout. I would also like to thank Akshaya Srikanth for her help throughout my masters. She has been a great source of support and a good friend.

I appreciate all the help and support from my lab members, Linhua Zhang for her help with protein purifications, Rouyu Ma for training me with cells and transfections, Moti Sobat for guiding me with the bioinformatics, Moujtaba Rostami, Amin Azimi, Daniel Moses, Saba Mohammaei, Arezoo Kamel and Xinyue Xu for being amazing labmates and for taking an interest in my project.

I would also like to thank previous lab members, Dr. Najmeh Nikpour, for introducing me to the world of RBP7910 and Ivan Lam, for helping with Thermophoresis.

I would also like to thank my good friend Rishi Rajesh who helped me in every step of this journey by being a constant support and motivation.

I will always be grateful for all the professors and friends in the department of Parasitology for accepting me and supporting me.

I would like to acknowledge the scholarship and DFW that I have received from the Faculty of Parasitology, which enabled me to conduct my research without any financial hardships.

Above all, I am indebted to my parents, who have always supported and motivated me throughout my academic career, without whom I could not have undertaken this journey. Their sacrifices and constant care made me achieve this.

5. Introduction and Literature review

5.1 Kinetoplastids: Occurrence, prevalence, and importance

Kinetoplastids are unicellular eukaryotic organisms with a peculiar mitochondrial genome known as the kinetoplast [1]. Unlike other eukaryotes, the kinetoplast contains a network of intertangled circular DNA, consisting of minicircles and maxicircles [2] [3]. Further research on the kinetoplast DNA (kDNA) revealed a novel post-transcriptional mRNA maturation process known as RNA editing. Several kinetoplastid species are parasitic, with many causing fatal diseases affecting millions around the globe. The major kinetoplastid pathogens in humans are *T. brucei*, *Trypanosoma cruzi (T.cruzi)*, and *Leishmania spp.,* that cause human African trypanosomiasis (HAT), chagas' disease, and leishmaniases, respectively [4]. These diseases collectively affect millions of people around the globe, and the current treatments available are not ideal due to toxicity, inefficacy, and the emergence of resistant strains [5] [6].

HAT, also known as sleeping sickness, is a vector-borne parasitic disease caused by *T. brucei* and transmitted to humans by tsetse flies (*Glossina* genus). Tsetse flies are mainly found in the sub-Saharan African region. There are approximately sixty-five million people in the sub-Saharan African region at risk of contracting the disease [7]. Two subspecies of *T. brucei* cause sleeping sickness, *T. brucei rhodesiense,* and *T. brucei gambiense*, prevalent in the east and west Africa, respectively. HAT is fatal and progresses in two stages. The hemolymphatic stage includes a lot of irregular and random symptoms such as skin lesions, cardiac problems, and more as the parasite proliferates through the bloodstream and the lymphatic system. The second stage is when the trypomastigotes enter the central nervous system and cause major neurological disorders.

Chagas' disease is caused by *T. cruzi*, which is endemic in Mexico, Central America, and South America, with approximately 8 million people currently infected. It is transmitted by the triatomine bug, which is found mainly in poor housing conditions. While acute infections present symptoms that are mild with skin lesions and swelling, but chronic infections entails cardiomyopathy with heart failure and ventricular arrhythmias that lead to sudden death. Measures have been taken for stopping or eradicating the disease, but the attempts have not been successful yet [8].

The several forms of **leishmaniases** are caused by approximately 21 species of *Leishmania* that are pathogenic to humans and transmitted by a family of sandflies. These parasites proliferate intracellularly in the phagolysosomes of macrophages, causing diseases that range from self-healing cutaneous infections to disfiguring mucocutaneous infections and visceral leishmaniasis [9]. Disease manifestations are greatly influenced by host-immune responses [10] and present public health challenges, including severe cases of HIV co-infection [11].

5.2 Life cycle of *T. brucei*

When a tsetse fly takes a blood meal from a mammalian host, it injects metacyclic trypomastigotes into the epidermal layer of the skin that transforms into the bloodstream form (BF) trypomastigotes. These trypomastigotes proliferate in the bloodstream and lymphatic system of the mammalian host, via binary fission. The cycle progresses when a tsetse fly takes a blood meal from an infected host, transforming into the procyclic form (PF) trypomastigotes inside the fly's midgut. They gradually migrate towards the salivary gland as epimastigotes transforming into the metacyclic forms ready for mammalian host infection. This process occurs in 3 weeks. The whole life cycle of *T. brucei* occurs in extracellular stages [12] shown in (Fig. 5.1).

These life stages have specific energy regulation and morphological forms. The parasite in the mammalian host evades the antibody responses through antigenic variation by variant surface glycoproteins (VSGs). Trypanosomes propagate in the mammalian host as long slender (LS) BF, transitioning into the non-propagative short stumpy (SS) BF that are ready to transform into the PF trypomastigotes [13].

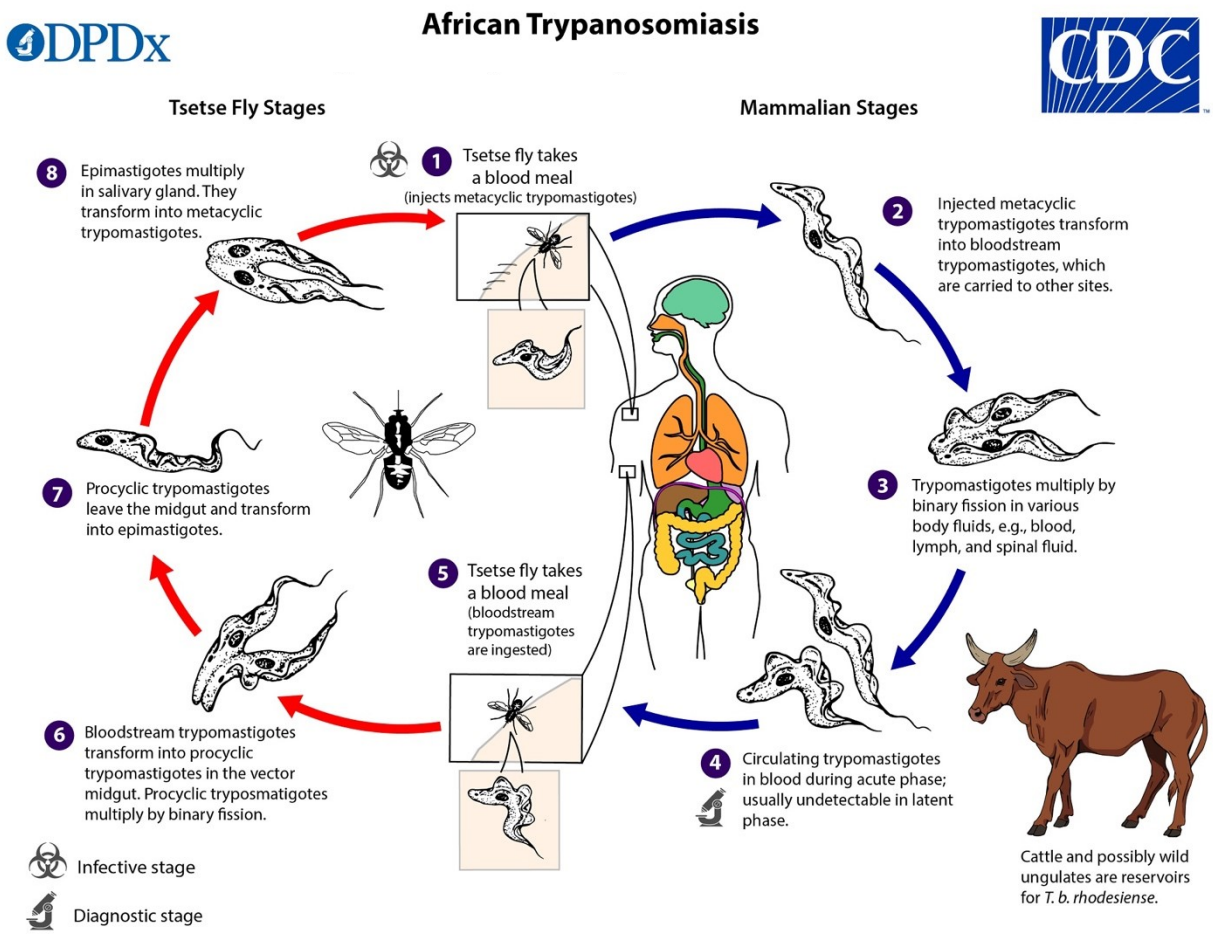


***Fig 5.1*:** *A detailed image of the life cycle of T. brucei [14].*

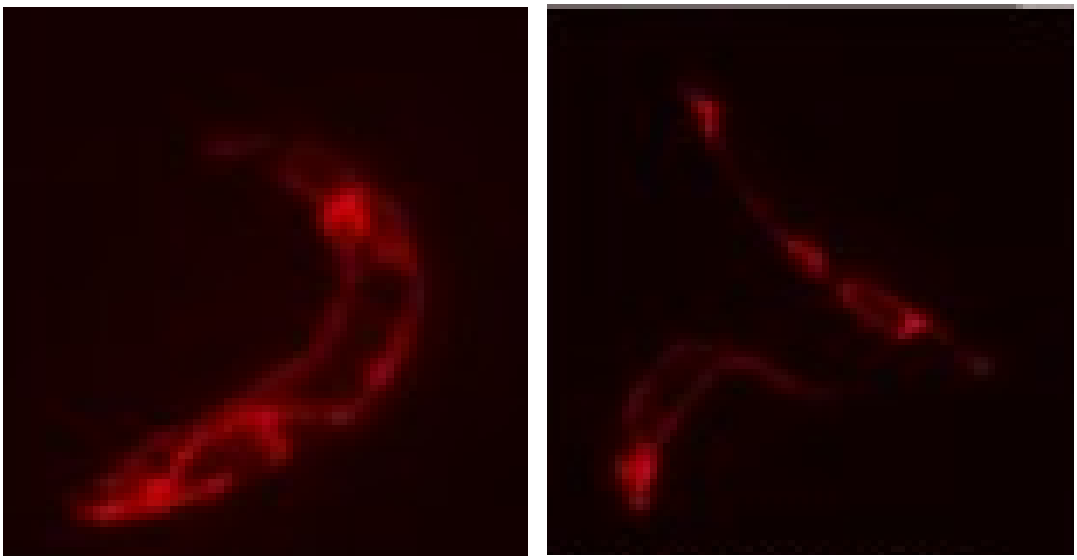
5.3 Energy metabolism in *T. brucei*:

The trypanosomes contain a large elongated mitochondrion that runs throughout the cells (Fig. 5.2). The mitochondrion in BF *T. brucei* lacks cristae and has reduced metabolic activity than in PF, as seen from the early microscopic analysis [15], as shown in figure 5.2A. Energy metabolism in BF is solely via glycolysis of the abundant blood glucose available [16]. On the other hand, the PF does not have the advantage of glucose as an abundant energy source and has a highly active mitochondrion for oxidative phosphorylation shown in Fig 5.2B. PF contains an acetate succinate CoA transferase and a succinyl CoA synthetase in addition to the Krebs cycle and electron transport chain. Acetate succinate CoA transferase serves in the mitochondria of some metazoan

and anaerobic protists. They generate adenosine triphosphate (ATP) by converting acetyl CoA to acetate [17].

A.

PF stage **BF stage**



B:

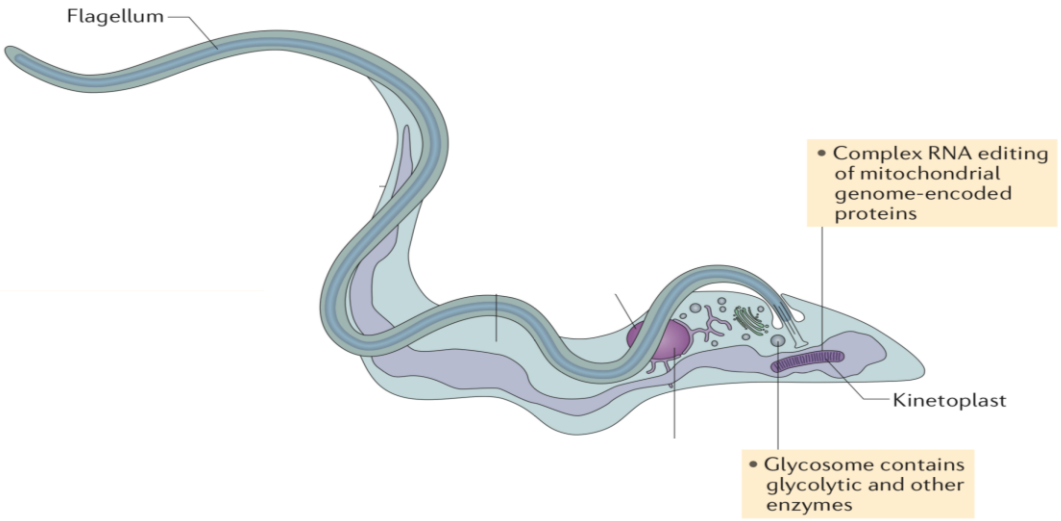


Fig 5.2: Mitochondrial structure in PF and BF. (A) *Mitochondria of a PF cell (left) that contain more cristae than in the BF cell (right). The cells are stained with mitotracker that is absorbed in the mitochondrion [5]. (**B**) A schematic representation of T. brucei containing the kinetoplast inside its single large mitochondrion, taken from [18] with permission from Spring Nature.*

5.4 Differentiation between PF and BF

During SS BF differentiation into PF trypomastigotes, the cells stop production of the VSG proteins, and a new coat of procyclin proteins replaces them [19] [20]. VSG expression is life stage-specific that is activated in the tsetse-fly salivary-gland and inactivated upon return to the tsetse-fly mid-gut. The coat provides a powerful protection as *T. brucei* cells occupy the bloodstream and tissue-spaces of their hosts and are fully exposed to immune vigilance. As an infection persists, a majority of the parasite population is periodically eliminated but does not get eradicated because of the VSG proteins [19] (Fig. 5.3).

Procyclin proteins are attached through glycosyl phosphatidylinositol (GPI), found in two distinct forms: EP procyclins that have GLU-PRO repeats [20], [21], and GPEET procyclins that have Gly-Pro-Glu-Thr repeats [22]. While the EP procyclins have 21-27 repeats (EP1,EP2,EP3…) the GPEET procyclins have 5 repeats [23], [24] (Fig. 5.3). The freshly transformed PF trypomastigotes express both GPEET and EP procyclins until day 7 (also determined *in vitro* [25]), from that point on, only EP procyclins are expressed [26]. EP procyclins are also found in PF epimastigotes in the salivary glands but are repressed in the metacyclic form. The LS BF transforms into the SS BF in optimal conditions. While the LS BF is exceptionally dependent on glucose, the SS BF adapt to poorly oxygenated environments as present in the fly midgut. There are changes in enzyme and mitochondrial functions, suggesting preparation for survival in an environment that is low in glucose and oxygen [25]. The cultured pleomorphic strains transition from the LS BF to SS BF, while the monomorphic strains do not. The monomorphic cells usually get arrested or locked in the G1/G0 phase of the life cycle and do not survive the newly provided conditions for differentiation. Studies have also shown that the SS BF is programmed to die without further division [27].

Morphologically, the cells undergoing differentiation are arrested in the G0 phase of division if the conditions are not optimal for cell cycle differentiation. The initiation is in the kinetoplast and

nucleus. Dividing cells can be found with two kinetoplasts and two nuclei. Cells in the pleomorphic strain gradually complete cellular division, however in the case of monomorphic strains, the nucleus and kinetoplasts divide and are arrested in the G0/GI phase and do not survive for longer periods [28]. This transition can occur in certain *in vitro* conditions. For example, citrate cis-aconitate (CCA) and temperature regulation are two main key factors for BF to PF differentiation [19]. In *T. brucei,* the optimal response to CCA requires the generation of SS BF [29], which are non-dividing and show partial adaptation for conditions in the midgut of tsetse flies, including elevated mitochondrial activity [30].

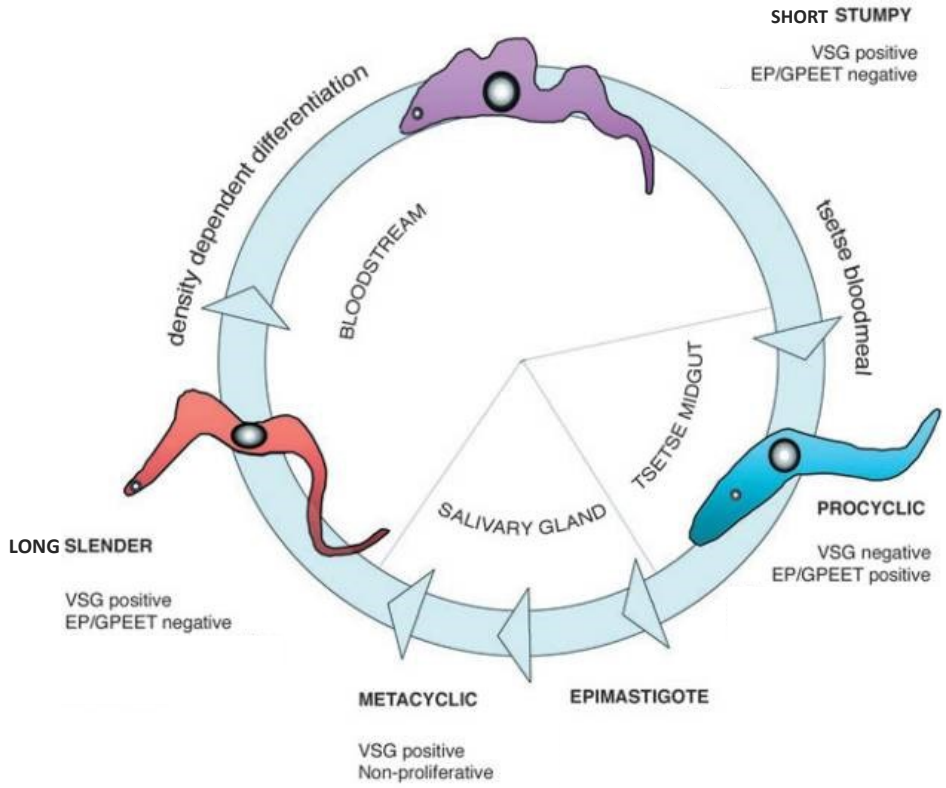


Fig 5.3: Transformation from LS BF to SS BF. *The LS BF transitions to SS BF (known as the early procyclic stage). The procyclin and VSG expressions have been indicated in specific life stages taken from [31] with permission from Journal of Science.*

5.5 Trypanosomatid RNA editing and its catalytic mechanism

RNA editing is a unique post-transcriptional maturation process that takes place in the mitochondrion of trypanosomatids. This process involves uridine (U) insertion and deletion in the premature transcripts originating from the kinetoplast to finally produce mature transcripts that code for multiple components of the electron transport chain. The specificity and amount of U inserted/deleted are dictated by guide RNA (gRNA) molecules.

The maxicircles encode all 18 protein-coding genes along with a few gRNA while the minicircles code for almost all the gRNA molecules. Six maxicircle genes exist as open reading frames, while 12 of them require U insertion/deletion RNA editing to produce translatable RNAs [32].

RNA editing requires simultaneous activities of various RNA binding protein complexes, including the RNA editing core complex (RECC) that catalyzes U-insertion and deletion [30] [33]. Editing initiates with the base-pairing of the 5' region of a gRNA with the segment downstream of the first editing site in the premature mRNA (Fig. 5.4) [34]. The remaining gRNA, including its U-tail, associates with the mRNA through complementary regions, mainly consisting of purine-rich segments that further stabilize the mRNA/gRNA duplex [35, 36]. The mismatches in the duplex on the editing site are recognized by an RNA editing endonuclease (KREN1/2), which cleaves the premature mRNA at the first unpaired nucleotide while the gRNA bridges the two fragments. The RNA editing TUTase 2 (KRET2) and RNA editing exoribonuclease 1/2 (KREX1/2) insert and deletes U residues at the 3' end of the 5' cleavage fragment, respectively. Finally, the edited sites are ligated by an RNA editing ligase 1/2 (KREL1/2) [6]. Once the sequence is edited in the mRNA, it serves as the precursor for priming the next gRNA [37]. RNA editing helicase (KREH) unwinds the mRNA gRNA duplexes during this process [38].

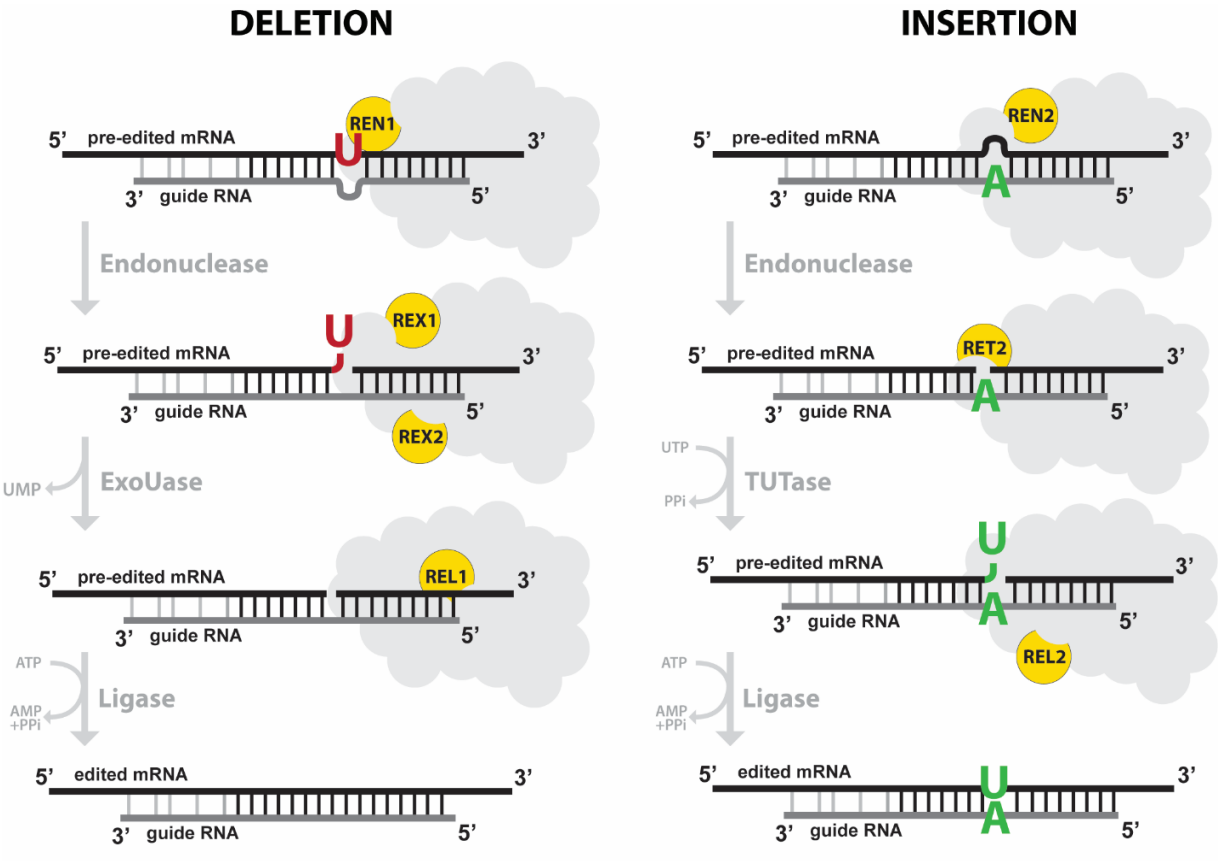


Fig 5.4: Catalytic steps in RNA editing. *The enzymatic steps involved in RNA editing. The left panel shows U deletion while the right panel shows U insertion. Editing initiates with an endonucleolytic cleavage at the editing site by REN1/2, proceeds with either U-deletion or insertion by an exoUase (REX1/2) or TUTase (RET2) and terminates with ligation of the edited site by REL1/2. Watson-Crick base pairing is indicated with black lines, while G:U wobble pairing is indicated in grey , taken from [39]*

5.6 RNA editing is regulated in a life-stage specific manner

RNA editing is differentially regulated in the BF and PF life stages, as represented in Figure 5.5. While the PF cells require all respiratory complexes for energy metabolism, oxidative phosphorylation is absent in the BF stage. Typically, complexes I, III, and IV pump protons into the mitochondrial intermembrane space to generate the mitochondrial membrane potential (MMP). While complexes III and IV are active in PF parasites [40], complex V adopts the unusual role of maintaining the MMP in BF parasites through the hydrolysis of ATP [41]. Regulation of different respiratory complexes are developmentally regulated at the level of RNA editing. Mitochondrial membrane

potential maintenance is important for several other essential cellular functions, including and not limited to lipid biogenesis, mitochondrial protein import, and calcium homeostasis [41]. Numerous knock-down and knock-out studies on RNA editing proteins indicate that this process is essential in both life stages [42] [43]. The mRNAs in the PFs are preferentially edited shown in red (Fig. 5.5), preferentially edited in bloodstreams in yellow and some are edited in both stages (green).

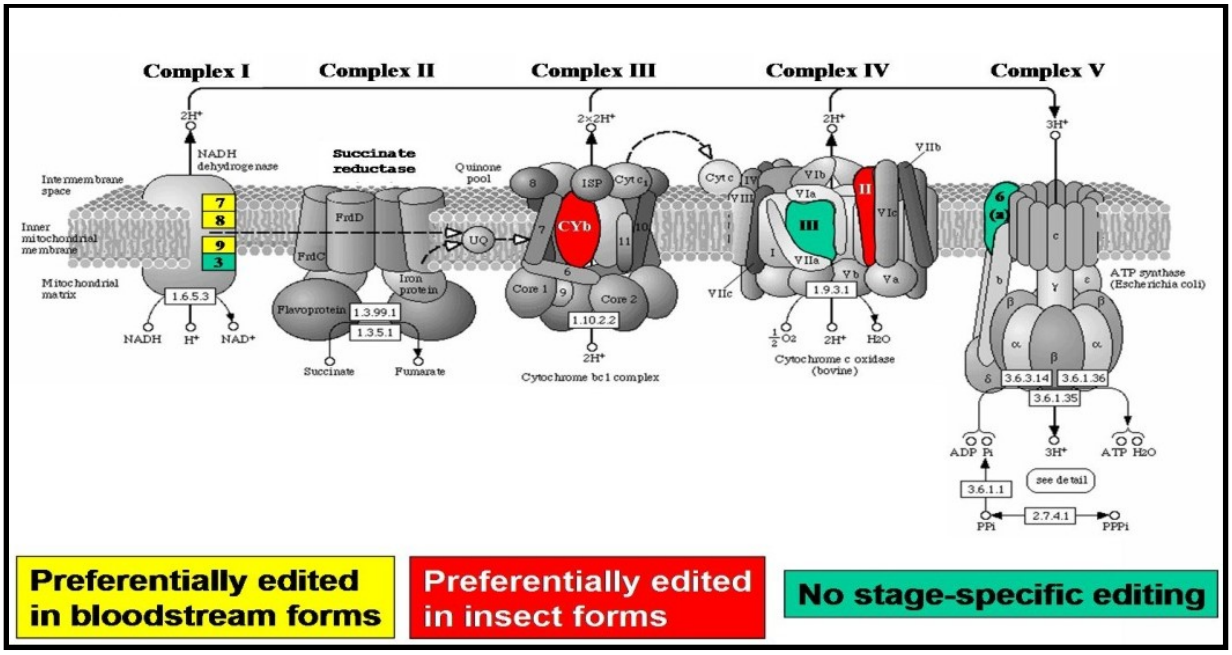


Fig 5.5: RNA editing regulation is life stage specific. *The image all five respiratory complexes that function in the oxidative phosphorylation system. These complexes play a role in generating a mitochondrial membrane potential and cellular energy through Krebs cycle products. Multiple components of these complexes are edited in specific life stages of the parasite, BF (yellow), PF (red) and both (green). Image adapted and modified from [4].*

5.7 Accessory complexes involved in RNA Editing

While the multi-protein RECC plays the catalytic role in U insertion/deletion RNA editing, several other protein complexes were identified to play a regulatory role that aids the RECC in specificity and fidelity. *In vivo* RNA editing is a complex mechanism, and the discovery and understanding of guide RNA binding complex (GRBC), also known as mitochondrial RNA binding complex 1 (MRB-1), was essential in understanding *in vivo* editing [44-46]. GRBC1 and GRBC2 proteins

that belong to the GRBC complex are functionally proven to bridge GRBC with the RNA editing process. GRBC1 and GRBC2 are responsible for gRNA stabilization [45, 47]. RNA editing initiation occurs in the GRBC complex, which was proven by knockdown studies on the components present in the GRBC/MRB1 complex [48-50]. GRBC is connected to the polyadenylation complex through polyadenylation mediator complex (PAMC) (Fig. 5.6), which plays a role in mRNA polyadenylation [33].

The GRBC also interacts with the TbRGG2 subcomplex, which is presently known as the RNA editing mediator complex (REMC) (Fig. 5.6). TbRGG2 interacts with RECC in an RNA dependent manner [33]. TbRGG2 also has an RNA binding domain (RBD) or RNA recognition motif (RRM), which is 70 amino acids in its structure. RBD/RRM containing proteins typically enhance protein-protein and protein-RNA/DNA interactions and is also a platform for the B-DNA/RNA to convert itself into a Z-form of RNA [51]. Knockdown analyses of the TbRGG2 protein suggest that it plays a role in the 3' to 5' progression of RNA editing, gRNA utilization, and RNA trafficking [48, 52, 53]. GRBC, REMC, and PAMC together form the RNA editing substrate-binding complex (RESC) (Fig. 5.6). The complete editosome holoenzyme consists of RECC and RESC [33]. Purified RECC run on glycerol gradients as ~20S (Svedberg) and ~40S complexes, with the latter possibly containing the RESC and endogenously bound [33]. Mitochondrial RNA binding proteins (MRP1 and MRP2) is a complex around 100kDa which associates with the core editosome with RNA at low concentrations of salt [45, 54-56]. They enhance the melting of gRNA stem loop I and annealing with the associated mRNA [56-59]. RNA interference analysis of the proteins reflected less stability of certain edited mRNAs [60, 61]. The RGG1 and RBP38 proteins are both hypothesized to play a role in RNA stability [44, 62, 63].

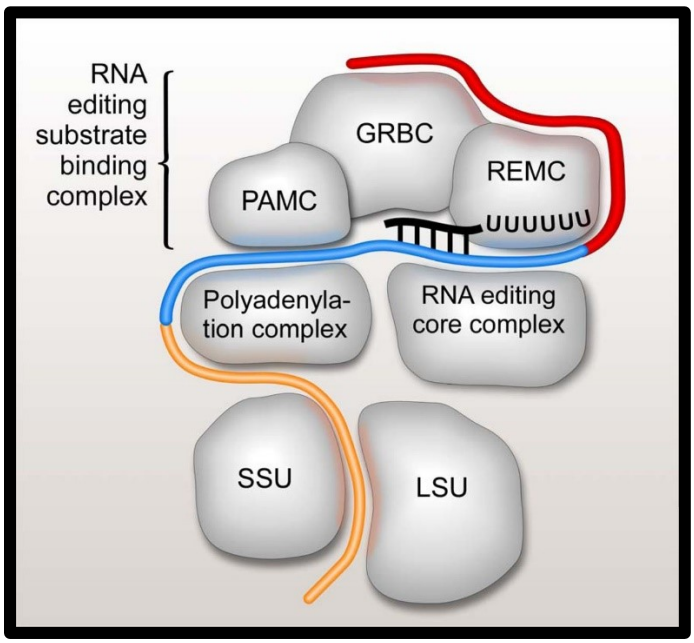


Fig 5.6: *A schematic representation of the multi-protein complexes involved in RNA editing. RESC, RECC, PAMC complexes. The RESC consists of GRBC, REMC, and PAMC. These complexes are involved in gRNA stabilization, recruitment of RNA editing core complexes, and polyadenylation complexes, respectively. The red line shows the pre-edited mRNA, edited mRNA in blue and the edited and polyadenylated/uridylated mRNA in orange. Taken from [33] with permission from American Society for Microbiology.*

5.8 RBP7910: Discovery and Importance

A *T. brucei* protein-protein interaction network analyzed the proteins associating with the RECC through co-fractionation on glycerol gradients (GG) and ion-exchange chromatography (IEX) [64]. The study identified 50 proteins that play a role in RNA editing, which include 17 proteins from the RECC and 21 proteins from the RESC. Some proteins directly associated with the RECC proteins while some proteins bound to the RECC via RNA (i.e., RNA dependent interaction). These interactions made the proteins dissociate into different clusters, although they co-sedimented together. Six novel proteins were identified in this study: Tb927.1.3010, Tb927.10.7910, Tb927.1.1730, Tb927.10.5830, Tb927.6.1200, and Tb927.10.1730, suggesting they played a role in RNA editing. Intriguingly, IEX fractionation patterns indicate Tb927.6.1200 and Tb927.10.1730 co-fractionated with members of the RECC and other accessory proteins. The

sedimentation profiles of the other four proteins in IEX chromatography suggested they were sensitive to salt, with Tb927.10.7910 (later termed RBP7910 [65]) being the most sensitive one. This observation potentially implicates RBP7910 to bind with its target RNA through charged interactions. Tetracycline inducible RNA interference (RNAi) knockdowns of the six novel proteins indicate that five are essential for the survival of PF *T. brucei*. As RBP7910 appeared to be one of the most salt-sensitive proteins which was also essential for the parasite's survival, the authors chose to perform additional experiments with that protein [64]. To verify the effect of RBP7910 knockdown on the editing of the mitochondrial transcripts, quantitative PCR analysis was also performed. RBP7910 knockdown surprisingly resulted in a reduced abundance of edited CyB (Cytochrome B) transcript and increased abundance of edited ATPase subunit 6 (A6) and edited Cytochrome Oxidase III (COIII) transcripts [64]. These data implicate a regulatory role for RBP7910 in RNA editing. RBP7910 is a potential component of the RESC

A recent study from 2018 indicates RBP7910's presence in the PPsome (termed as MERS3 in [66]). The transcription-defined 5′ terminus is converted into a monophosphorylated state by the pyrophosphohydrolase complex, termed the PPsome. Composed of the MERS1 NUDIX enzyme, the MERS2 pentatricopeptide repeat RNA-binding subunit, and MERS3 polypeptide, the PPsome binds to specific sequences near mRNA 5′ termini. Most guide RNAs lack PPsome-recognition sites and remain triphosphorylated. RBP7910 (MERS3) in this study was described to lack motifs similar to the other two proteins (MERS1 and MERS2), and the proteins interact directly with each other [66]. However, recent publications suggest it plays a role in the RESC with recommended renaming to RESC19 in [67]. RBP7910 has also been pulled down with TbRGG2 in a previously published work [68] via RNA [64].

5.9 RBP7910 is predicted to contain Z-DNA/RNA binding domains

Homology based predictions of RBP7910 using the HH-PRED online server identified two Z-DNA binding domains (https://toolkit.tuebingen.mpg.de/tools/hhpred). These domains are found in Z-DNA binding proteins (ZBP) such as the Human adenosine deaminase acting on RNA (hADAR) [65]. The prediction placed the Z-DNA binding domains, namely Z-α and Z-β, toward the N and C termini of RBP7910, respectively [69]. The Z-DNA binding domain family belongs to the superclass of the winged helix turn helix (WHTH) domains, which is largely represented in

the DNA-binding domains of prokaryotic and eukaryotic transcription factors. This domain specifically recognizes the Z-form of DNA/RNA molecules in a conformation-specific manner. The secondary structure prediction showed three-helix bundles and three β-sheets with an αβααββ topology for both domains (Fig. 5.8B) [65].

Z-folded nucleic acids are a left-handed conformation, usually observed on purine-pyrimidine repeats such as poly d(GC) in high concentrations of salt [70]. The left-handed helix is built on a dinucleotide repeat where the anti-conformation of bases alternate with the unusual syn form where the purine-pyrimidine base goes over the deoxyribose ring. This conformation gives rise to a zig-zag structure known as Z-DNA (Fig. 5.7). The proteins which contain a Z-DNA/RNA binding domain are classified as ZBPs. Crystallographic studies on ZBPs reveal the junction where the Z-form arises in a B-folded DNA, known as the B-Z junction, upon interaction with a ZBP (Fig. 5.7) [70]. This condition is favored around non-Watson-Crick base pairs such as mismatches, and often found occurring with processive enzymes acting on DNA such as polymerases and helicases. Z-DNA is stabilized both by negative super helical stress and by binding of the Zα domain of Z-binding proteins. ZBPs interact with their Z-folded nucleic acid target in a conformation-specific manner [70].

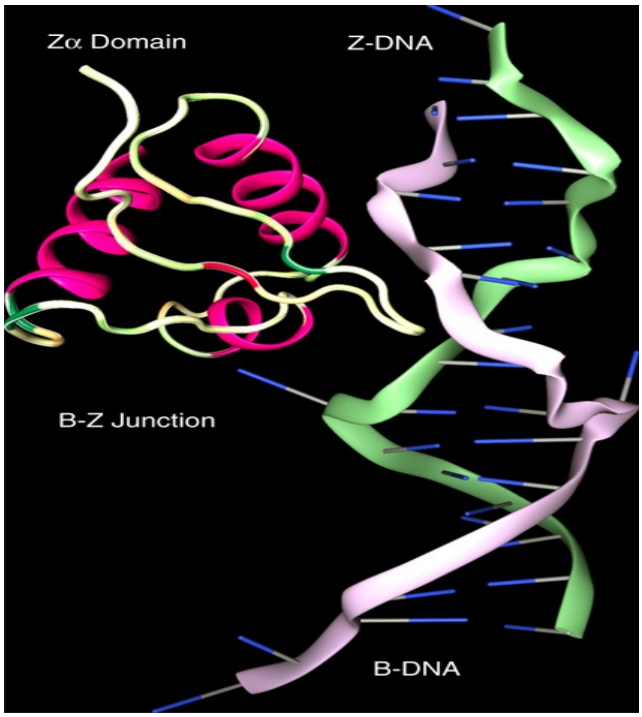


Fig 5.7 Shift from B-form- Z-form: *The shift from B-DNA to a Z-DNA and the B-Z junction between the two forms. The green helix represents one strand of DNA, and the pink strand represents the second strand of DNA. The point of contact of the DNA helix and the Z- α forms the Z-form taken from [70] with permission from Spring Nature.*

A well-known example of a ZBP is the hADAR, which was one of the top hits in the HH-PRED analysis on RBP7910 [65]. hADAR plays a role in A-I RNA editing, wherein the target adenosine is converted into an inosine (read as guanosime during translation) by a deaminase domain on the C-terminus of the protein. The protein also contains multiple double-stranded RNA binding domains (dsRBD) that aid in target recognition (Fig 5.8A). hADAR functions in the nucleus, editing transcripts that code for proteins important in neurotransmission. Mutations or deletions of hADAR genes lead to seizures, un-coordination, and neurodegeneration [71]. Recent studies have shown hADAR interaction with a poly d(GC) [70]. In contrast to hADAR, RBP7910 only contains the two Z-DNA binding domains, Z-α and Z-β (Fig. 5.8A and B).

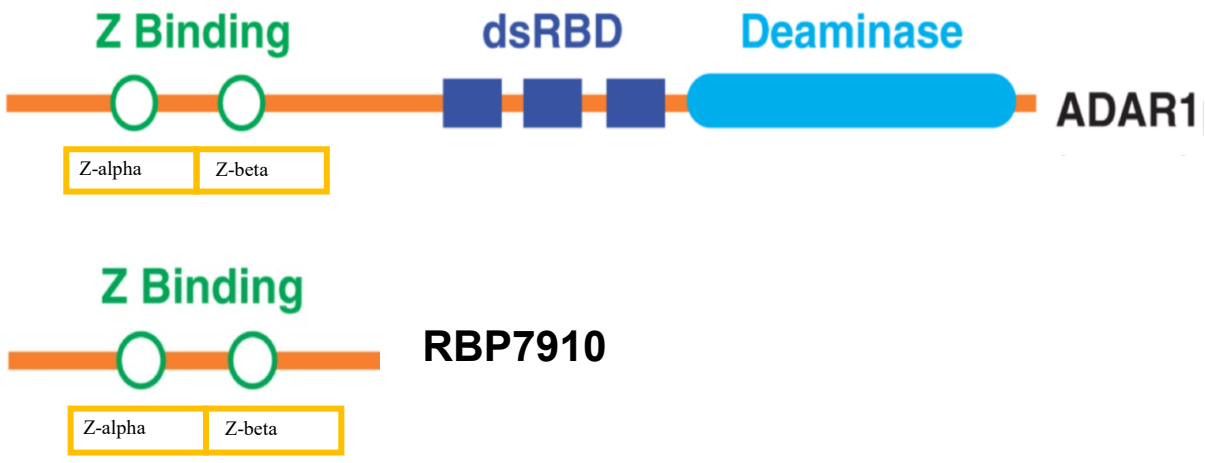
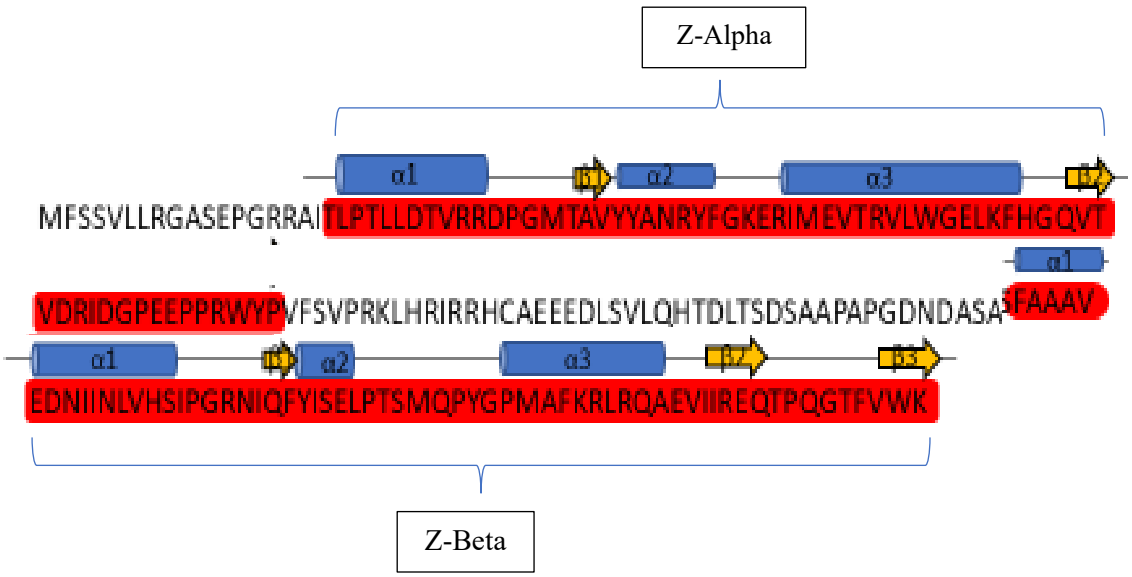


Fig 5.8A- Human ADAR and RBP7910 Z-binding domains: *The structure of Human ADAR with 2 Z-binding domains (Z- α and Z- β), double stranded RNA binding domains and a deaminase region [72] and RBP7910 with a Z- α and Z- β domains.*



RBP7910

Fig 5.8B: *RBP7910 a homolog shows a Z-alpha region and a Z-beat region in the structure, taken from [65] with permission from Spring Nature.*

5.10 RBP7910 interacts with AU enriched sequences

Adenosine-Uridine (AU) repeats similar to d(GC) are also purine-pyrimidine repeats, which are most often found in mitochondrial mRNA in *T. brucei*, such as the AU elements of pre-edited Cyb mRNA that have been shown to regulate RNA editing [73]. Moreover, studies also show that AU rich sequences are important for the formation of pre-edited mRNA/gRNA duplex as A to C point mutations and the formation of gRNA duplex was reduced by 80% [74]. Another important structure of AU repeat is found in the post-editing AU extension of the primary short A tail of pre-edited transcripts. On the other hand, the long AU tail, is found in the translation process of fully edited transcripts. RNA binding proteins affect the stability of mitochondrial mRNAs, which contain AU tails [75]. Since HH-PRED analysis predicted RBP7910 to be an Z-RNA binding protein, electro-mobility shift assays (EMSA) were performed to confirm its RNA binding capacity with various RNA substrates, such as Cyb and A6 mRNA segments, and native gRNA with a U tail [65]. Moreover, RBP7910 exhibited a high affinity toward synthetic RNA with AUUU repeats (Fig. 5.9). Competition assays with RNA containing AUCUA repeats do not perturb that interaction, substantiating RBP7910's specificity for AUUU stretches (Fig. 5.9).

However, poly U RNA was the most effective competitor to perturb the interaction between RBP7910 and AUUU (Fig. 5.9). Poly A and poly G substrates were not found to be substantial competitors. Hence, the study concluded that RNA with AU and U stretches to be the targets of RBP7910 [65].

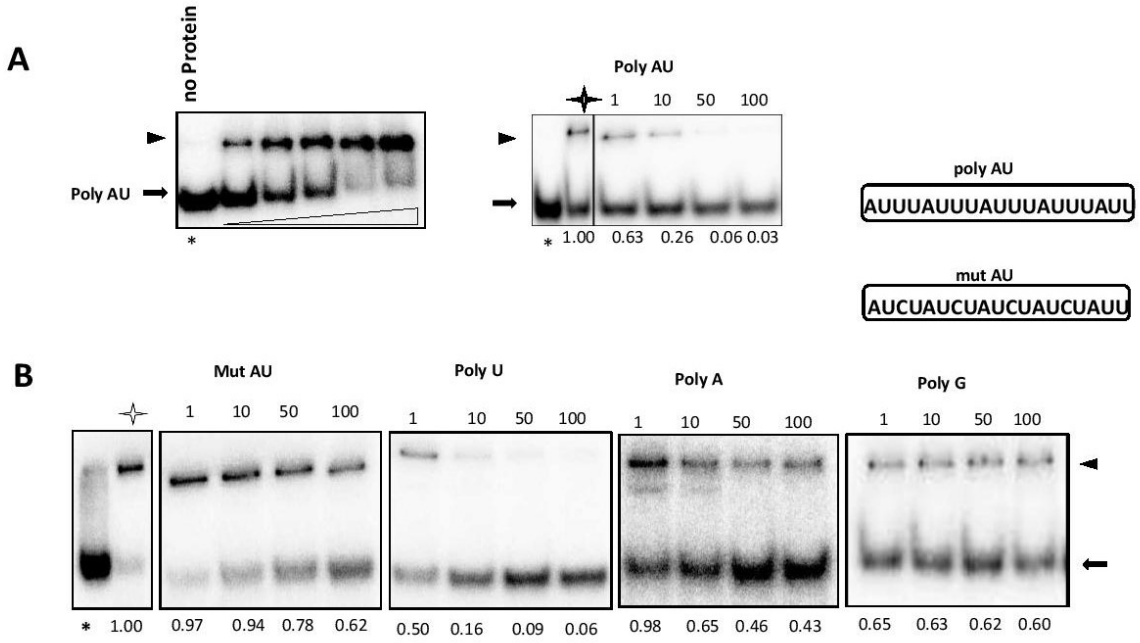


Fig 5.9: Competition assays to determine the binding affinity of RBP7910 with Poly AU sequence. ***Panel A*** *shows the binding of RBP7910 with non-radiolabelled (cold) and radiolabelled (hot) AU sequence while* ***Panel B*** *shows the competition assays and the % of shifts due to binding. Clearly, poly U and Poly AU compete out the rest of the substrates and conclude to be a good binding partner [65]*.

Since RBP7910 binds to AU and U stretches, it could be playing a functional role in regulating RNA editing through interactions with AU and U segments in the mRNA AU/U-tails, gRNA U-tails, or the U-inserted sequences in the edited sites of the mature mRNA (Fig. 5.10). The objective of this study was to determine the accurate mode of action through which RBP7910 confers its regulatory role in RNA editing.

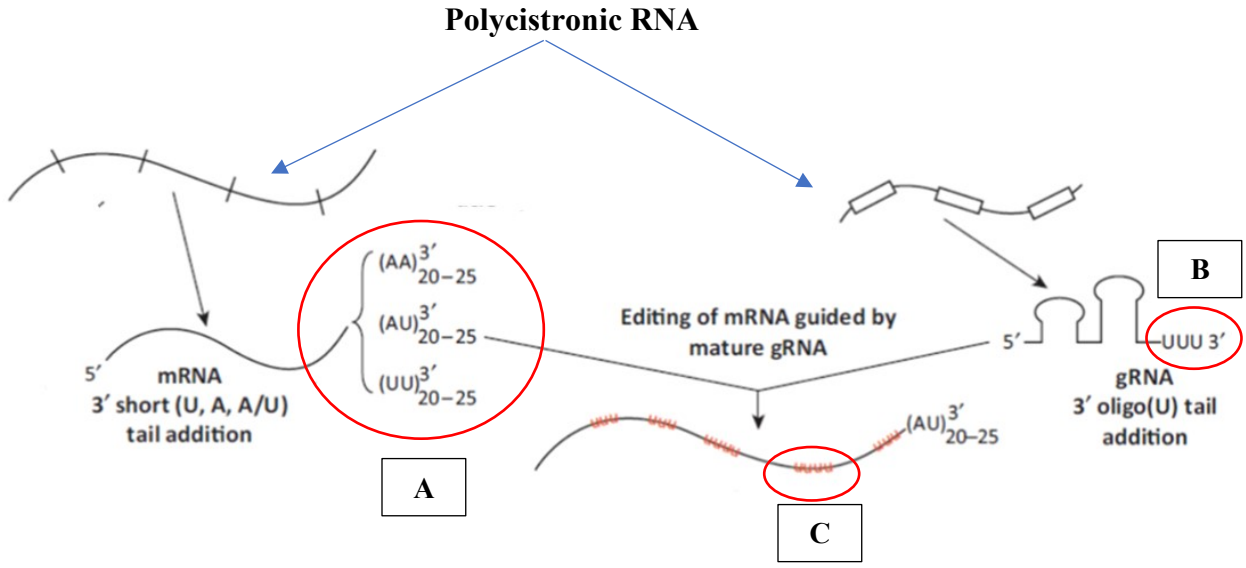


Fig 5.10: Possible RNA targets of RBP7910 in trypanosomatid RNA editing. (A) AU/U-tails of mRNA, (B) gRNA U-tails and (C) U stretches resulting from U-insertion editing. *Image used from [76]*.

6. Hypothesis

Given that RBP7910 is a mitochondrial RNA binding protein with a predicted Z-DNA binding motif and found associated with RNA editing and life-stage specific transcripts (i.e., CYB), RBP7910 may play a role in controlling differential editing. It does so potentially by interacting with other mitochondrial proteins involved in life-stage specific mRNA processing.

7. Objectives

The main aim of this research was to determine the mechanism of RBP7910 gene regulation. Accordingly, to achieve this goal, we pursued the following three specific objectives. The first was to investigate the RNA binding mode of RBP7910 further, by introducing an ectopic copy of tagged RBP7910 in the parasite, and monitoring its fractionation upon various treatments (i.e., RNase, high salt). The second was to find the interacting protein partners of RBP7910 by immunoprecipitation and mass spectroscopy. The third was to examine the essentiality of the RBP7910 for the BF by creating transgenic RNA interference (RNAi) knockdown cell lines and the effect of RBP7910 overexpression in BF and differentiation to the insect form of the parasite.

8. Materials and methods

8.1 Cloning RBP7910 into the pLew100-3v5 plasmid

RBP7910 was cloned into the pLew100-3v5 plasmid, under the control of a tetracycline regulatable T7 promoter. It also contains a phleomycin resistance gene for selecting positive transfectants. RBP7910 gene was amplified from *T. brucei* genomic DNA using the high fidelity Q5 polymerase following manufacturers' recommendations (NEB) with primers containing a Hind III and BamHI restrictions sites. The purified plasmid and RBP7910 DNA were digested with Hind III and BamHI restriction endonucleases following recommended protocol by NEB and purified via agarose gel electrophoresis and gel extraction kit (NEB). The two DNA fragments were then ligated with T4 DNA ligase (NEB) overnight at 16°C and transformed into DH5α competent bacteria. Plasmids were then purified from a few colonies for sequencing (Genome Quebec) and the colony containing the plasmid without any errors was inoculated for the large-scale MidiPrep plasmid purification (Qiagen). The sequences for RBP7910, C-terminal v5 tag, and the complete RBP7910-3v5 are given below, respectively.

RBP7910 sequence

ATGTTTTCCAGCGTATTACTGCGCGGGGCATCGGAACCCGGCCGGAGGGCAATAAC
GCTGCCCACACTGCTCGATACGGTGCGCCGCGATCCAGGAATGACGGCCGTTTATTA
TGCCAATCGTTACTTTGGAAAAGAACGAATAATGGAGGTAACACGCGTCCTGTGGG
GTGAACTGAAGTTCCACGGCCAGGTCACGGTGGATCGTATAGACGGGCCAGAAGAG
CCCCCACGGTGGTACCCCGTGTTTTCTGTGCCGCGCAAACTCCACCGAATACGTCGC
CACTGTGCGGAGGAGGAAGATCTCAGCGTGCTGCAGCACACTGATTTAACATCGGA
TAGCGCGGCACCGGCACCGGGAGATAATGATGCAAGTGCATCCTTCGCTGCCGCTG
TCGAGGACAACATCATCAATTTAGTCCACTCTATCCCGGGGCGCAACATCCAGTTTT
ATATTAGCGAACTTCCCACCTCCATGCAGCCATATGGCCCGATGGCGTTTAAACGGT
TGCGGCAAGCTGAGGTAATCATTAGGGAACAGACACCGCAGGGAACTTTTGTATGG
AAGTAA

3v5 sequence

GGTAAGCCTATCCCTAACCCTCTCCTCGGTCTCGATTCTACGCGAGGTAAGCCAATC
CCAAACCCACTTCTTGGTCTCGACTCAACACAGCTGGGTAAGCCCATCCCCAACCCC
CTCCTCGGTCTTGATAGCACGTAGTAA

3v5 tag protein sequence (v5 underlined)

GKPIPNALLGLDSTRGKPIPNPLLGLDSTQLGKPIPNPLLGLDST

RBP7910 in pLew100-3v5 plasmid

AGACTTCAATTACACCAAAAAGTAAAATTCACAAGCTTATGTTTTCCAGCGTATTAC
TGCGCGGGGCATCGGAACCCGGCCGGAGGGCAATAACGCTGCCCACACTGCTCGAT
ACGGTGCGCCGCGATCCAGGAATGACGGCCGTTTATTATGCCAATCGTTACTTTGGA
AAAGAACGAATAATGGAGGTAACACGCGTCCTGTGGGGTGAACTGAAGTTCCACGG
CCAGGTCACGGTGGATCGTATAGACGGGCCAGAAGAGCCCCCACGGTGGTACCCCG
TGTTTTCTGTGCCGCGCAAACTCCACCGAATACGTCGCCACTGTGCGGAGGAGGAA
GATCTCAGCGTGCTGCAGCACAGTGATTTAACATCGGATAGCGCGGCACCGGCACC
GGGAGATAATGATGCAAGTGCATCCTTCGCTGCCGCTGTCGAGGACAACATCATCA
ATTTAGTCCACTCTATCCCTGGGCGCAACATCCAGTTTTATATTAGCGAACTTCCTAC
CTCCATGCAGCCATATGGCCCGATGGCGTTTAAACGGTTGCGGCAAGCTGAGGTAAT
CATTAGGGAACAAACACCGCAGGGAACTTTTGTATGGAAGGGATCCGGTAAGCCTA
TCCCTAACGCTCTCCTCGGTCTCGATTCTACGCGAGGTAAGCCAATCCCAAACCCAC
TTCTTGGTCTCGACTCAACACAGCTGGGTAAGCCCATCCCCAACCCCCTCCTCGGTC
TTGATAGCACGTAGTAA

Sequences highlighted in red – Restriction sites (Hind III and Bam H1)

Sequences highlighted in gray- RBP7910 sequence

Sequences highlighted in green- 3v5 tag

8.2 DNA precipitation and transfection

For transfection, 10µg of the cloned pLew100-RBP7910-3v5 vector was linearized with 60 U of the Not I restriction endonuclease in a 400µL reaction volume overnight at 37°C. This linearization results in a DNA molecule containing flanking regions from the rDNA spacer from *T. brucei* genome. Following digestion, 40µl of 3M sodium acetate and 880 µl of ice cold 100% ethanol were added and incubated at -80°C for 30 mins. The precipitated DNA was then pelleted by centrifugation at 13,200xg for 30 mins at 4°C. The supernatant was discarded and 300 µL of 70% ice-cold ethanol was added to the pellet to remove residual salt. The tube was centrifuged again in the same conditions as above, and the 70% ethanol wash was repeated once more. Finally, after the last spin, the supernatant was carefully aspirated in the biosafety cabinet (BSC), and the pellet was allowed to air dry and then resuspended in 20µL of sterilized milli-Q water.

For transfection, a total of 2.5 x 10^8 PF cells (29.13 strain) were grown to late-log phase 1-2 x 10^7 cells/ml. The cells were harvested in JA 25.50 rotor (Beckman Coutler) at 2000rpm in at 4°C for 7 mins. The supernatant was discarded and the cell pellet was washed with 3ml Cytomix solution (120 mM KC1, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4 pH 7.6, 25 mM Hepes pH 7.6, 2 mM EGTA pH 7.6, 5 mM $MgCl_2$, final pH adjusted with KOH to 7.6 if required and sterilized through 0.2 µ syringe filter in the BSC). The cells were washed again in the Cytomix solution as done above and finally resuspended in 200 µl of the same and transferred to two transfection cuvettes, each containing 100 µl (Fisher Scientific). Next, 20 µl of the resuspended linearized vector and sterile milli-Q water was added to the cuvettes and immediately electroporated with Amaxa biotechnology Nucleofector II following manufacturer's recommended conditions. The electroporated cells were transferred to 25 cm^2 culture flasks containing 10 ml SDM-79 media supplemented with 10% fetal bovine serum (FBS), 15 µg/ml G418 and 50 µg/ml hygromycin and incubated for growth at 28°C with agitation in Mandel Max 4000 incubator at 70 rpm. The rDNA flanking regions in the linearized vector guide homologous recombination into the parasite genome (in the rDNA spacer); detailed explanation is provided in results section 10.1.

8.3 Selecting RBP7910-3v5 PF positive transfectants

The cloned pLew100-RBP7910-3v5 plasmid contains a phleomycin resistance gene for use in selecting positive transfectants. After 24h of incubation at 28°C, phleomycin was added to the culture flasks to a final concentration of 2.5 μg/ml. From this flask, the culture was serially diluted in 24-well plates using SDM-79 media containing phleomycin as well, in order to obtain clonal transfectants. While the cells in the control flask did not survive beyond 4 days of phleomycin addition, measurable growth appeared in the serially diluted culture containing transfected cells around day 8 after phleomycin addition. The positive transfectants obtained from the 24-well plate were then transferred to 25 cm^2 culture flasks and monitored over the next week to measure their growth rate, which should be approximately 16h compared to the wild-type PF cell line.

8.4 Mitochondrial extract preparation and glycerol gradient sedimentation

The final positive transfectant cells, RBP7910-3v5 PF, were grown to the late log phase 2 x 10^7 cells/ml in a 1 L volume containing tetracycline (0.5 μg/ml) for inducing protein expression. The cells were harvested in JA 16.25 rotor (Beckman Coutler®) at 6000xg, 10min, 4°C. The cell pellets were washed with ice-cold PBS supplement with 6 mM glucose, then re-suspended in 25 ml hypotonic DTE buffer [1 mM Tris-HCl pH 8.0 and 1 mM EDTA] and incubated on ice for 5 mins. The cells were then gently lysed with 10 strokes using a 40 ml sterile tight-fitting Dounce homogenizer on ice. The lysates were immediately made isotonic by adding a sucrose to a final concentration of 250 mM and centrifuged at 15,800 x g for 10 min at 4°C. The pellet containing crude mitochondria was re-suspended in 4.6 ml STM buffer [20 mM Tris-HCl pH 8.0, 250 mM sucrose and 2mM $MgCl_2$] with 0.3 mM $CaCl_2$ and 40units of RNase-free DNase I (Roche) for 1 hr on ice. The DNase activity was quenched by the addition of 4.6 ml STE buffer [20 mM Tris-HCl pH 8.0, 250 mM sucrose and 2mM EDTA] and centrifuged again as above. The crude mitochondrial pellet was then resuspended into 1 ml lysis buffer [10 mM Tris-HCl pH 7.2, 10 mM $MgCl_2$, 100 mM KCl, 1 μg/ml pepstatin, 1 mM DTT, and 1x complete EDTA-free protease inhibitor (Roche)] and lysed with the addition of triton X-100 to a final concentration of 1% (v/v).

The lysates were rotated on a tube rotator for 20 mins at 4°C, and centrifuged at 16,000xg, 20 mins at 4°C twice, to eliminate debris. The resulting 1 ml cleared lysate was split into 4 tubes and subjected to the following treatments and incubated at RT for 1h: RNase A (0.1 mg/ml, Sigma), DNase 1 (400 U, Roche) and 300 mM KCl.

In the meantime, 10-30% glycerol gradients [in 25 mM Hepes pH 7.9, 50 mM KCl, 10 mM $Mg(OAc)_2$, and 1 mM EDTA] were prepared using the Gradient Maker 108 (Biocomp) following manufacturer's protocol. The lysates were then loaded onto the gradients and centrifuged at 36,000 rpm for 6h at 4°C in an SW41 rotor using the Beckman Coulter ultracentrifuge. The gradients were manually fractionated in 500 ml aliquots from top to bottom, resulting in a total of 24 fractions from each gradient. For the salt gradients, the cleared mitochondrial lysates were treated with 50mM, 500mM and 1M KCl and incubate at RT for 1h as done above. The fractions obtained were then analysed on western blot (procedure mentioned below) using antibodies against RECC proteins (KREPA1, KREPA2, KREL1, and KREPA3) and anti-v5 antibody for recognizing RBP7910-3v5. This procedure was adapted and modified from [77-79].

All the samples collected were stored in 2X SDS. The samples were boiled to 95°C for 6mins, and they were loaded on an SDS -Page gel for electrophoresis. The proteins were separated for a good 40mins after which they were transferred on a PVDF blot at 4°C for 1 hour. The blot was then blocked with 5% milk and PBS-Tween. The blot was then probed with anti-V5 (1:20,000) (Biorad) as the primary and anti-mouse HRPO (1:10,000) (Biorad) as the secondary antibody. The blot was then developed in a Chemidoc™ MP Imaging system (Biorad) for further analysis.

8.5 Purification of RBP7910-3v5 via immuno-precipitation

The pLew100-RBP7910-3v5 PF transfectant cells were grown upto late log phase in 3 x 1 L SDM-79 media, 2 L with and 1 L without (control) tetracycline. The cells were harvested in JA 10.5/16.25 rotors (Beckman Coulter) at 6000xg, 10mins, 4°C. They were then washed in ice-cold PBS supplemented with 6mM glucose and harvested again as above. The cell pellets were each re-suspended in 8ml cold IPP50 (10mM Tris-HCl pH 8.0, 50mM NaCl, and 0.1% NP40),

containing 1% BSA and 1 complete mini EDTA-free protease inhibitor cocktail tablet (Roche). The cells were lysed with the addition of 1ml 10% Triton X-100 to a final concentration 1% and incubated on ice for 20 mins with occasional mixing. The cell debris was cleared off the lysates by centrifugation in JLA 25.5 rotor (Beckman Coulter) at 13,200 rpm, 15 min and 4°C). In the meantime, 10 ml of IPP50 and 200µl anti-V5 Sepharose beads (Sigma Aldrich) were added to three 10ml disposable columns (Biorad), washed for 5 min at 4°C and drained by gravity flow. The cleared lysate was then added to the column and rotated overnight at 4°C. To one of the columns containing the cleared lysate from tetracycline induced cells, RNase A (Sigma Aldrich) was added to a final concentration 0.1mg/ml, and incubated further for 3h at 4°C. The columns were drained and washed with cold IPP50 twice. The resulting beads were re-suspended in 140µl elution buffer (1 mg/ml v5 peptide (Sigma), 1 x PBS) and incubated at 37°C for 10mins. The supernatants were collected and snap-frozen in liquid nitrogen before storage at -80°C. Aliquots were collected through the procedure for western blot analysis. The outline of the method was adapted and modified from [77]

8.6 RECC ligase adenylating assay

The RNA editing ligases (REL1 and REL2) were detected in the immuno-precipitations via an adenylation assay using ATP, [α-^{32}P] (Perkin Elmer). The adenylation reactions were carried out on the three immuno-precipitated eluates from the previous section (i) no induction (-tetracycline) (negative control) (ii) induced (+tetracycline) (iii) induced (+tetracycline) with RNase A treatment. The reactions composed of 25 mM Tris-HCl [pH 8.0], 10 mM $Mg(OAc)_2$, 0.5 mM DTT, 1% BSA, 10% DMSO and 20µCi [α-^{32}P] ATP and were incubated for 15 min at RT (final volume is 30 µl) (protocol adapted from [80]). Adenylation was stopped by the addition of SDS-PAGE loading dye, and the proteins were resolved on a denaturing SDS-PAGE gel. The radiolabeled RELs were detected using Phosphor Imaging.

8.7 Liquid chromatography coupled to tandem mass spectrometry

Proteomics analyses were performed by the Center for Advanced Proteomics Analyses, a Node of the Canadian Genomic Innovation Network that is supported by the Canadian Government through Genome Canada. The following protocol was performed by Dr. Eric Bonneil in University of Montreal (eric.bonneil@umontreal.ca).

The immuno-precipitated eluate obtained in section 8.7 were reconstituted in 50 mM ammonium bicarbonate with 10 mM TCEP [Tris (2-carboxyethyl) phosphine hydrochloride; (Thermo Fisher Scientific) and vortexed for 1 h at 37°C. Chloroacetamide (Sigma-Aldrich) was added for alkylation to a final concentration of 55 mM. Samples were vortexed for another hour at 37°C. 1µg of trypsin was added, and digestion was performed for 8 h at 37°C. Samples were dried down and solubilized in 5% ACN-0.2% formic acid (FA). The samples were loaded on a 1.5 ul pre-column (Optimize Technologies, Oregon City, OR). Peptides were separated on a home-made reversed-phase column (150-µm i.d. by 200 mm) with a 56-min gradient from 10 to 30% ACN-0.2% FA and a 600-nl/min flow rate on a Easy nLC-1000 connected to a Q-Exactive HF (Thermo Fisher Scientific, San Jose, CA). Each full MS spectrum acquired at a resolution of 60,000 was followed by tandem-MS (MS-MS) spectra acquisition on the 15 most abundant multiply charged precursor ions. Tandem-MS experiments were performed using higher energy collision dissociation (HCD) at a collision energy of 27%. The data were processed using PEAKS X (Bioinformatics Solutions, Waterloo, ON) and a *T. brucei* database. Mass tolerances on precursor and fragment ions were 10 ppm and 0.01 Da, respectively. Fixed modification was carbamidomethyl (C). Variable selected posttranslational modifications were oxidation (M), deamidation (NQ), phosphorylation (STY). The data were visualized with Scaffold 4.3.0 (protein threshold, 99%, with at least 2 peptides identified and a false-discovery rate [FDR] of 1% for peptides).

8.8 Purifications of recombinant (r) RBP7910 and human (h) ADAR Zα domain

8.8.1 Expression and purification of rRBP7910

RBP7910 was cloned into the bacterial expression vector pET30a, between Nde I and Xho I restriction sites that results in a C-terminal 6x histidine tag upon expression. The first 8 amino acids predicted as the mitochondrial import signal were deleted for better expression, as the protein natively functions without the import signal in *T. brucei* [65]. The purified pET30a plasmid and amplified RBP7910 DNA were digested with Nde I and Xho I restriction enzymes and ligated with T4 DNA ligase before transformation into DH5α competent bacteria according to NEB's recommendations.

RBP7910 protein sequence with import signal (striked through)

~~MFSSVLLR~~GASEPGRRAITLPTLLDTVRRDPGMTAVYYANRYFGKERIMEVTRVLWGE
LKFHGQVTVDRIDGPEEPPRWYPVFSVPRKLHRIRRHCAEEEDLSVLQHTDLTSDSAAPA
PGDNDASASFAAAVEDNIINLVHSIPGRNIQFYISELPTSMQPYGPMAFKRLRQAEVIIRE
QTPQGTFVWK

RBP7910 cloned into the pET30a plasmid was transformed into T7 Express lysY/Iq Competent E.coli (NEB) and a starter culture was grown at 37°C, 200 RPM overnight in 50 ml Luria-Bertani (LB) broth containing 25 µg/ml kanamycin, until the OD 600 measurement was 3-5. Inoculum from this culture was then added to 1 L LB media with kanamycin and grown at 37°C until an OD of 0.6. IPTG was then added to the culture to a final concentration of 0.5 mM to induce protein expression for 3h at 37°C. All subsequent steps were done at 4 °C. The proteins were purified essentially to homogeneity under nondenaturing conditions as follows. The cells were then harvested at 7000xg for 15mins at 4°C and re-suspended in 10ml lysis buffer (Hepes 20mM pH 7.5, NaCl 1M, β-mercaptoethanol 10 mM , 0.1mM EDTA, 1% Triton, 10 mM Imidazole, 1mg/ml Lysozyme, 100 µM PMSF, 1 x protease inhibitor cocktail from Roche©) and incubated on ice for 30 minutes. The lysate was then sonicated for a total of 1min, 10secs pulse on and 10secs pulse off using a sonicator (Fisher Scientific). The cell debris was cleared from the lysate by

centrifugation (15000xg, 30mins, at 4°C) with the soluble protein retained in the supernatant. In the meantime, 600µl of Ni-NTA agarose slurry was placed in 10-Eco-column (Biorad), resulting in a bed volume (Ni-NTA agarose) of 300µl. The beads were then washed with 3ml water thrice and equilibrated with 6ml equilibration buffer (20 mM Hepes pH 7.5, 1M NaCl, 0.1 mM EDTA, 1% Triton and 10% Glycerol). The equilibrated beads were then transferred into a tube containing the cleared supernatant and was left to bind for an hour at 4°C. The protein bound beads were then collected by passing the slurry through 10ml Eco - column through gravity flow, and washed thrice with the wash buffer (20mM Hepes pH 7.5, 1M NaCl, β-mercaptoethanol, 0.1mM EDTA, 1% Triton, 10% Glycerol, 20mM Imidazole). Bound RBP7910 was eluted with 0.8ml Elution buffer (20mM Hepes pH 7.5, 1M NaCl, 10 mM β-mercaptoethanol, 0.1mM EDTA, 1% Triton 1%, 10% Glycerol , 500mM Imidazole). The elutions were pooled together and dialyzed with a 10kDa cutoff dialysis cassette (Thermofisher) in PBS supplemented with 20% glycerol at 4°C. The proteins were then collected and snap frozen in liquid nitrogen before storage at -80°C. Fractions were analyzed by denaturing SDS-polyacrylamide gel electrophoresis (PAGE) on 15% gels.

8.8.2 Expression and purification of hADAR Zα domain

The hADAR Zα sequence was cloned into the pZa77 for bacterial expression by Genscript. The plasmids received were transformed into T7 Express lysY/Iq Competent E. coli (NEB) as above.

hADAR Zα sequence:
LSIYQDQEQRILKFLEELGEGKATTAHDLSGKLGTPKKEINRVLYSLAKKGKLQKEAGTP
PLWKIAVSTQAWNQHSG

The transformed cells were grown and induced identically as above. The cell pellet obtained after harvesting was resuspended in 15 ml buffer A (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM imidazole, 5 mM β-mercaptoethanol, and 100µM phenylmethylsulfonyl fluoride) and the cells were lysed by sonication similarly. The lysate was then centrifuged for 30 min at 15,000 × *g*, and the cleared supernatant was bound to 0.8 ml of Ni^{2+}-nitrilotriacetic acid metal affinity resin (Qiagen) for 1 h. The resin was washed three times with 8 ml of buffer A in the column and then washed with 40 ml of buffer B (50mM Tris-HCl pH 8.0, 1M NaCl, 10mM imidazole, and 5mM

β-mercaptoethanol). His tagged fusion protein was eluted in 0.5ml buffer C (50mM Tris-HCl pH 8.0, 500mM NaCl, and 5mM β-mercaptoethanol) four times, each with 30 minutes incubation. Eluted fractions containing the protein were pooled and dialyzed against buffer D (PBS with 20% glycerol). Fractions were analyzed by denaturing SDS-polyacrylamide gel electrophoresis (PAGE) on 15% gels.

8.9 Micro-scale thermophoresis (MST)

Four nucleic acid substrates were designed with a 3' Cy5 fluorophore prepared by Integrated DNA Technologies (IDT) (Table 1), for analysis with Monolith NT.115 Red (NanoTemper®). The GC ssDNA Cy5 substrate was designed with a non-fluorescent antisense GC complementary DNA. The two strands were annealed together by mixing equimolar amounts and heating the mixture to 95°C in a water bath and letting the water bath cool down to RT. This annealing protocol was followed for the ssRNA sequences as well, without any complementary molecule. The GC DNA substrate should fold into a Z conformation when bound to hADAR Zα (control, [70]) while the AU RNA substrate were designed for binding with rRBP7910 as used previously [65]. A third version of the AU substrate was designed without the random flanking regions, containing only AUUU repeats. The mutated version AUCUA here is used as a control.

Table 1: RNA and DNA substrates labelled with Cy5 for binding experiments:

GC-Cy5	5'-GGTGGGTTATCGCGCGCGCGCGATCAACTGGGC/3Cy5/-3'
GC antisense	5'-GGTGGGTTATCGCGCGCGCGCGATCAACTGGGC-3'
AUUUA-Cy5 (long)	5'-AUUCUUUUCGCGUCUCUUUUAUUUAUUUAUUUAUUUAUUUUUUUCUCUCUUAUU/3Cy5/ -3'
AUUUA-Cy5 (truncated)	5'- UUAUUUAUUUAUUUAUUUAUU/3Cy5/- 3'
AUCUA-Cy5	5'- AUUCUUUUCGAGUCUCUUUUAUCUAUCUAUCUAUCUAUUUUUUUCUCUCUUAUU/3Cy5/-3'

Following annealing, a pre-test and binding check were performed on the RNA and DNA substrates to optimize the fluorophore readout. The pretest was performed using various concentrations of both substrates ranging from 5nM – 5 μM to test for aggregation and adsorption

to the capillaries onto which the samples were loaded. These experiments indicate usage of 50nM substrate for optimal fluorescence output in downstream experiments.

Protein binding experiments were performed with rRBP7910 and hADAR Zα, wherein 50nM of the respective nucleic acid substrate were incubated with 1:2 serially diluted proteins, concentrations ranging from 5 μM to 0.53nM. The reactions were incubated at RT for 30 mins and absorbed onto the capillaries and introduced into the machine for MST analysis. The fluorescence change between two time points (laser on until laser off) is measured for each sample and an affinity curve is generated.

8.10 RNA interference in the bloodstream stage

The p2T7 vector containing the RBP7910 sequence cloned previously was used here [64]. The plasmid was linearized by the Not I restriction enzyme for integration into the rDNA spacer genome via homologous recombination. BF *T. brucei* 427 single marker (transfected with the tetracycline repressor and T7 RNA polymerase genes on a single selection marker, G418) cells were used here for transfection. The transfection and selection were performed as done in sections 8.2 and 8.3, except that each transfection required only 2 x 10^7 cells and that the cells were cultured in HMI-9 media at 37°C. RNAi induction from the resulting transgenic cell line is regulated by a tetracycline operator.

8.11 Localization of RBP7910 using Immunofluorescence assay (IFA)

A BF variant of the over-expressing pLew100-RBP7910-3v5 was created by the transfection of the respective plasmid into BF SM-427 as done above. The resulting transfectants were used for cellular localization of RBP7910-3v5 by IFA. The BF transfectants were grown to late-log phase ($2x10^6$ cells/ml) in HMI-9 media supplemented with 10% FBS, 1 μg/ml hygromycin, 1 μg/ml phleomycin. The cells were cultured with 1 μg/ml tetracycline (induced sample) and without tetracycline (uninduced samples) at 37°C for 24 h. A final concentration of 100 nM MitoTracker® (Invitrogen) was added to cells to probe the mitochondria. After 1 h of treatment, Tetracycline-induced and uninduced parasites were spun down, washed in PBS supplemented with 6mM

glucose and finally resuspended in TDB buffer pH 7.4 [20 mM NA2HPO4, 2 mM Na2HPO4, 5mM KCl, 80 mM NaCl, 1 mM MgSO4, and 10 mM Glucose] at a density 7 x 10^6 cells/ml. In the meantime, round coverslips (Fisher Scientific) were placed into a 24 well plate and 125 µl 0.1% Poly-L-lysine solution (Sigma-Aldrich) was dropped on the top face of the coverslip and incubated at RT for 1 h and was washed twice with 1 ml milli-Q water for 5 min on a shaker. Finally, 1 ml of the resuspended cells were then added onto each slide and incubated at RT on a shaker. After 1 h incubation, the unbound cells were removed and adhered cells were fixed onto the slides with 1 ml 4% paraformaldehyde (pH 7.4) (Fisher Scientific) in PBS for 10 min at RT. The slides were then washed twice with PBS for 5 min, followed by permeabilization with 1 ml 0.2% Triton X-100 in PBS for 15 min at RT. After two washes with PBS, the coverslips were blocked with 3% BSA in PBS for 1 h and then washed with PBS for 5 min. Anti-V5 antibody (1:10,000 in PBS) was then incubated with the coverslips for 2h at RT. Following extensive washes with PBS, the secondary goat anti-mouse conjugated with Alexa Flour®488 (Abcam) (1:10,000) was prepared in PBS, added to the coverslips and incubated for 1 h at RT in the dark. Coverslips were then washed four times with PBS before staining nucleic acids with DAPI (4',6-diamidino-2-phenylindole) (100 ng/ml), which strongly binds to A-T rich regions in DNA. After 10-15 min, the slides were washed three times with PBS. A mounting solution, Fluoromount-G (Southern Biotech) was dropped on a microscope slide (Fisher Scientific), and each coverslip was carefully flipped with parasites down on a microscopy slide. After 24 h, slides were imaged in the dark using a Nikon Eclipse E800 upright wide field microscope.

9. Results

10. Objective 1. To investigate the RNA binding mode of RBP7910 by monitoring its sedimentation profile upon various treatments.

10.1 Ectopic copy of RBP7910-3v5 introduced in the rDNA spacer of the PF *T. brucei*

The gene (RBP7910-3v5) was introduced in the rDNA locus (a non-transcribed spacer). rDNA spacer is also called the silent region as it is not affected by the read-through transcription of the genome. 18S is an rRNA which belongs to the ribosomal RNA in the spacer. The red dot represents the Tetracycline operator to which the tetracycline repressor binds. Tetracycline regulatable PARP (procyclic acidic repetitive protein) promoter is responsible for the expression of RBP7910-3v5, while a T7 promoter expresses a Phleomycin resistant gene [81].

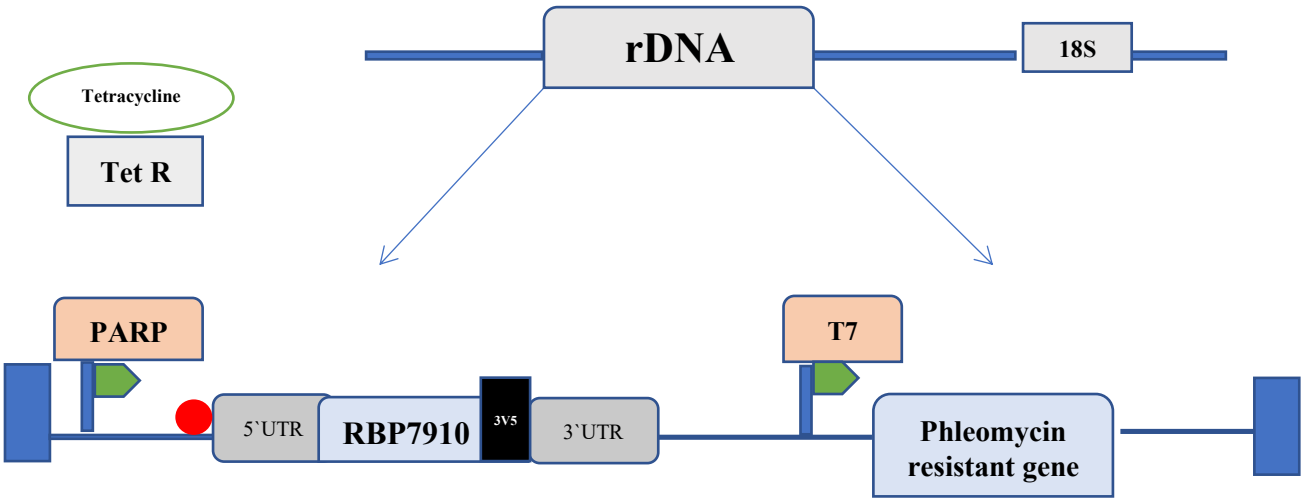


Fig 10.1: The ectopic copy introduced in the rDNA spacer of T.brucei. *Over-expression of RBP7910 was achieved by this construct and RBP7910 was regulated in the rDNA/silent spacer. The tetracycline binds to the Tetracycline repressor and allows the PARP promoter to transcribe the RBP7910 gene and a T7 promoter is present downstream with a phleomycin resistant gene which selects the positive transfectants. The UTRs are from the vector which would help with the homologous recombination rDNA spacer. This construct was adapted and modified from [81].*

Procyclic form cells (Wild type 29.13) was transfected with RBP7910 with a v5 tag using pLew100-3v5 plasmid. The gene recombines in the rDNA spacer of the genome and expresses as an ectopic copy, as controlled by a tetracycline operator. In this case, RBP7910 gets overexpressed in the cells, as the cells may already be expressing the endogenous version. Procyclic form cells non-treated with tetracycline were used as a negative control without any ectopic RBP7910-3v5 expression while with increasing amounts of tetracycline, we obtained robust expression of the protein (Fig. 10.2).

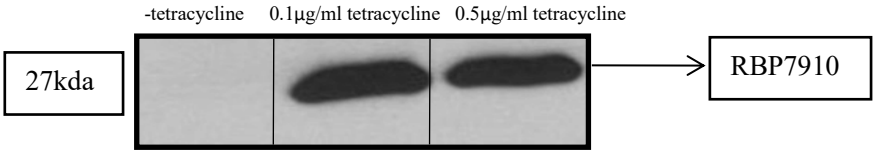


Fig 10.2: Procyclic (insect stage) cells expressing RBP7910 tagged with v5. *Tetracycline regulated cells in procyclic stage. Western blot analysis after introducing the gene in the rDNA spacer. Anti- v5 primary antibody used for detection of the v5 tag on RBP7910 which has a molecular weight of 27kDa.*

10.2 RBP7910 co-migrates with the core editosome complex and RNase A treatment of mito extract affects co-migration of RBP7910

To verify which multi-protein complex RBP7910 belongs to, glycerol gradient sedimentation experiments were conducted. Proteins and protein complexes sediment on the glycerol gradients based on their density, with the heavier complexes sedimenting towards the bottom while single proteins and lighter complexes sediment in the lower concentrations of glycerol on top. In Fig 10.3, the western blot shows the control (untreated) in Fig. 10.3A, RNase A treatment in Fig. 10.3B, DNase I treatment as a control to RNase A in 10.3C and salt (300 mM KCl) treatment in 10.3D. RBP7910-3v5 co-migrates with the catalytic editosome complex (RECC) and this co-migration is lost upon the treatment with RNase A, indicating that RBP7910 may be associating with RECC through RNA. Moreover, RNase A treatment reduces the RECC complexes to lower fractions indicating that the treatment has broken down the heavier 40S editosomes into 20S complexes. While high salt tends to have a similar effect on the RECC, it does not perturb the co-sedimentation of RBP7910 with RECC proteins, implying that its interaction with the RECC is not mediated electrostatically. On the other hand, DNase I control treatment does not affect the co-sedimentation profiles of RBP7910 and RECC significantly, further corroborating the effect obtained after RNase A treatment.

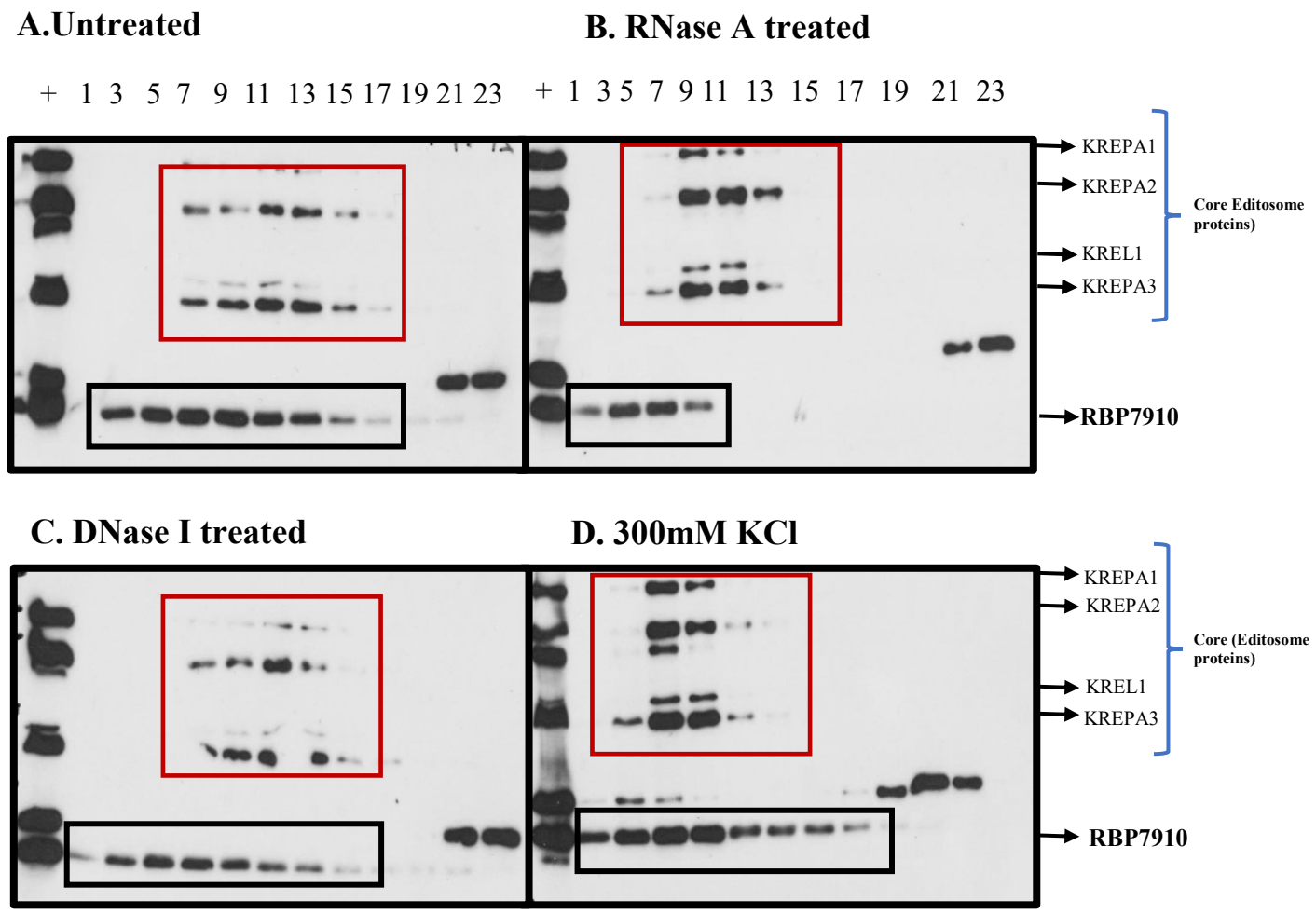


Fig 10.3: Western blot analysis; RBP7910 co-migrates with the core editosome complex and its sedimentation is affected by RNase A. *Migration of RBP7910 and the RECC on 10-30% glycerol gradients. The numbers on the top represent the fraction numbers from the top to the bottom of gradients. Only odd fractions were loaded on the gel. Bands seen on fractions 21 and 23 are a result of cross reactivity, hence classified to be non-specific bands. Different fractions of crude mitochondrial lysates were prepared and run after treatments with RNase, DNase and 300 mM KCl prior to running on the gradients. Antibodies against the RECC proteins (KREPA1, KREPA2, KREL1 and KREPA3) and against v5 (to represent the tagged RBP7910) were used to analyze the migration of the respective proteins on the gradients.* ***(A)*** *Migration of the proteins under untreated "normal" conditions,* ***(B)*** *following RNase A treatment.* ***(C)*** *Following DNase treatment, and* ***(D)*** *under conditions containing 300 mM KCl. The antibodies used for this western were; primary: anti- editosome/core complex and anti v5. Secondary- Anti Mouse HRPO.*

10.3 RBP7910 co-sedimentation with the editosome proteins is not greatly affected by increasing salt concentrations

To confirm if high KCl concentrations indeed had no effect on RBP7910 co-sedimentation with the RECC, glycerol gradients were prepared with 50 mM, 500 mM and 1 M KCl. Salt concentrations generally affect protein complexes as it perturbs protein-nucleic acid and protein-protein interactions that rely on charge [70]. Similar to 300 mM KCl treatment in the previous section, none of the KCl concentration tested, including 1M KCl, had a major effect on RBP7910 co-sedimentation with RECC (Fig. 10.4). While the RECC proteins moved toward the lower density fractions on the gradient, so did RBP7910. This effect intensifies upon increasing concentrations of KCl, leading to an accumulation of the complex around fractions # 5-7. This observation implies that RBP7910 may potentially interact with the RECC via RNA, while its presence in fractions heavier than #5-7 indicates its RNA-mediated interaction with other complexes as well. Furthermore, these data also demonstrate that RBP7910 is stable in high salt conditions, which is also a feature of other Z-DNA/RNA binding proteins. To elaborate, its RNA-mediated occupancy in multi-protein complexes may be governed by its recognition of substrate RNA structure and fold rather than sequence or charge, as Z-DNA/Z-RNA binding proteins do [70].

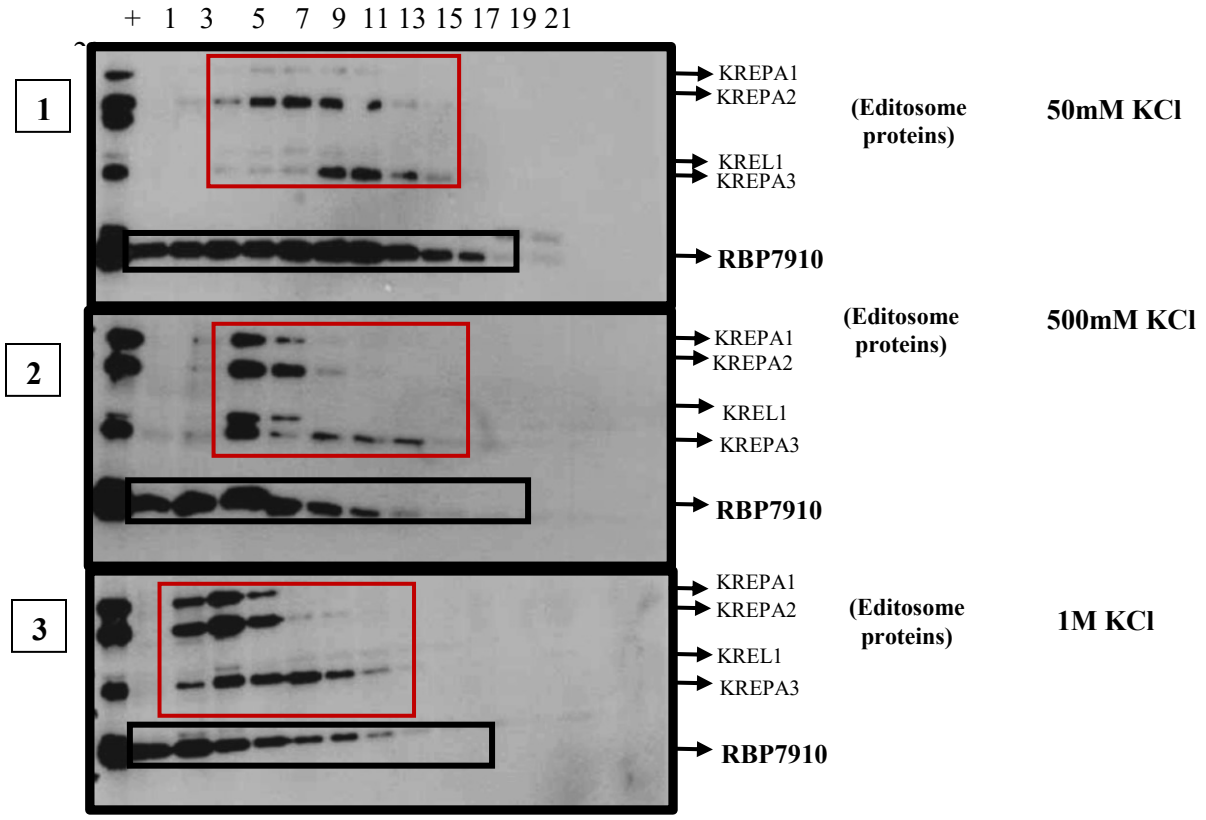


Fig 10.4: Western blot analysis of Salt gradients ranging from 50nM to 1M. *Panel 1, 2 and 3 are western blots wherein the different fractions of mitochondrial lysates were treated with 50mM, 500mM and 1M KCl respectively. The antibodies used for this western blot were primary: anti- editosome/core complex to detect the RECC complex (KREPA1, KREPA2, KREL1, and KREPA3) marked in the red box and anti v5 Ab to detect RBP7910-3v5 in the black box. Secondary Ab was Anti Mouse HRPO. The numbers on the top represent the fraction numbers from the gradients. Only odd fractions were loaded on the gel.*

10.4 Optimizing interaction kinetics of recombinant RBP7910 and its RNA substrates using Microscale Thermophoresis (MST) for future analysis

To experimentally establish if RBP7910 indeed interacts with Z-form of RNA, bacterial protein expression vectors were designed for purifying recombinant (r) RBP7910, rRBP7910 Z-alpha domain, and human ADAR Z-alpha domain (control).

10.4.1 Purifications and quantification of recombinant proteins.

All three proteins are his-tagged to facilitate purification, as seen in the western blot image below (*Fig 10.5*). Purification for RBP7910 Z-alpha domain was not as successful as the other two proteins, as it may be prone to degradation. We, therefore, decided to optimize MST with RBP7910 full length (FL) without the first 8 amino acids (import signal) and human ADAR Z-alpha proteins. The proteins were quantified using Image lab with calculating the intensities of the protein with the his-tag standards since accurate concentrations were required in order to analyze the binding affinities between the protein and substrate. Bradford assays were inaccurate since an error in one dilution may result in a faulty linear curve. The concentrations were quantified using the Imagelab software are summarized in the table below. His-tag markers were used as a standard to compare the protein concentrations.

Table 2: Concentrations of the proteins measured in µM with the protein weight in kDa:

Protein	Size of the protein	Concentration in µM
RBP7910-D8	27kDa	8.09µM
Human ADAR Z-alpha	8kDa	12.4µM

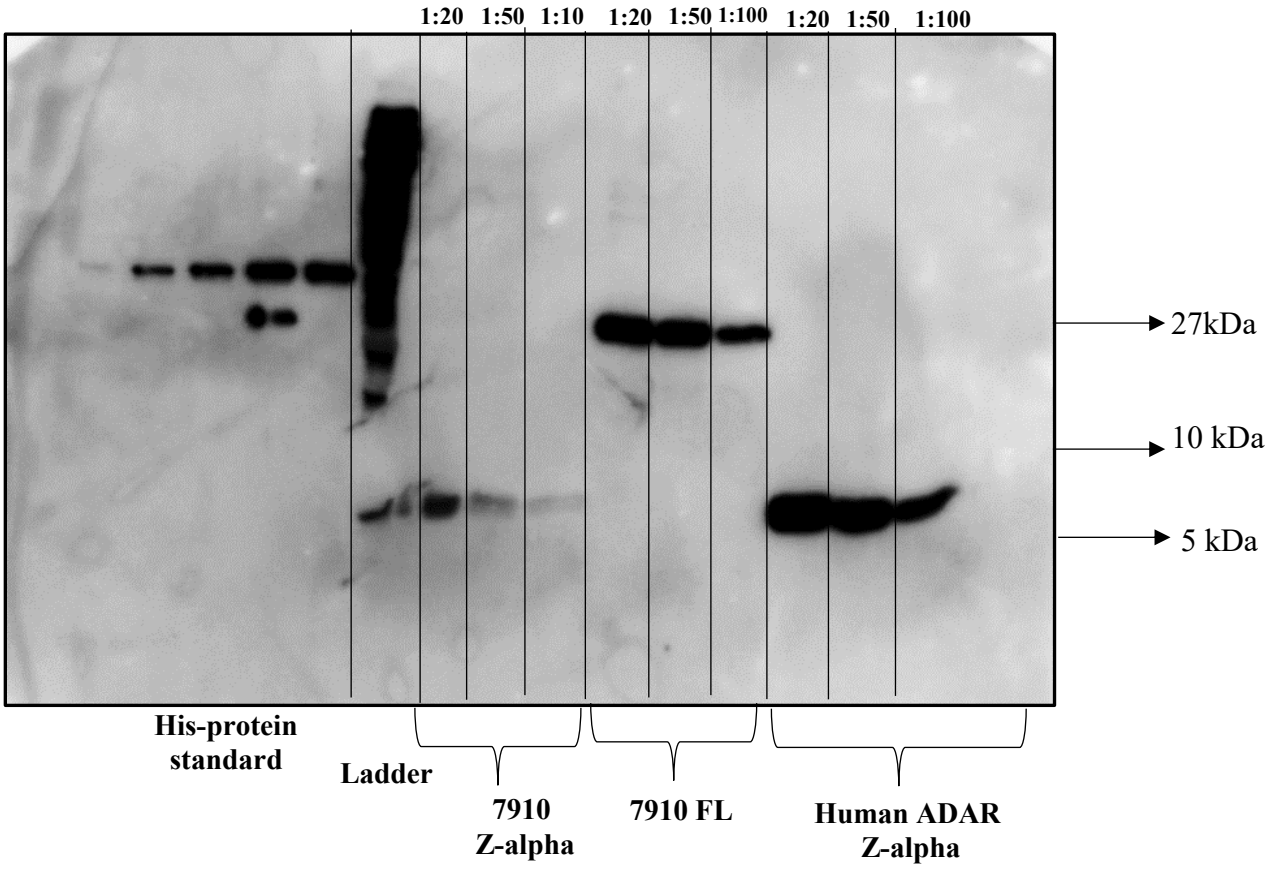


Fig 10.5: Purification of recombinant RBP7910 full length, RBP7910 Z-alpha and ADAR Z-alpha. *Western blot analysis in order to quantify the proteins purified.* ***His*** *tag standard protein was loaded in increasing amounts to create a standard curve. The concentrations of the standard were (20ng, 40ng, 80ng, 100ng, 200ng). Anti-6xhis as primary and anti-mouse HRPO as secondary antibodies were used for detecting the proteins on western blot. The protein quantifications were performed on Image lab software from Bio-Rad.*

10.4.2 Optimizing the detection limit of the fluorescently labeled target nucleic acids for RBP7910 and human ADAR Z-alpha on MST.

Nucleic acid substrates were labeled with Cy5 at the 3' end by Integrated DNA Technologies. The RNA molecule prepared as a target for RBP7910 contained AUUU repeats in its sequence as determined previously [65], whereas the DNA molecule prepared as a target for ADAR contained GC repeats [70]. Different concentrations ranging from 5µM-5nM of substrates were tested to obtain an optimal fluorescent concentration. The concentrations tested were 5µM, 1 µM, 500nM, 200nM, 100nM, 50nM, 20nM, 10nM, 5nM of both AUUU-Cy5 RNA and GC-Cy5 DNA. The two curves represent two technical replicates of the same sample. Fluorescence reading pretest in Fig 10.6 demonstrates that 5µM of both AUUU-Cy5 and GC-Cy5 had high levels of signal output in the capillary count leading to saturation in the substrates. On the other hand, 10nM concentrations were too low to be detected with sufficient output. Cy5-labeled probes at 50nM or 100nM of both AUUU and GC repeat sequences are ideal for analysis as they not only provide a good capillary count (above 200), but also follow MST traces as recommended by NanoTemper without fluorescence variation and aggregation. Therefore, 50nM of these probes were used for further downstream analyses.

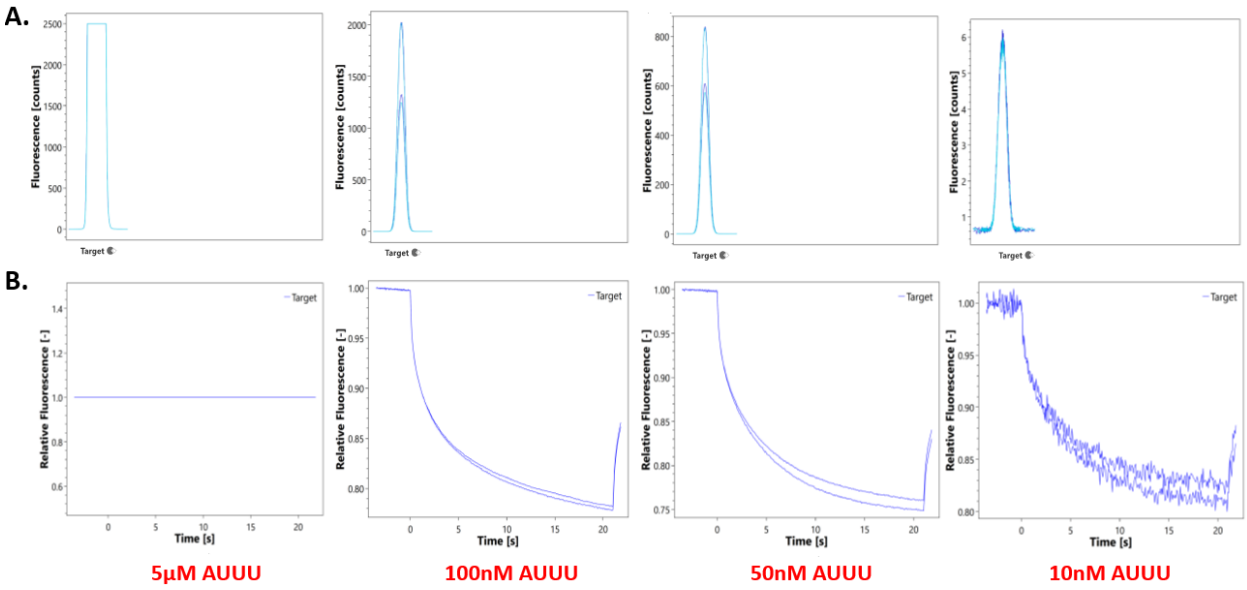


Fig 10.6: Cy5 fluorescence output in capillary curves and MST traces with 10nM, 50nM, 100nM and 5µM of the AUUU (truncated) repeat RNA probe: *Panel A shows the capillary scans and Panel B shows MST traces. The x-axis in panel A represents the Cy5 labeled probe while the y-axis represents the fluorescence counts. The x-axis in panel B represents the time in seconds during the read and the y-axis represents relative fluorescence output.*

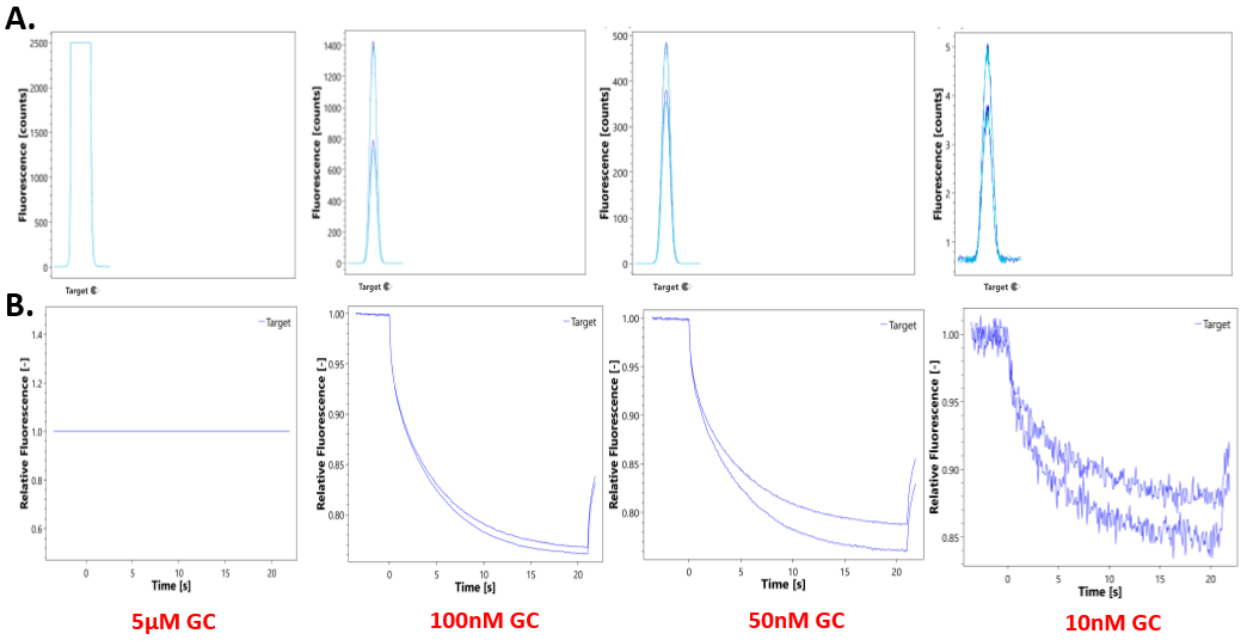
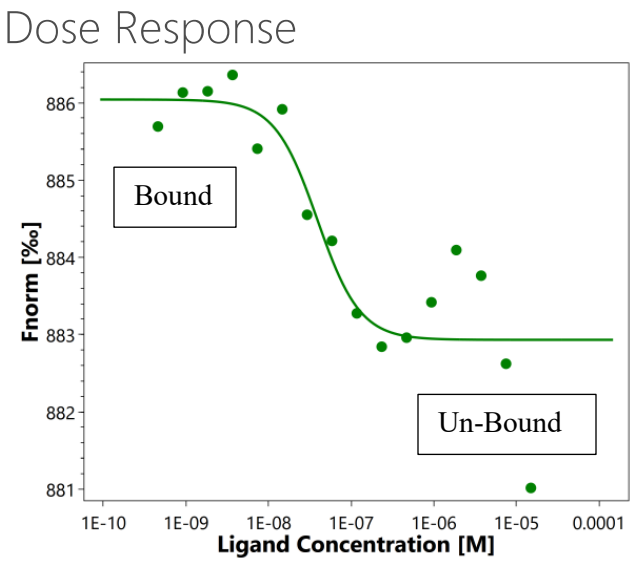


Fig 10.7: Cy5 fluorescence output in capillary curves and MST traces with 10nM, 50nM, 100nM and 5µM of the GC repeat DNA probe: *Panel A shows the capillary scans and Panel B shows MST traces. The x-axis in panel A represents the Cy5 labeled probe while the y-axis represents the fluorescence counts. The x-axis in panel B represents the time in seconds during the read and the y-axis represents relative fluorescence output.*

10.4.3 Equilibrium dissociation constants (k_d) of RBP7910 and ADAR Z-alpha proteins with their respective nucleic acid probes.

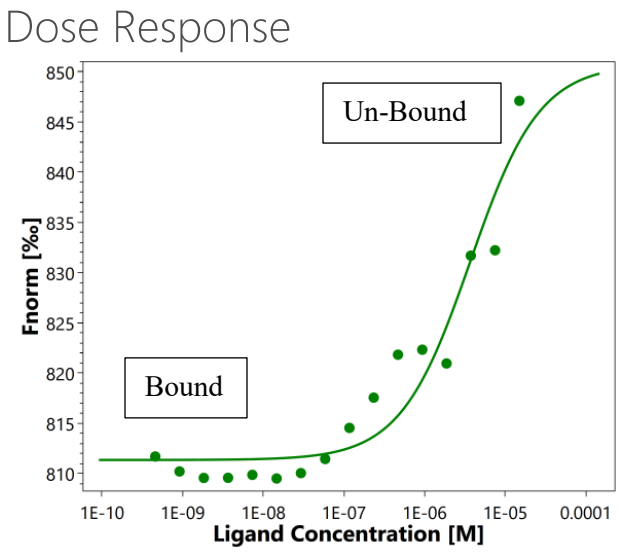
Based on the optimization results of fluorescence outputs from the labeled nucleic acid probes for optimal measurement in the previous section, we chose to perform kinetic experiments with 50nM substrate with serially diluted recombinant proteins. MST analysis on the interactions between RBP7910 and AUUU repeats resulted in a k_d value of 39nM while the interactions between ADAR Z-alpha and GC repeats resulted in a k_d value of 3.6μM. The response is calculated from the point the laser turns on for about 20 seconds until it switches off (see Figure 10.6 and 10.7). The fluorescence emitted by the nucleic acids drops upon increasing duration of the read as a microscopic temperature gradient is applied by the infrared laser to each capillary. The movement of the molecules in the temperature gradient change upon interaction. The final MST signal is plotted against the protein concentration to obtain a dose-response curve, from which we calculated k_d. The affinity curves look different because the curves are dependent on the F normalization [82].

While the mass diffusion represents the kinetics of depletion, S_T determines the steady-state concentration ratio c_{hot}/c_{cold} under a temperature increase ΔT. The normalized fluorescence $F_{norm}=F_{hot}/F_{cold}$ measures mainly this concentration ratio. Binding affinity is based on fluorescence and time (laser on to laser off). Therefore, the binding affinity is calculated between the bound and un-bound points [82].



RBP7910 FL- AUUU

k_d - 39nM



Human ADAR Z-alpha- GC

k_d - 3.6µM

Fig10.8: Equilibrium dissociation constants (k_d) calculated from dose response curve obtained through MST analysis on the interactions between RBP7910 – AUUU repeats and hADAR – GC repeats. *The x-axis represents the ligand concentrations of serially diluted RBP7910 and hADAR. The y-axis represents the Fnorm in %.*

11. Objective 2. To identify the interacting protein partners of RBP7910 by immuno-precipitation and mass spectrometric analysis.

11.1 Pulldown and elution of v5-tagged RBP7910

RBP7910-3v5 cell line developed in the previous objective was used for purifying the native proteins and its proteins partners from procyclic *T. brucei*. As RBP7910 was observed to co-migrate with the RECC complex, immunoprecipitations were performed in order to not just confirm its interaction with RECC, but also identify the other proteins / protein complexes it interacts with. Since the proteins were bound to anti-v5 conjugated agarose beads, elution via v5 peptide competition was the only conditions ideal for native purifications. Three conditions were applied to these purifications: no Tetracycline (no RBP7910-3v5 expression), +Tetracycline (expression) and +Tetracycline (expression) with +RNase A treatment during precipitation. In the western blot (Fig 11.1), the first panel contains immunoprecipitations performed without Tetracycline induction as a negative control, hence the elution does not contain any recognizable proteins. The second panel contains Tetracycline induced purification (positive) and the third panel contains Tetracycline induced and treated with RNase A before elution from the beads. RNase A treatment of the immune-precipitated protein eliminates proteins bound to RBP7910-3v5 indirectly through RNA and should only contain proteins bound directly. The three eluates obtained were used for further downstream analyses such as mass spectrometry and RNA sequencing.

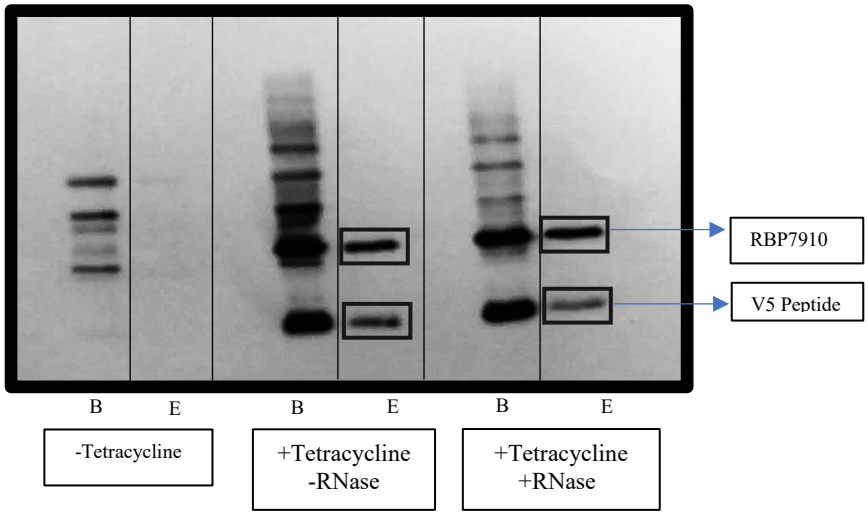


Fig 11.1: Purification of v5-tagged RBP7910 using anti-v5 conjugated agarose beads eluted through v5-peptide competition. *The antibodies used in this western blot are anti-v5 (primary) and anti-mouse HRPO (secondary) to detect the tag on RBP7910. RBP7910-3v5 and v5 peptide are marked through the blue arrows. Proteins bound on beads (B) and the final eluates (E) are indicated on the western blot.*

11.2 RBP7910-3v5 MS analysis: Its RNA dependent and independent interacting proteins partners

The three eluates prepared in the previous section were analyzed by Liquid Chromatography coupled to Mass Spectrometry (LC-MS). The eluate prepared without Tetracycline induction was used as a control to eliminate all the false positives, i.e. the proteins that bind to the beads non-specifically. While the second eluate (Tetracycline induced) provided a list of all proteins precipitated with RBP7910-3v5, the RNase A treated eluate helped to identify the proteins that interacted with RBP7910-3v5 via RNA. Over 200 proteins were processed in four steps with stringent parameters to obtain an accurate set of proteins that interacted truly with RBP7910 (Appendix 1). First, the proteins count obtained in both induced eluates were normalized with respect to the target protein, RBP7910. For instance, the RNase A treated eluate contained 9 counts of RBP7910 while the untreated (no RNase A) contained 3 counts of the same, so the list of

proteins in the untreated MS analysis was multiplied by 3 to match the RBP7910 count in the treated one. The resulting list of proteins was then annotated with their gene IDs, names and biological function (if available). Second, all proteins with a peptide count < 2 were eliminated as it is considered inaccurate. Third, in order to eliminate the false positives, protein counts obtained from MS analysis of the uninduced eluate were subtracted from the induced ones. Moreover, the proteins that do not belong to the mitochondrion, where RBP7910 is found, were also eliminated using https://tritrypdb.org/tritrypdb/ as reference (Table 3). Finally, the Normalized Spectral Abundance Factor (NSAF) values for each identified protein (spectra) was calculated to analyze the edge length. The spectral counts were summed for all identified proteins and divided by their respective lengths, generating spectral abundance factor (SAF) values. The SAF value of each identified subunit (a component of a recognized multi-protein complex) was then normalized against the sum of all SAFs within an individual biological sample, resulting in the normalized SAF (NSAF). The NSAF values correspond to the edge length and the peptide count corresponds to the peptide abundance. The MS-based protein abundance feature originates from spectral counts, which, in the simplest form, counts the number of MS spectra assigned to one protein as its abundant feature. To do so, normalized spectral abundance factor (NSAF) was used:

$$(\mathrm{NSAF})_J = \frac{(\mathrm{Sc}/L)_J}{\sum_{i=1}^{N} (\mathrm{Sc}/L)_i}$$

The definition of NSAF, wherein Sc is the spectral count of protein *J* and *L* is the lengths of protein *J* and *N* is the total number of proteins [83].

According to NSAF values and the abundance of the protein, they were differentiated into RNA dependent and RNA independent interactors. The raw data of proteins pulled down with RBP7910 is mentioned the ***Appendix 1.***

Table 3: The final set of proteins obtained after the stringent analysis: *The highlighted proteins display the relevant protein partners after going through the analysis mentioned above.*

ANNOTATIONS	Tetracycline + (-RNase)	Tetracycline (+RNase)
RNA editing mediating complex (REMC)		
Tb927.10.10130 REMC1	9	2
Tb927.2.1860 REMC2	6	2
Tb927.8.8180 REMC4	6	0
Tb.927.7.800 REMC3	24	6
Tb927.8.8180 REMC 5A	24	0
Universal minicircle sequence binding protein (UMSBP)		
Tb927.10.6060 UMSBP1	3	0
Tb10.70.0800 UMSBP2	21	0
Tb927.10.10830 RGG2	6	16
Tb927.6.2230 RGG1	30	0
Tb927.7.320 RBP8	3	0
Tb927.10.7910 RBP7910	9	9
Tb927.11.15640 MERS1	9	4
Polyadenylation mediating complex (PAMC)		
Tb927.1.1730 PAMC1	36	3
Tb927.6.1200 PAMC2	6	0
Tb927.10.1730 PAMC3	9	0
Tb927.1.3010 PAMC4	18	1
Guide RNA binding complex (GRBC)		
Tb927.7.2570 GRBC 1	18	0
Tb927.2.3800 GRBC 2	12	1
Tb927.11.16860 GRBC3	12	3
Tb11.02.5390 GRBC4	30	0
Tb927.10.11870 GRBC5	9	2
Tb927.5.3010 GRBC6	15	3

Tb927.4.1500 REH 2	21	0
Tb927.4.3020 ATP-dependent DEAH-box RNA helicase	3	0
Tb927.2.3180 KPAF1	9	0
Tb927.9.12770 KPAF3	3	0
Tb927.10.5830 Protein coding	9	6
Tb927.6.1440 LSU ribosomal protein, mitochondrial	4	0
Tb927.3.2230 Succinyl-CoA synthetase alpha subunit	4	0
Tb927.11.1710 gBP21 guide RNA-binding protein	4	0
Tb927.7.2700 B5R NADH-cytochrome b5 reductase	9	0
Tb927.10.4080 Hypothetical protein	3	0

Table 4: The calculated NSAF values of each protein predicting the edge and interactions with RBP7910:

ANNOTATIONS	Tetracycline+ (-RNase)	Tetracycline+ (+RNase)	Protein length (L)	Sc/L (-RNase)	Sc/L (+RNase)	NSAF (-RNase)	NSAF (+RNase)
Tb927.10.10130 REMC1	9	2	545	0.016513761	0.003669725	0.019993966	0.022474078
Tb927.2.1860 REMC2	6	2	872	0.006880734	0.002293578	0.008330819	0.014046299
Tb927.8.8180 REMC4	6	0	934	0.006423983	0	0.00777781	0
Tb.927.7.800 REMC3	24	6	543	0.044198895	0.011049724	0.053513623	0.067670568
Tb927.8.8180 REMC 5A	24	0	934	0.025695931	0	0.031111239	0
Tb927.10.6060 UMSBP	3	0	213	0.014084507	0	0.017052757	0
Tb10.70.0800 UMSBP	21	0	140	0.15	0	0.181611859	0
Tb927.10.10830 RGG2	6	16	320	0.01875	0.05	0.022701482	0.306209318
Tb927.6.2230 RGG1	30	0	794	0.037783375	0	0.04574606	0
Tb927.7.320 RBP8	3	0	180	0.016666667	0	0.020179095	0
Tb927.10.7910 RBP7910	9	9	189	0.047619048	0.047619048	0.057654559	0.291627922
Tb927.11.15640 MERS1	9	4	395	0.02278481	0.010126582	0.027586612	0.062017077
Tb927.1.1730 PAMC1	36	3	693	0.051948052	0.004329004	0.062895882	0.026511629

Tb927.6.1200 PAMC2	6	0	468	0.012820513	0	0.015522381	0
Tb927.10.1730 PAMC3	9	0	357	0.025210084	0	0.030523002	0
Tb927.1.3010 PAMC4	18	1	372	0.048387097	0.002688172	0.058584471	0.016462867
Tb927.7.2570 GRBC 1	18	0	473	0.038054968	0	0.04607489	0
Tb927.2.3800 GRBC 2	12	1	492	0.024390244	0.00203252	0.029530384	0.012447533
Tb927.11.16860 GRBC3	12	3	482	0.024896266	0.006224066	0.030143047	0.038117343
Tb11.02.5390 GRBC4	30	0	1087	0.027598896	0	0.033415246	0
Tb927.10.11870 GRBC5	9	2	310	0.029032258	0.006451613	0.035150682	0.03951088
Tb927.5.3010 GRBC6	15	3	516	0.029069767	0.005813953	0.035196097	0.035605735
Tb927.4.1500 REH 2	21	0	2167	0.009690817	0	0.011733115	0
Tb927.4.3020 ATP-dependent DEAH-box RNA helicase	3	0	1440	0.002083333	0	0.002522387	0
Tb927.2.3180 KPAF1	9	0	1003	0.008973081	0	0.010864119	0
Tb927.9.12770 KPAF3	3	0	1022	0.002935421	0	0.003554048	0
Tb927.10.5830 Protein coding	9	6	546	0.016483516	0.010989011	0.019957347	0.067298751
Tb927.6.1440 LSU ribosomal protein, mitochondrial	4	0	258	0.015503876	0	0.018771252	0
Tb927.3.2230 succinyl-CoA synthetase alpha subunit	4	0	301	0.013289037	0	0.016089644	0
Tb927.11.1710 gBP21 guide RNA-binding protein	4	0	936	0.004273504	0	0.005174127	0
Tb927.7.2700 B5R NADH-cytochrome b5 reductase	9	0	287	0.031358885	0	0.037967636	0
Tb927.10.4080 hypothetical protein	3	0	1183	0.002535926	0	0.003070361	0
				-rnase SUM sc/l	+rnase SUM sc/l		
N = total # of proteins	384	58		0.825937252	0.163286997		
0.825937252							

The Cytoscape software was used in order to create a protein-protein interaction map. The edges serve as a visual indicator of the abundance of proteins (Fig 11.2). Interacting protein partners were classified into two groups: RNA dependent and RNA independent protein partners. Functions of some of the important proteins have been mentioned below.

RBP7910 appeared to mainly interact with proteins in the RESC such as GRBC, PAMC and RGG2 (a protein containing an RNA recognition motif). Two novel unannotated proteins Tb927.10.5830 (RNA independent binder) and Tb927.10.4080 (RNA-dependent binder) known to be essential for parasite growth and survival in the PF, BF and also for differentiation from BF to PF based on data obtained from a large scale RNAi study [84]. UMSBP1 and UMSBP2 (RNA dependent interactors) are minicircle proteins that were found in abundance and help in the kinetoplast development in a parasite. The UMSBPs play a role in kDNA replication initiation and segregation as well as in mitochondrial and nuclear division, which implicates a potential role for UMSBPs in linking kDNA replication and segregation to the nuclear S-phase control during the trypanosome cell cycle [85].

MRP1(RNA dependent interactor) formerly known as the gRNA binding proteins, gBP21 and gBP25, the mitochondrial RNA binding proteins 1 and 2 (MRP1/2) form a ~100 kDa stable heterotetrametric complex [86], [87] that associates with the RECC and the RESC in a RNA-dependent manner at low salt concentrations [88]. In addition to its RNA binding activity, this complex promotes RNA annealing [86]. Moreover, immuno-depletion of MRP1 from mitochondrial fractions active in RNA editing leads to suppression of *in vitro* editing [89]. Succinyl CoA and B5R NADH cytochrome are proteins (RNA dependent interactors) which are involved in the oxidative phosphorylation and play a major role in the stability and energy production of the protein.

As RBP7910 is a potential Z-binding protein, REH2 (RNA helicase 2) (an RNA dependent interactor) may potentially assist Z-conformation which is a characteristic behavior of enzymes such as polymerases and helicases. Z-DNA is stabilized both by negative super helical stress and by binding of the Zα domain of Z-binding proteins [90]. REH2 functions with the core editosome (RECC), which plays a role in unwinding the gRNA mRNA duplexes during RNA editing [91]. RBP8 is an RNA binding protein which interacts with RBP7910 in an RNA dependent manner.

A very interesting protein interactor, TbRGG2 (RNA independent interactor) was identified in abundance in the RNase A treated eluate vs untreated eluate, indicating it to be a direct interactor of RBP7910. Previous studies also show that immuno-precipitated RGG2 interacted with RBP7910 [92]. RGG2 also contains an RNA recognition motif which potentially could assist RBP7910 in its Z-form. MERS1 (RNA independent interactor) which belongs to the PPsome interacts with RBP7910 directly and this bond has previously been studied. MERS 1 targets mRNA 5′ termini *in vivo* and displays an *in vitro* hydrolase activity which is stimulated by the RNA-editing substrate-binding complex (RESC) [66].

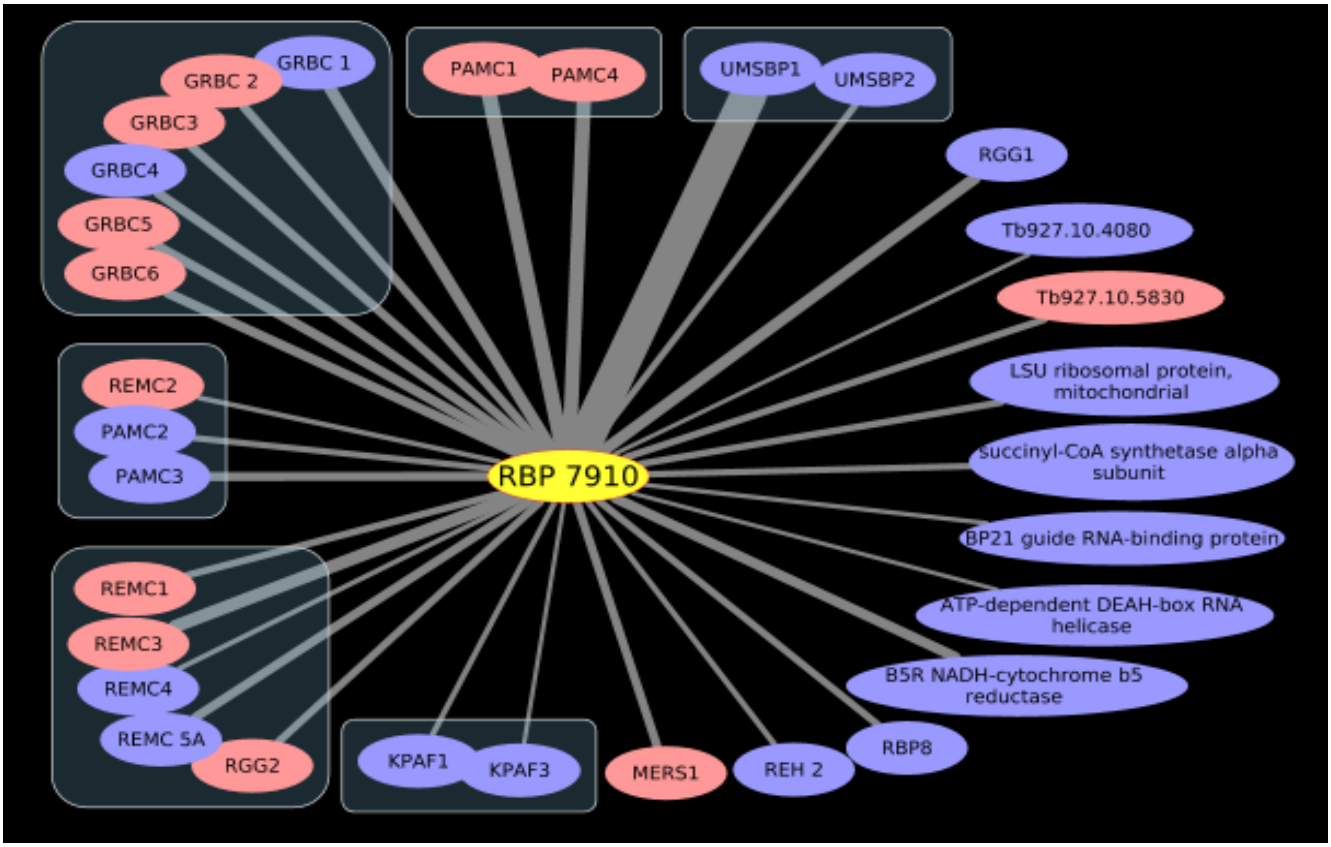


Fig11.2: Cytoscape analysis of RBP7910 and its RNA dependent and RNA independent protein partners. ***Edges****: protein abundance in –RNase condition.* ***Pink****: proteins that remain in complex after RNase treatment (RNA-independent interactors)* ***Violet****: proteins that are not detected when treated with RNase. (RNA-dependent interactors).*

11.3 RNA editing ligase (REL) adenylation assay demonstrates that RBP7910 interacts with the RECC weakly through RNA

The eluted proteins obtained from the three conditions were subjected to a western analysis against RECC proteins using antibodies against KREPA1, KREPA2, KREL1 and KREPA3. However, these proteins were not detected using western analysis (data not shown). Therefore, we decided to perform a ligase adenylation assay that detects the ligase-adenylate intermediate formed after covalent catalysis of radioactive [α - ^{32}P] ATP, for increased sensitivity. As seen in figure 11.3, immunoprecipitation of RBP7910-3v5 does recruit the RECC in an RNA dependent manner. Since we did not observe these RECC proteins on a western blot of MS analysis, we conclude that RBP7910's interaction with the RECC is limited compared to the other complexes involved in RNA editing.

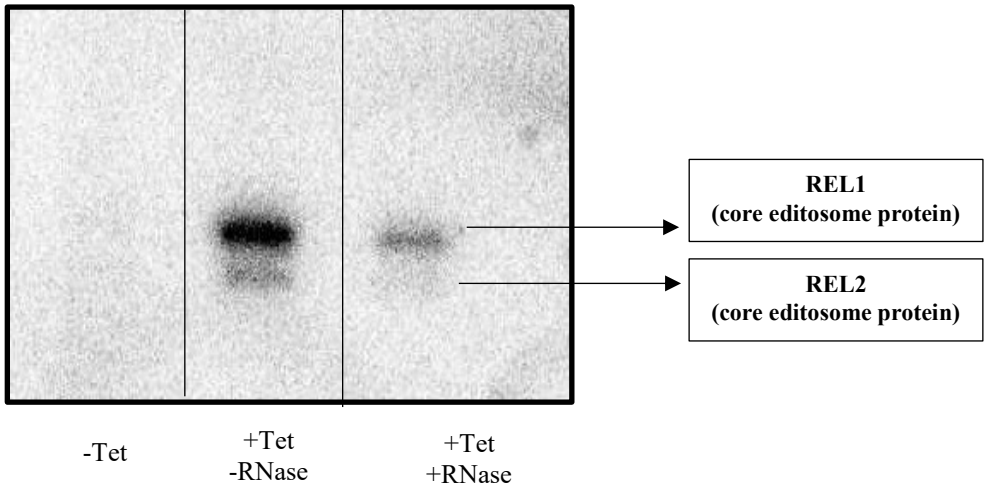


Fig11.3: RNA editing ligase (REL) adenylation assay indicates RBP7910 interaction with the RECC is RNA dependent. *Ligase-adenylate intermediates (for REL1 and REL2) were detected using [α – ^{32}P] ATP. Samples run on denaturing PAGE and developed using Phosphorimaging.*

12. Objective 3. Understanding the importance of RBP7910 in mammalian bloodstream form (BF) *T. brucei*.

12.1 RBP7910 is essential for the growth and survival of BF parasites

Wild-type BF single marker cell line was transfected with p2T7-7910 RNA interference (RNAi) vector that recombines in the rDNA spacer and expresses a double-stranded RNA fragment of the RBP7910 gene for downstream silencing of the gene, as done previously with the PF cell line [64]. The BF growth curve exhibits robust growth of the cultured cell line without induction (-tetracycline) of RBP7910 silencing (Fig. 12.1). Upon induction, the cell suffers severe growth defect with a decline in the parasite population as early as the 14 – 21h period post knockdown, implying that RBP7910 is essential in the BF stage of the parasite as well. This effect is, however, considerably drastic in the BF cells than the PFs in comparison. By 49 h post-induction, no surviving parasites were observed on the microscope, indicating that RBP7910 silencing is fatal.

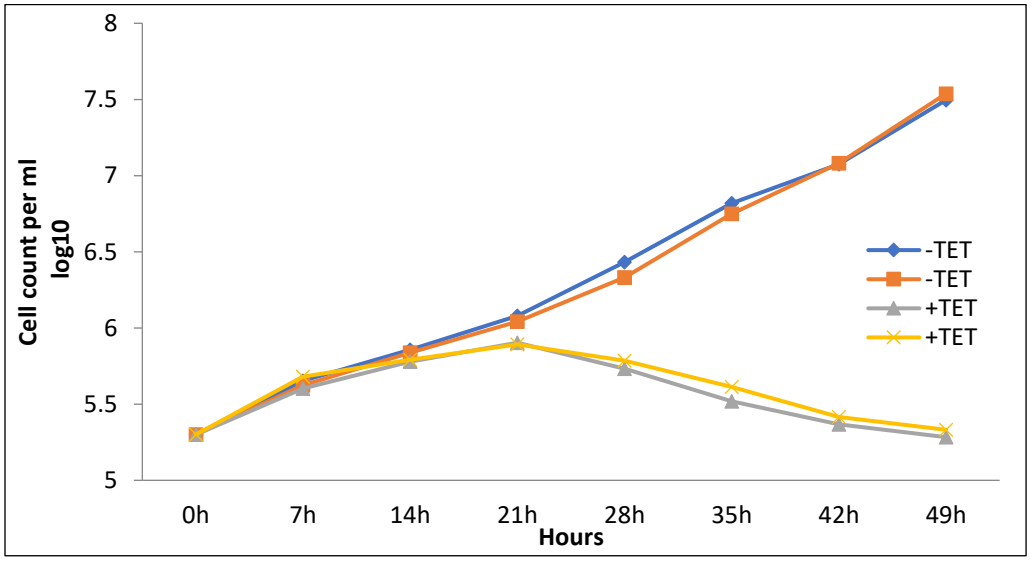


Fig12.1: Growth curve of RBP7910 knockdown in BF T. brucei. *The x-axis represents the time in hours post RNAi induction and the y-axis represents the cell count/ml. The orange and the blue data set are the growth curves of the uninduced cell line (–tetracycline) and the yellow and grey data sets are the growth*

curves of the induced cells (+tetracycline). A growth defect was observed after starting 14h post RNAi induction.

12.2 RBP7910-3v5 ectopic copy introduced in the rDNA spacer of BF *T. brucei*

BF single marker (SM427) cells were transfected with pLew100-3v5 vector containing RBP7910 into its rDNA spacer of the genome similar to the PF transfection done in section 8.1. Protein expression from this ectopic gene copy is controlled by a tetracycline promoter, overexpressing the protein as this adds to the native copies of RBP7910 expressed endogenously. The expression works similarly upon tetracycline induction as compared to the PF cell line.

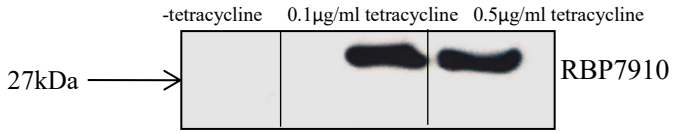


Fig12.2: BF cells expressing RBP7910-3v5 regulated by tetracycline. *Tetracycline regulated cells in BF. Western blot analysis after introducing the gene in the rDNA spacer. Anti- v5 primary antibody used for detection of the v5 tag on RBP7910 with a molecular weight of 27kDa.*

12.3 RBP7910 is localized in the mitochondrion of BF stage

Immunofluorescence assay (IFA) was performed to confirm the mitochondrial localization of RBP7910 as determined previously in the PF cells [64]. The cells were incubated with DAPI for staining the nuclei and kinetoplast, and Mitotracker to stain the mitochondria. RBP7910-3v5 was localized with anti-v5 primary antibody and anti-mouse secondary antibody conjugated with Alexa fluoro 488. The DIC image was taken to visualize the cells. From Fig.12.3, it is clear that RBP7910 staining matches with that of mitotracker, validating the mitochondrial localization of the protein.

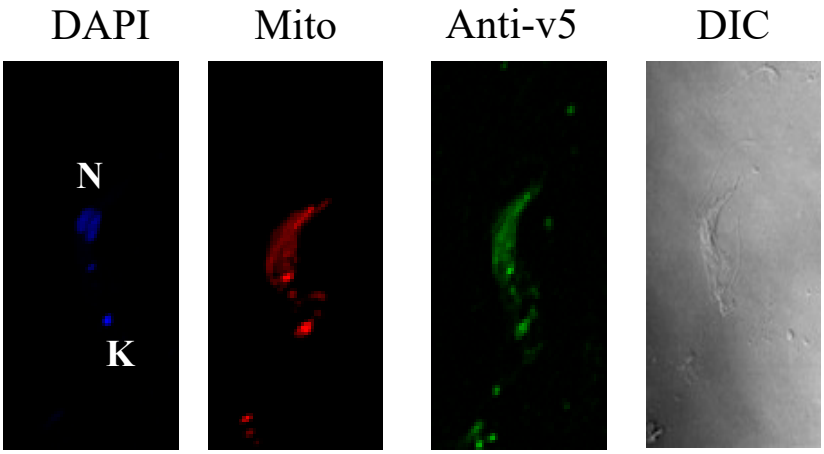


Fig 12.3: RBP7910 is localized in the mitochondria of BF stage. *IFA was performed using the stains (DAPI, Mito tracker FM and Alexa fluoro 488) to localize RBP7910-3v5 in the mitochondria of BF parasite.*

12.4 Over expression of RBP7910 in BF results in cells with larger mitochondrion

The mitochondrion was larger in several cells from the induced culture (Fig 12.4B) when compared to the un-induced condition (Fig 12.4.A), under the fluorescent microscope. Moreover, the mitochondrion appears more developed than usual, as normally observed in PF cells.

Cells not treated with tetracycline were used as un-induced control where RBP7910 is not over expressed, while the induced cells exhibit the larger mitochondrion with developed cristae. RBP7910 over expression is potentially regulating RNA editing and expression of specific protein components of the electron transport chain that are usually only produced in the PF stage. The assumption that RBP7910 may be regulating RNA editing of specific mitochondrial genes comes from an RNAi study, wherein RBP7910 knockdown in PF reduced the abundance of Cyb, RPS12 and MURF2 and increased the abundance of COIII, COII and ATPase 6 edited mRNA [64] . Therefore, it is likely that overexpression may induce a similar effect and lead to the expression of components that drive the mitochondrial respiration that occurs only in PF and the short stumpy BF stage that are ready to differentiate into the PF. Further analysis, such as qPCR on the pre-edited and edited mRNA will be required to prove this hypothesis. Also, the use of a pleomorphic

cell line will benefit this experiment as they are amenable to BF to PF differentiation experiments, as opposed to a monomorphic cell line that was used here. RBP7910 may potentially be the first mitochondrial protein to induce such a life cycle shift differentiation.

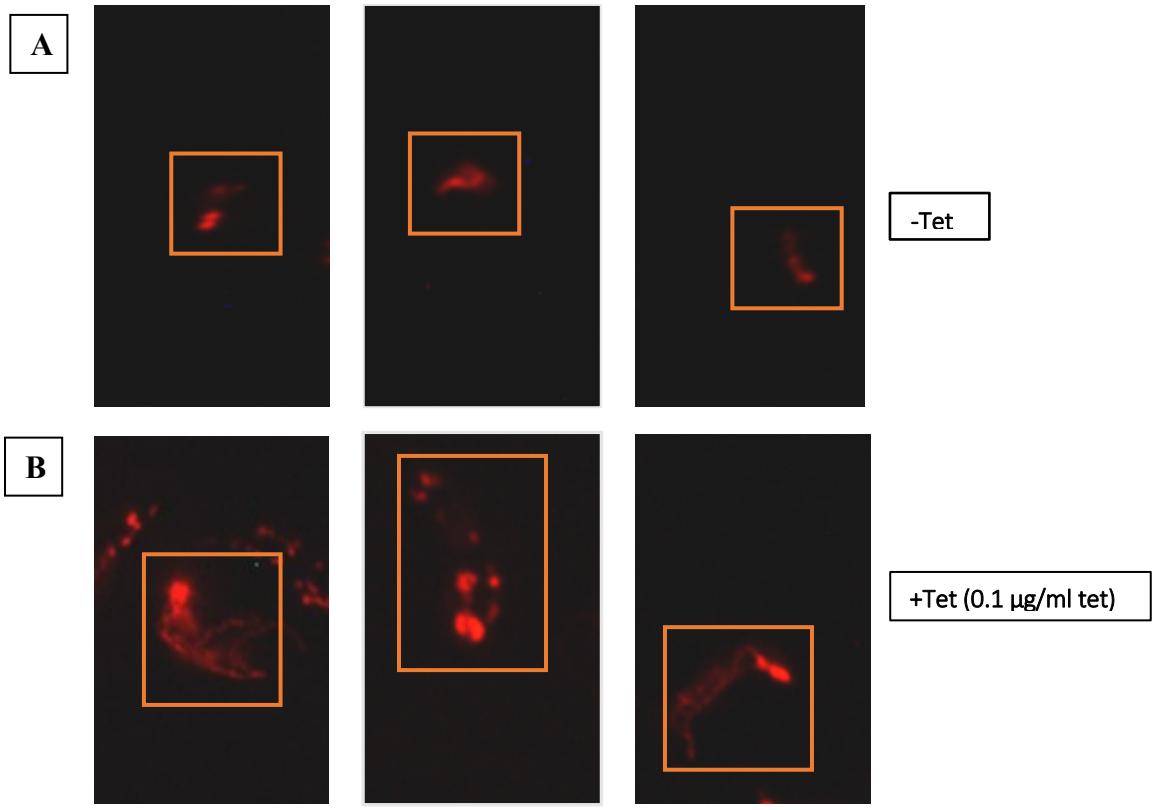


Fig12.4: Over expression of RBP7910 leads to a more developed mitochondrion in the bloodstream stage. *IFA was performed using Mito tracker FM to study the effect of overexpressing RBP7910 in the BF parasite. The 3 panels on the top 12.4.A contain -Tetracycline cells (RBP7910-3v5 un-induced) which serves as a control and represents normal BF cells, while the 3 panels below 12.4.B contain BF cells induced with Tetracycline (over-expression of RBP7910), and exhibit larger mitochondria.*

12.5 Overexpression of RBP7910 results in cells containing multiple nuclei possibly arresting the cells in their G0 phase of division

Normally dividing BF cells contain one nucleus and one kinetoplast (1K1N) in each cell (Fig.12.5 A). Upon over-expressing, RBP7910 cells with multiple nuclei and kinetoplasts (2K2N) were observed (Fig 12.5.B). The uninduced cells (-tetracycline) that do not overexpress RBP7910; a majority of these cells contained a single kinetoplast and nuclei. On the other hand, overexpression of RBP7910 (+tetracycline) cells with multiple nuclei and kinetoplasts were observed. Multiple nuclei indicate that the cells may be arrested in their G0 phase of the division [93]. As mentioned in the previous section, performing this experiment with a pleomorphic cell line (that are amenable to differentiation experiments) as opposed to the monomorphic cell lines used here, will benefit our understanding of this phenotype in greater detail.

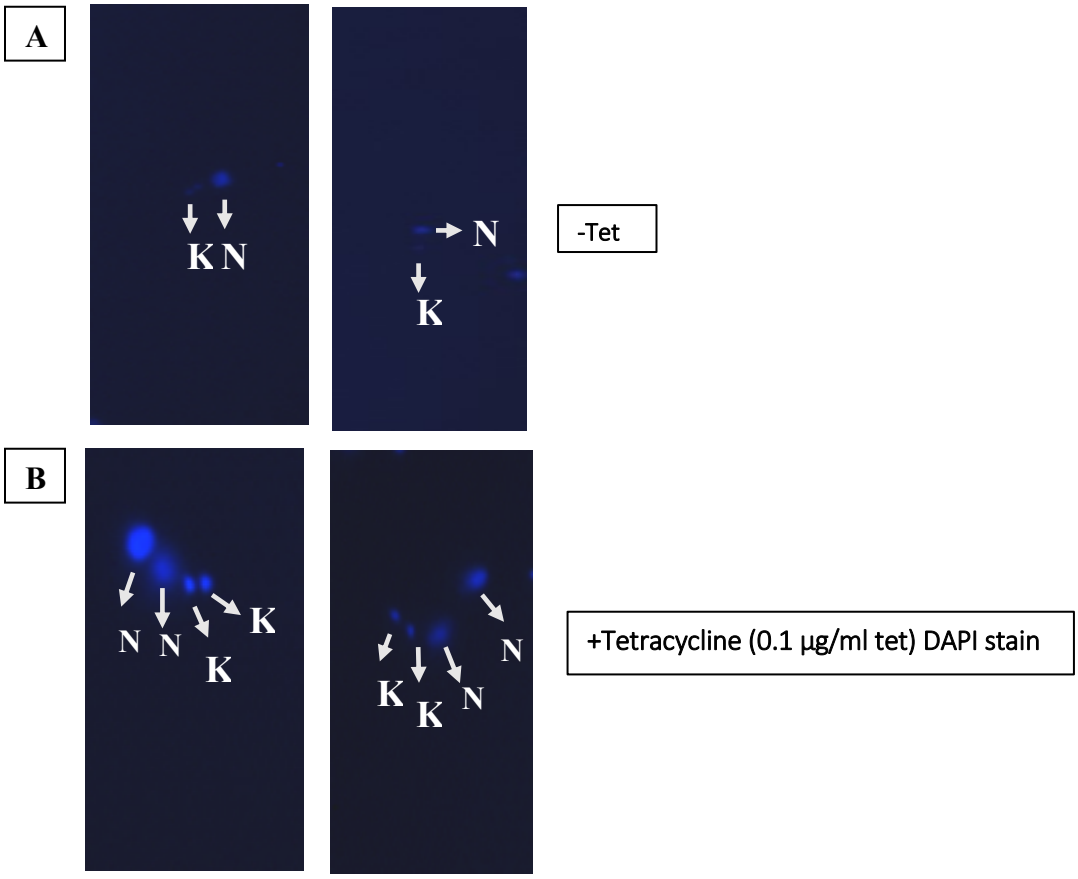


Fig12.5: Overexpression of RBP7910 results in a possible cell cycle arrest at G0 phase. *IFA was performed using DAPI stain to analyse the nucleus and kinetoplasts of the BF parasite. The 2 panels on the top 12.5.A contain un-induced cells (-Tetracycline) which serves as control and generally contain a single nucleus (1N) and kinetoplast (1K) while the 2 panels below 12.5.B contain RBP7910 over expression with cells containing multiple nuclei and kinetoplasts(2K2N) indicating that the cells are entering the arrest phase or the G0 phase.*

12.6 RBP7910 over expression in BF induces expression of precursor procyclins proteins

To complement the conclusions made from the IFA experiments, western blot analysis was performed using procyclin antibody, a protein marker that is expressed in PF stage. This antibody was against the EP1 (GLU-PRO repeats) of the procyclin family. The (Fig 12.6) below indicates presence of precursor procyclins (EPs) in BF cells overexpressing RBP7910. While the first three lanes in the western blot are whole cell lysates from un-induced BF cells (-tetracycline) and RBP7910-3v5 overexpressing (induced) BF cells, the last three lanes are different quantities of wild-type PF cells used as a control for the procyclin antibody. The procyclins detected in the BF appear to be unprocessed (precursor) procyclins (early EP1) as the cells are potentially arrested in their G2 phase before they could evolve into PF. Antibody against Phosphoglucokinase (PGK A, B, C isoforms) was used as a loading control. *T. brucei* has three PGK genes. One, PGKA, encodes a minor glycosomal variant that is expressed at low levels in both BF and PF. The second gene, PGKB, encodes the major cytosolic enzyme PGKB, which is present only in PF. The third gene, PGKC, encodes the major glycosomal enzyme PGKC, expressed only in BF [94].

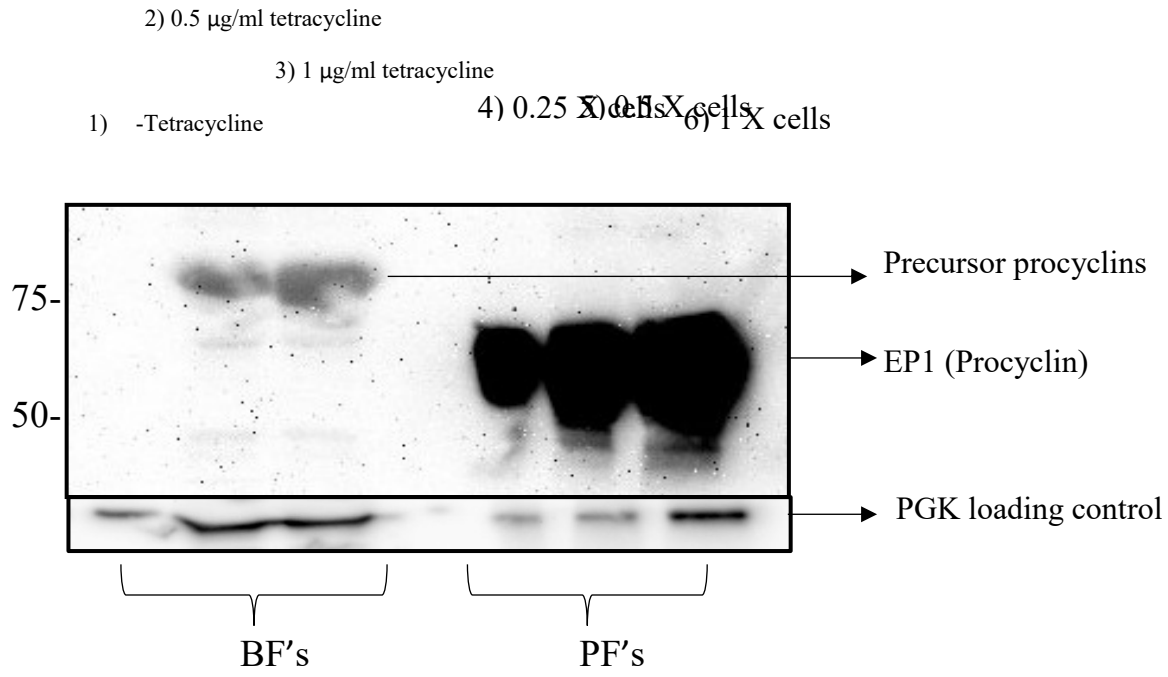


Fig 12.6: Overexpression of RBP7910 in BF results in expression of precursor procyclins. *Western blot analysis of overexpressing RBP7910-3v5 in the bloodstreams showed traces of precursor procyclins in increasing amounts as the Tetracycline concentrations increased. Sample 1 shows –Tetracycline (un-induced protein) which shows no traces of procyclins. Samples 2 and 3 show bloodstream stage cells producing early procyclins on RBP7910 over-expression. As a control, Procyclic cells were loaded in increasing amounts (samples 4, 5, 6). The loading control was PGK A, B, C shown in the bottom panel of the western blot. Primary Ab was α-procyclin Cedarlane® and secondary Ab was α-mouse HRPO (Biorad™).*

13. Discussion

Of most trypanosomatid RNA binding proteins, RBP7910 is unique as it is predicted to contain Z-RNA binding domains. The objective of this project was further to characterize RBP7910 as an RNA binding protein further. According to the glycerol gradients, RBP7910 was notably sensitive to RNase treatment as it lost contact with the core editosome hence suggesting that RBP7910 interacts with RNA editing complexes via RNA. KREL1, KREPA2, KREPA3, and KREPA4 are part of the core catalytic editosome complex (a.k.a. RECC) that here represents the whole complex on western blots. RBP7910 did not have any effect on treatment with DNase, which was used as control, confirming it as an RNA binding protein.

Contrary to RNase treatment, salt (including gradients containing up to 1M salt) did not have a considerable effect or change on RBP7910 co-sedimentation with the RECC and into fractions with heavier densities. This observation was contrasting with the IEX (Ion exchange chromatography) fractionation data published previously that suggested the protein to be salt-sensitive [64]. While the gradient sedimentations were performed at 4°C, the IEX was performed at room temperature, and possibly causing RBP7910 to salt out due to temperature sensitivity. Recombinant RBP7910 purification is sensitive to degradation when not provided optimum conditions (4°C), and therefore the native RBP7910 may behave the same which could explain why it fractionated first on IEX analysis. As most Z-DNA/RNA binding proteins bind their substrates structurally and via charged residues [70], RBP7910's tolerance to salt in the glycerol gradients indicates that it may potentially be a Z-binding protein. Prediction of the Z-binding domains have been quite intriguing, as it provides an additional mechanism through which RNA editing could be regulated *in vivo*. This Z-binding property could be tested *in-vitro* as well as *in-vivo.* Microscale thermophoresis is useful in analysing protein binding kinetics and providing accurate k_d values in a variety of conditions tested. Previous analysis via electromobility shift assay (EMSA) on rRBP7910 binding kinetics estimated a k_d value for the wild-type (WT) interacting with the U-tail-bearing A6 guide RNA substrate to be 0.21 ± 0.01 nM [65]. The k_d for RBP7910-AUUU with thermophoresis was noted to be 39nM. The difference between the two binding affinities could be an attribute of the two different RNA substrates used along with the different methodologies implemented in estimating the WT RBP7910 k_d. The k_d measured for human ADAR Z-α domain with its poly GC repeats Z-DNA substrate using circular dichroism

spectroscopy is reported to be 4 nM [95]. The k_d of hADAR measured on MST by us was found to be 3.6μM as the protein-ligand complex did not reach saturation. However, this experiment was a preliminary proof to show that hADAR binds to d(GC) and is a positive control for all the binding experiments. The two main advantages of MST over gel shift assays are (a) it is relatively quick and (b) does not require radioactivity. However, Circular Dichorism (CD) spectroscopy will aid in concluding that the RNA substrate that is bound to RBP7910 is indeed in Z-form since the spectra for every form of nucleotides vary. For example, a B-form, A-form and Z-form of DNA will have different spectras as they have different conformations. Since CD-spectroscopy is a very sensitive to conformational changes, even a single mutation in a nucleotide may result in different spectral values validating the Z-form [96]. Therefore CD-spectroscopy may help in concluding the binding interactions between RBP7190 and the potential binder (AUUU) sequences.

Since the glycerol gradients showed co-migration of RBP7910 with the RECC, we wanted to confirm this interaction. Western blotting analysis against the RECC proteins on RBP7910-3v5 pulldowns were insufficient to conclude this as the RECC proteins were present at low concentrations. In order to increase our sensitivity, we performed a radioactive ligase adenylation assay that forms ligase-adenylate intermediate (on KREL1 and KREL2 ligases of the RECC) easily visible on phosphor-imaging. Three samples were used in order to verify this experiment. -Tetracycline (no RBP7910-3v5 induced), +Tetracycline (RBP7910-3v5 overexpressed), +Tetracycline +RNase (RBP7910-3v5 with the effect of RNase). The uninduced sample did not show any adenylation activity as there was no purified protein while RBP7910 without RNase showed adenylation activity, and RBP7910 with RNase treatment had a diminished activity proving that RBP7910 interacts with the core editosome via RNA. The observation that the ligase-adenylation decreases in RNase treated RBP7910-3v5, confirms that the RECC binds to RBP7910 via RNA in low amounts. The KRELs are identified on the basis of their auto-adenylation capability, which is common in ATP-dependent ligases (of the nucleotidyltransferase family) such as T4 RNA ligase 2, DNA ligases, and mRNA-capping enzymes [97-100].

RBP7910 migrated to lighter fractions than RECC proteins on the salt gradients, indicating that RBP7910 is interacting with a complex larger than the RECC. In order to identify this complex, RBP7910-3v5 immunoprecipitations were performed to at physiological conditions to pull down RBP7910-3v5 and its interacting proteins/protein complexes. Since anti-V5 antibody conjugated beads were used for purifying RBP7910-3v5, multiple methods were explored to elute the protein.

Glycine pH 2 was unsuccessfully used in the first elution buffer which degraded the protein. Urea was notably harsh on retaining the interactions with the proteins that RBP7910-3v5 was bound to. SDS buffers had a similar effect as urea and also it split the antibody into the heavy chain and light chain, complicating downstream analysis further. Ultimately, RBP7910-3v5 was eluted using a concentrated solution of v5 peptide that competed out the proteins bound on the beads. This method proved to be gentle on retaining the interacting proteins and consistent. These eluates were sent for in-solution mass spectrometry (MS) analysis.

RBP7910-3v5 eluates were prepared in 3 conditions, -Tetracycline (uninduced), +Tetracycline -RNase, +Tetracycline +RNase. The uninduced sample was used as a negative control, and all the false positives from MS analysis were eliminated. While MS analysis on the +Tetracycline –RNase eluate provided a list of all interacting proteins, MS analysis of +Tetracycline +RNase eluate identified the proteins interacting with RBP7910 via RNA. Through elimination of these proteins we obtained candidate proteins that potentially interacted directly with RBP7910.

The first aim was, Mass spectrometry to analyse all the proteins pulled down by RBP7910. The mass spectrometry data was analysed on a Scaffold viewer, which sorts the proteins according to the peptide count. The identified proteins were then annotated using the tritryp database (https://tritrypdb.org/tritrypdb/). Next, the proteins which did not belong to the mitochondria were eliminated. After the three-step elimination process was done, the proteins which remained were individually investigated. A map was created on Cytoscape with direct interactors and indirect interactors (via RNA). The protein RBP7910 appeared to mainly interact with proteins in the RESC (RNA editing substrate-binding complex) such as GRBC (Guide RNA binding complex), PAMC (Poly adenylating mediator complex), and RGG2 (a protein with an RNA recognition motif). A distinct protein that was a potential direct interactor (protein-protein) with RBP7910 was MERS1, which also belongs to RESC, and this interaction has been published before wherein RBP7910 is referred to as MERS3 [66]. Interestingly the analysis provided us with two novel proteins without any functional annotation: Tb927.10.5830 (RNA-independent) and Tb927.10.4080 (RNA-dependent) are essential proteins in the procyclic and bloodstream stage, and also during differentiation from one stage to another [58]. Since RBP7910 is a potential Z-binding protein, REH2 (RNA helicase 2) may be involved in the transformation from the B-form to the Z-form, which occurs when processive enzymes such as polymerases and helicases generate

underwound DNA in their wake. Z-DNA is stabilized both by negative super-helical stress and by binding of the Zα domain of Z-binding proteins.

Another interesting interacting protein obtained from the MS analysis is TbRGG2, shown below in Fig.13.1. TbRGG2 belongs to the RNA editing mediating complex (REMC) and contains an RNA recognition motif (RRM) or an RNA binding domain (RBD). Typical Z-binding proteins such as Human ADAR have RNA-recognition motif (RRM) or RNA binding domains along with its Z-DNA binding domains [72]. Human ADAR, for instance, has 2 RBD's and since previous data showed that Human ADAR is a homolog of RBP7910, there is a possibility that via direct interaction between TbRGG2, RBP7910 may form a Z-binding form with the help of the RRM/RBD provided *in trans*.

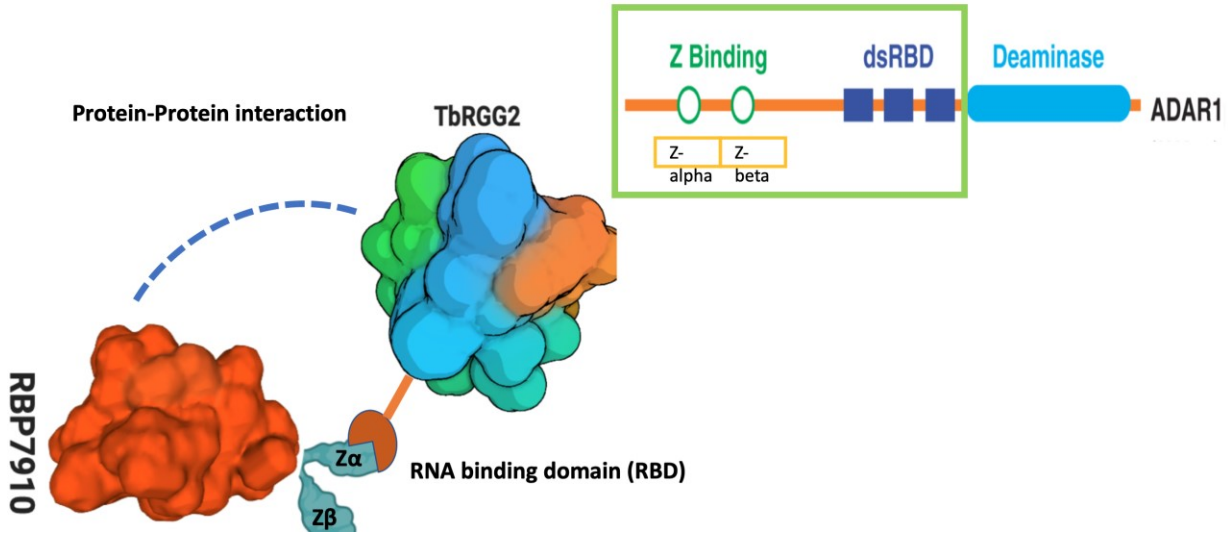


Fig 13.1: RBP7910 replicating its homolog, Human ADAR. *Mass spectrometry results suggest that TbRGG2 and RBP7910 may be direct interactors. Since TbRGG2 has an RRM/RBD, RBP7910 may interact with it to form a Z-binding protein. Human ADAR on the other hand has 2 RBDs and a deaminase connected to the Z-binding sites.*

The other interesting observation from the MS analysis was a protein interacting directly to RBP7910, MERS1, which belongs to the RESC and aids in gRNA stabilization [66]. On the other hand, MERS2 wasn't found in the mass spectrometry of RBP7910 contradicting the published

data. Our MS samples had a control in which the ectopic copy wasn't expressed (-Tetracycline) and a +Tetracycline +RNase which helped in eliminating all false positives and obtained true positives. Since this method wasn't used in the previously published data, the authors may have included the false positives along with the peptide count [66]. In conclusion, MERS1 and TbRGG2 are potential direct interactors of RBP7910, and conducting additional analysis such co-expression, RNA-binding and competition experiments should help us understand RBP7910's function in more detail. Moreover, analyzing the role of the two novel unannotated proteins (Tb927.10.5830 and Tb927.10.4080) may also help understand in vivo regulation of RNA editing.

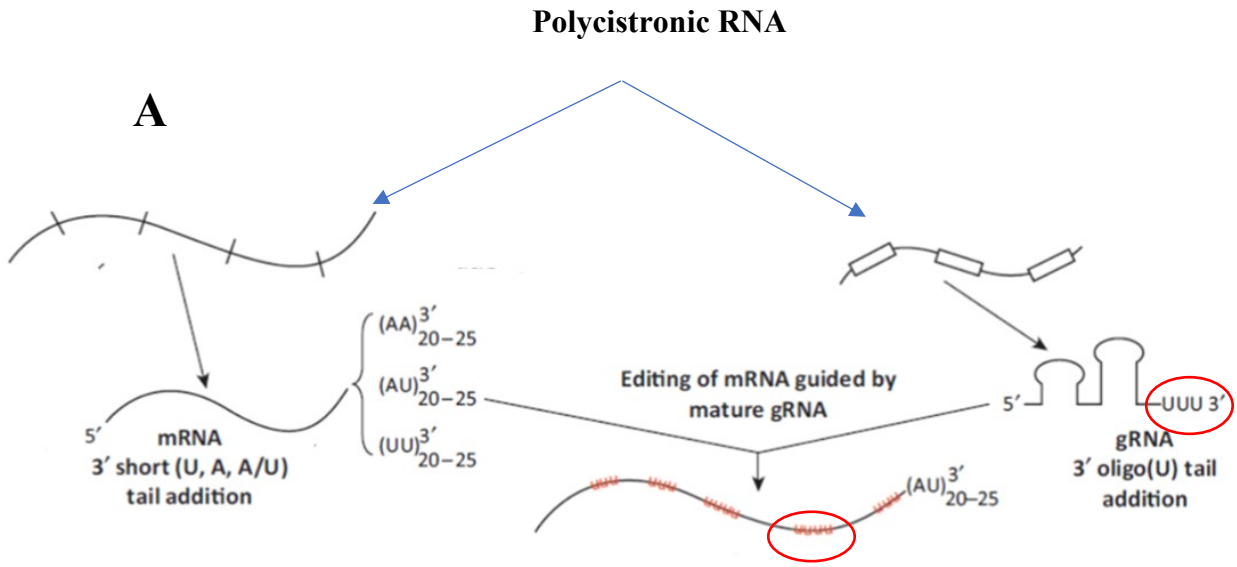
While RBP7910 was shown to be essential for the growth of procyclic stage parasites, we show here that it has a pronounced effect in the bloodstream stage, affecting its survival within 24 hours of RNAi induction, indicating that it may be more essential in this stage. The differentiation property of RBP7910 is quite intriguing and a novel phenomenon, as this has only been shown to occur with RNA binding protein from the cytoplasm previously [101]. RBP7910 on over expression in bloodstream forms leads to procyclic-like cells or short stumpy BF cells. The well-defined mitochondrion is a property of these aforementioned cell types. Mitochondria is less developed in the long slender bloodstream stage cells since these cells rely purely on glycolysis for energy generation. Usually RNA binding proteins carry on the process of differentiation [102]. RBPs have been universally shown to have a combinatorial mode of action, i.e. many RBPs are bound to the same mRNA and some may have opposing roles in RNA metabolism. Additionally, most RBPs have multiple mRNA targets. The widespread, overlapping and often different effects of RBPs on the transcriptome present a great challenge in delineating the function of trypanosome RBPs in regulating gene expression and development. Few of the cytoplasmic RBPs aid in the differentiation activity of the parasite. RBP6 (an RRM containing protein) assists the differentiation from PFs to BFs. RBP6 localizes in the cytoplasm, but RBP7910 maybe the first mitochondrial protein to initiate a differentiation. Since the strain used for transition was not a pleomorphic strain, we couldn't obtain a complete transformation to the procyclic stage cells. Further studies using a pleomorphic strain and qPCR analysis will aid monitoring the differentiation more accurately. MS analysis could be a good data to find out the proteins interacting with RBP7910 in the bloodstream forms and possibly provide with an explanation on its ability to differentiate cells.

14. Major Conclusions

RBP7910 potentially has a direct interaction with TbRGG2 that is a core component of the REMC.

The MS analysis revealed a list of potential protein interactors of RBP7910. Among the list, TbRGG2 was the most intriguing as it contains an RNA recognition motif (RRM), also known as an RNA binding domain (RBD). Other Z-DNA binding proteins such as human ADAR contain an RRM along with the Z-alpha and Z-beta binding domains. This interaction between RBP7910 and TbRGG2 implies that the RRM could be provided *in trans* to RBP7910 for target RNA recognition and function.

Moreover, as the REMC is part of the larger RESC that altogether plays a role in gRNA utilization and editing progression from 3' to 5' region, RBP7910 may play a regulatory role through these complexes. This protein may potentially be involved in an interaction with the U-rich sequences found in gRNA U-tails or in the edited segments of the mRNA (Fig 14.1).



B

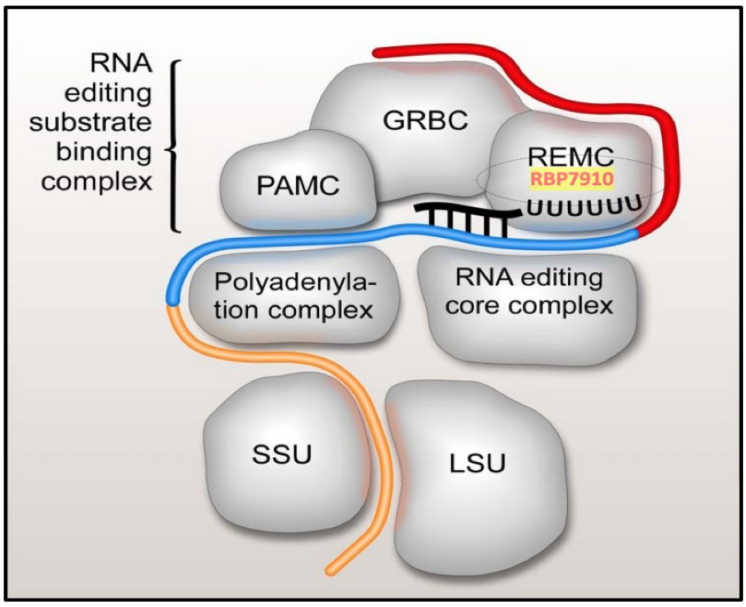


Fig 14.1: Role of RBP7910. *(**A**): Functional role of RBP7910 by interaction with U-rich sequences in gRNA U-tails and U stretches in the edited segments of mRNA [76]. (**B**): RBP7910 a novel component of the REMC. Figure adapted and modified from [33] with permission from American Society for Microbiology.*

RBP7910 is associated with two novel unannotated proteins.

Two novel proteins were identified to be interacting with RBP7910; Tb927.10.5830 which potentially is a direct interactor and Tb927.10.4080 which interacts with RBP7910 via RNA. These proteins are essential for the survival of the parasite in both the life cycle stages and for the differentiation from BF to PF based on a high throughput RNAi study [84]. Further analyses into the roles of these proteins may help us understand RNA editing regulation in greater detail.

Overexpression of RBP7910 in BF likely leads to procyclic-like cells.

Differentiation of the parasite from BF to PF is important in understanding their biology and morphology. The transition from long slender forms to short stumpy forms (precursor procyclics) was possibly initiated by increased expression of RBP7910. Traces of precursor procyclins found on a western blot from RBP7910 over-expressing BF cells and presence of cells containing larger

mitochondrion with developed cristae on a fluorescent microscope indicate the development of short stumpy BF cells. Further analysis into this through qPCR of edited transcripts and differentiation experiments with a pleomorphic cell line will confirm this finding. RBP7910 would be the first reported mitochondrial protein to induce such a change.

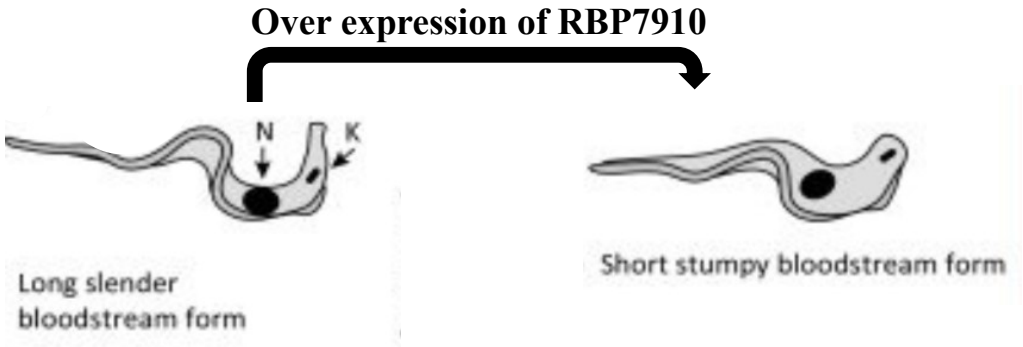


Fig 14.2: Over-expression of RBP7910 in BFs potentially assists differentiation to PFs. *The long slender forms of BFs transition to PFs when RBP7910 is over-expressed in BFs.*

15. Future directions

The future directions of this project would be to analyze the RNA sequences RBP7910 binds to *in vivo* and to prove RBP7910 is a Z-binding protein by extracting the RNA bound to the proteins eluted from pulldowns. The direct interaction of RBP7910 with TbRGG2 could be verified using Yeast 2 hybrid and confirm its presence in the REMC.

For the differentiation experiments, qPCRs could be used to check the gene expressions and to look for procyclin levels. A different strain preferably a polymorphic strain could be used to show the accrual differentiation from one stage to another.

To confirm that RBP7910 is indeed a Z-binding protein, Conditions optimized using thermophoresis should be amenable to CD spectroscopy, which is essential to determine if the target RNA is in Z- form. Since CD spectroscopy is a sensitive process which detects spectras according to conformations as explained before in section 13, it will determine the binding interaction between RBP7910 and its RNA substrate and confirm the conformation of the binding.

16. References

1. Benne, R., et al., *Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA.* Cell, 1986. **46**(6): p. 819-826.
2. Shapiro, T.A., *Kinetoplast DNA maxicircles: networks within networks.* Proceedings of the National Academy of Sciences, 1993. **90**(16): p. 7809.
3. Laurent, M. and M. Steinert, *Electron Microscopy of Kinetoplastic DNA from <em>Trypanosoma mega</em&gt.* Proceedings of the National Academy of Sciences, 1970. **66**(2): p. 419.
4. Schnaufer lab. *Generation of membrane potential in trypanosomes.* 2011.
5. Dong, L.-F., et al., *Horizontal transfer of whole mitochondria restores tumorigenic potential in mitochondrial DNA-deficient cancer cells.* eLife, 2017. **6**: p. e22187.
6. Salavati, R., et al., *Inhibitors of RNA editing as potential chemotherapeutics against trypanosomatid pathogens.* International journal for parasitology. Drugs and drug resistance, 2011. **2**: p. 36-46.
7. Organization, W.H. *Trypanosomiasis, human African (sleeping sickness)*. 2020.
8. Centrer for Diesease Control and Prevention. *Parasites - American Trypanosomiasis (also known as Chagas Disease)*. 2019.
9. Herwaldt, B.L., *Leishmaniasis.* Lancet, 1999. **354**(9185): p. 1191-9.
10. Tripathi, P., V. Singh, and S. Naik, *Immune response to leishmania: paradox rather than paradigm.* FEMS Immunol Med Microbiol, 2007. **51**(2): p. 229-42.
11. Lindoso, J.A., et al., *Leishmaniasis-HIV coinfection: current challenges.* HIV AIDS (Auckl), 2016. **8**: p. 147-156.
12. Centers for Disease Control and Prevention. *Parasites - African Trypanosomiasis (also known as Sleeping Sickness)*.
13. Simpson, R.M., et al., *Trypanosome RNA Editing Mediator Complex proteins have distinct functions in gRNA utilization.* Nucleic acids research, 2017. **45**(13): p. 7965-7983.
14. Zíková, A., et al., *A paradigm shift: The mitoproteomes of procyclic and bloodstream Trypanosoma brucei are comparably complex.* PLoS pathogens, 2017. **13**(12): p. e1006679-e1006679.
15. Vickerman, K., *Polymorphism and Mitochondrial Activity In Sleeping Sickness Trypanosomes.* Nature, 1965. **208**(5012): p. 762-766.
16. Parsons, M., *Glycosomes: parasites and the divergence of peroxisomal purpose.* Mol Microbiol, 2004. **53**(3): p. 717-24.
17. Bochud-Allemann, N. and A. Schneider, *Mitochondrial substrate level phosphorylation is essential for growth of procyclic Trypanosoma brucei.* J Biol Chem, 2002. **277**(36): p. 32849-54.
18. Field, M.C., et al., *Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need.* Nature reviews. Microbiology, 2017. **15**(4): p. 217-231.
19. Ziegelbauer, K., et al., *Synchronous differentiation of Trypanosoma brucei from bloodstream to procyclic forms in vitro.* Eur J Biochem, 1990. **192**(2): p. 373-8.

20. Roditi, I., M. Carrington, and M. Turner, *Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of Trypanosoma brucei.* Nature, 1987. **325**(6101): p. 272-4.
21. Mowatt, M.R. and C.E. Clayton, *Developmental regulation of a novel repetitive protein of Trypanosoma brucei.* Molecular and Cellular Biology, 1987. **7**(8): p. 2838.
22. Mowatt, M.R., G.S. Wisdom, and C.E. Clayton, *Variation of tandem repeats in the developmentally regulated procyclic acidic repetitive proteins of Trypanosoma brucei.* Molecular and Cellular Biology, 1989. **9**(3): p. 1332.
23. Mowatt, M.R. and C.E. Clayton, *Polymorphism in the procyclic acidic repetitive protein gene family of Trypanosoma brucei.* Molecular and Cellular Biology, 1988. **8**(10): p. 4055.
24. Mehlert, A., A. Treumann, and M.A. Ferguson, *Trypanosoma brucei GPEET-PARP is phosphorylated on six out of seven threonine residues.* Mol Biochem Parasitol, 1999. **98**(2): p. 291-6.
25. Vickerman, K., et al., *Biology of African trypanosomes in the tsetse fly.* Biol Cell, 1988. **64**(2): p. 109-19.
26. Vassella, E., et al., *A major surface glycoprotein of trypanosoma brucei is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia.* Genes & development, 2000. **14**(5): p. 615-626.
27. Seed, J.R. and M.A. Wenck, *Role of the long slender to short stumpy transition in the life cycle of the african trypanosomes.* Kinetoplastid biology and disease, 2003. **2**(1): p. 3-3.
28. Zimmermann, H., et al., *A quorum sensing-independent path to stumpy development in Trypanosoma brucei.* PLoS pathogens, 2017. **13**(4): p. e1006324-e1006324.
29. Vickerman, K., *DEVELOPMENTAL CYCLES AND BIOLOGY OF PATHOGENIC TRYPANOSOMES.* British Medical Bulletin, 1985. **41**(2): p. 105-114.
30. Bringaud, F., L. Rivière, and V. Coustou, *Energy metabolism of trypanosomatids: adaptation to available carbon sources.* Mol Biochem Parasitol, 2006. **149**(1): p. 1-9.
31. Matthews, K.R., *The developmental cell biology of Trypanosoma brucei.* Journal of cell science, 2005. **118**(Pt 2): p. 283-290.
32. Read, L.K., J. Lukeš, and H. Hashimi, *Trypanosome RNA editing: the complexity of getting U in and taking U out.* Wiley interdisciplinary reviews. RNA, 2016. **7**(1): p. 33-51.
33. Aphasizheva, I., et al., *RNA binding and core complexes constitute the U-insertion/deletion editosome.* Molecular and cellular biology, 2014. **34**(23): p. 4329-4342.
34. Blum, B., N. Bakalara, and L. Simpson, *A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information.* Cell, 1990. **60**(2): p. 189-98.
35. Lukes, J., H. Hashimi, and A. Zikova, *Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates.* Curr Genet, 2005. **48**(5): p. 277-99.
36. Koslowsky, D.J., et al., *Evidence for U-tail stabilization of gRNA/mRNA interactions in kinetoplastid RNA editing.* RNA Biol, 2004. **1**(1): p. 28-34.
37. Maslov, D.A. and L. Simpson, *The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing.* Cell, 1992. **70**(3): p. 459-67.

38. Li, F., et al., *Trypanosome REH1 is an RNA helicase involved with the 3'-5' polarity of multiple gRNA-guided uridine insertion/deletion RNA editing.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3542-7.
39. Mehta, V., *Dissecting RNA editing mechanism «in vitro» with genetic and chemical genetic approaches*, in *Department of Biochemistry*. 2020, McGill University: McGill University.
40. Horváth, A., et al., *Downregulation of the nuclear-encoded subunits of the complexes III and IV disrupts their respective complexes but not complex I in procyclic Trypanosoma brucei.* Mol Microbiol, 2005. **58**(1): p. 116-30.
41. Schnaufer, A., et al., *The F1-ATP synthase complex in bloodstream stage trypanosomes has an unusual and essential function.* The EMBO journal, 2005. **24**(23): p. 4029-4040.
42. Schnaufer, A., et al., *An RNA Ligase Essential for RNA Editing and Survival of the Bloodstream Form of <em>Trypanosoma brucei</em&gt.* Science, 2001. **291**(5511): p. 2159.
43. Carnes, J., et al., *An essential RNase III insertion editing endonuclease in Trypanosoma brucei.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(46): p. 16614-16619.
44. Hashimi, H., et al., *TbRGG1, an essential protein involved in kinetoplastid RNA metabolism that is associated with a novel multiprotein complex.* RNA, 2008. **14**(5): p. 970-80.
45. Weng, J., et al., *Guide RNA-binding complex from mitochondria of trypanosomatids.* Mol Cell, 2008. **32**(2): p. 198-209.
46. Panigrahi, A.K., et al., *Mitochondrial complexes in Trypanosoma brucei: a novel complex and a unique oxidoreductase complex.* Mol Cell Proteomics, 2008. **7**(3): p. 534-45.
47. Hashimi, H., et al., *Kinetoplastid guide RNA biogenesis is dependent on subunits of the mitochondrial RNA binding complex 1 and mitochondrial RNA polymerase.* RNA, 2009. **15**(4): p. 588-99.
48. Read, L.K., J. Lukes, and H. Hashimi, *Trypanosome RNA editing: the complexity of getting U in and taking U out.* Wiley Interdiscip Rev RNA, 2016. **7**(1): p. 33-51.
49. Ammerman, M.L., et al., *MRB3010 is a core component of the MRB1 complex that facilitates an early step of the kinetoplastid RNA editing process.* RNA, 2011. **17**(5): p. 865-77.
50. Ammerman, M.L., et al., *A core MRB1 complex component is indispensable for RNA editing in insect and human infective stages of Trypanosoma brucei.* PLoS One, 2013. **8**(10): p. e78015.
51. Barraud, P. and F.H.T. Allain, *ADAR proteins: double-stranded RNA and Z-DNA binding domains.* Current topics in microbiology and immunology, 2012. **353**: p. 35-60.
52. Ammerman, M.L., et al., *TbRGG2 facilitates kinetoplastid RNA editing initiation and progression past intrinsic pause sites.* RNA, 2010. **16**(11): p. 2239-51.
53. Fisk, J.C., et al., *TbRGG2, an essential RNA editing accessory factor in two Trypanosoma brucei life cycle stages.* J Biol Chem, 2008. **283**(34): p. 23016-25.
54. Allen, T.E., et al., *Association of guide RNA binding protein gBP21 with active RNA editing complexes in Trypanosoma brucei.* Mol Cell Biol, 1998. **18**(10): p. 6014-22.

55. Osato, D., et al., *Uridine insertion/deletion RNA editing in trypanosomatid mitochondria: In search of the editosome.* RNA, 2009. **15**(7): p. 1338-44.
56. Aphasizhev, R., et al., *A 100-kD complex of two RNA-binding proteins from mitochondria of Leishmania tarentolae catalyzes RNA annealing and interacts with several RNA editing components.* RNA, 2003. **9**(1): p. 62-76.
57. Muller, U.F. and H.U. Goringer, *Mechanism of the gBP21-mediated RNA/RNA annealing reaction: matchmaking and charge reduction.* Nucleic Acids Res, 2002. **30**(2): p. 447-55.
58. Muller, U.F., L. Lambert, and H.U. Goringer, *Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21.* EMBO J, 2001. **20**(6): p. 1394-404.
59. Schumacher, M.A., et al., *Crystal structures of T. brucei MRP1/MRP2 guide-RNA binding complex reveal RNA matchmaking mechanism.* Cell, 2006. **126**(4): p. 701-11.
60. Vondruskova, E., et al., *RNA interference analyses suggest a transcript-specific regulatory role for mitochondrial RNA-binding proteins MRP1 and MRP2 in RNA editing and other RNA processing in Trypanosoma brucei.* J Biol Chem, 2005. **280**(4): p. 2429-38.
61. Fisk, J.C., et al., *Distinct and overlapping functions of MRP1/2 and RBP16 in mitochondrial RNA metabolism.* Mol Cell Biol, 2009. **29**(19): p. 5214-25.
62. Vanhamme, L., et al., *Trypanosoma brucei TBRGG1, a mitochondrial oligo(U)-binding protein that co-localizes with an in vitro RNA editing activity.* J Biol Chem, 1998. **273**(34): p. 21825-33.
63. Sbicego, S., et al., *RBP38, a novel RNA-binding protein from trypanosomatid mitochondria, modulates RNA stability.* Eukaryot Cell, 2003. **2**(3): p. 560-8.
64. Gazestani, V.H., et al., *A Protein Complex Map of Trypanosoma brucei.* PLoS neglected tropical diseases, 2016. **10**(3): p. e0004533-e0004533.
65. Nikpour, N. and R. Salavati, *The RNA binding activity of the first identified trypanosome protein with Z-DNA-binding domains.* Scientific Reports, 2019. **9**(1): p. 5904.
66. Sement, F.M., et al., *Transcription initiation defines kinetoplast RNA boundaries.* Proceedings of the National Academy of Sciences, 2018. **115**(44): p. E10323.
67. Aphasizheva, I., et al., *Lexis and Grammar of Mitochondrial RNA Processing in Trypanosomes.* Trends Parasitol, 2020. **36**(4): p. 337-355.
68. Fisk, J.C., et al., *TbRGG2, an essential RNA editing accessory factor in two Trypanosoma brucei life cycle stages.* The Journal of biological chemistry, 2008. **283**(34): p. 23016-23025.
69. Söding, J., A. Biegert, and A.N. Lupas, *The HHpred interactive server for protein homology detection and structure prediction.* Nucleic Acids Res, 2005. **33**(Web Server issue): p. W244-8.
70. Herbert, A., *Z-DNA and Z-RNA in human disease.* Communications Biology, 2019. **2**(1): p. 7.
71. Savva, Y.A., L.E. Rieder, and R.A. Reenan, *The ADAR protein family.* Genome Biology, 2012. **13**(12): p. 252.
72. Zinshteyn, B. and K. Nishikura, *Adenosine-to-inosine RNA editing.* Wiley interdisciplinary reviews. Systems biology and medicine, 2009. **1**(2): p. 202-209.
73. Brown, L.M., et al., *A cis-acting A-U sequence element induces kinetoplastid U-insertions.* J Biol Chem, 1999. **274**(10): p. 6295-304.

74. Oppegard, L.M., A.L. Kabb, and G.J. Connell, *Activation of guide RNA-directed editing of a cytochrome b mRNA.* J Biol Chem, 2000. **275**(43): p. 33911-9.
75. Aphasizheva, I., et al., *Ribosome-associated pentatricopeptide repeat proteins function as translational activators in mitochondria of trypanosomes.* Mol Microbiol, 2016. **99**(6): p. 1043-58.
76. Hashimi, H., et al., *Dual core processing: MRB1 is an emerging kinetoplast RNA editing complex.* Trends Parasitol, 2013. **29**(2): p. 91-9.
77. Moshiri, H., et al., *Naphthalene-based RNA editing inhibitor blocks RNA editing activities and editosome assembly in Trypanosoma brucei.* J Biol Chem, 2011. **286**(16): p. 14178-89.
78. Moshiri, H., V. Mehta, and R. Salavati, *RNA catalyst as a reporter for screening drugs against RNA editing in trypanosomes.* J Vis Exp, 2014(89).
79. Moshiri, H., et al., *Pilot-scale compound screening against RNA editing identifies trypanocidal agents.* J Biomol Screen, 2015. **20**(1): p. 92-100.
80. Mehta, V., et al., *Mutational analysis of Trypanosoma brucei RNA editing ligase reveals regions critical for interaction with KREPA2.* PloS one, 2015. **10**(3): p. e0120844-e0120844.
81. Wirtz, E., et al., *A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei.* Mol Biochem Parasitol, 1999. **99**(1): p. 89-101.
82. Moon, M.H., et al., *Measuring RNA-Ligand Interactions with Microscale Thermophoresis.* Biochemistry, 2018. **57**(31): p. 4638-4643.
83. Deng, N., et al., *freeQuant: A Mass Spectrometry Label-Free Quantification Software Tool for Complex Proteome Analysis.* TheScientificWorldJournal, 2015. **2015**: p. 137076-137076.
84. Alsford, S., et al., *High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome.* Genome research, 2011. **21**(6): p. 915-924.
85. Milman, N., et al., *Mitochondrial origin-binding protein UMSBP mediates DNA replication and segregation in trypanosomes.* Proceedings of the National Academy of Sciences, 2007. **104**(49): p. 19250.
86. Simpson, L., S. Sbicego, and R. Aphasizhev, *Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business.* RNA (New York, N.Y.), 2003. **9**(3): p. 265-276.
87. Aphasizhev, R. and I. Aphasizheva, *Mitochondrial RNA processing in trypanosomes.* Research in microbiology, 2011. **162**(7): p. 655-663.
88. Grams, J., M.T. McManus, and S.L. Hajduk, *Processing of polycistronic guide RNAs is associated with RNA editing complexes in Trypanosoma brucei.* The EMBO journal, 2000. **19**(20): p. 5525-5532.
89. Zíková, A., et al., *Structure and function of the native and recombinant mitochondrial MRP1/MRP2 complex from Trypanosoma brucei.* Int J Parasitol, 2008. **38**(8-9): p. 901-12.
90. Kafková, L., et al., *Functional characterization of two paralogs that are novel RNA binding proteins influencing mitochondrial transcripts of Trypanosoma brucei.* RNA (New York, N.Y.), 2012. **18**(10): p. 1846-1861.

91. Hernandez, A., et al., *REH2 RNA helicase in kinetoplastid mitochondria: ribonucleoprotein complexes and essential motifs for unwinding and guide RNA (gRNA) binding.* The Journal of biological chemistry, 2010. **285**(2): p. 1220-1228.
92. Ammerman, M.L., et al., *TbRGG2 facilitates kinetoplastid RNA editing initiation and progression past intrinsic pause sites.* RNA (New York, N.Y.), 2010. **16**(11): p. 2239-2251.
93. Kabani, S., M. Waterfall, and K.R. Matthews, *Cell-cycle synchronisation of bloodstream forms of Trypanosoma brucei using Vybrant DyeCycle Violet-based sorting.* Molecular and biochemical parasitology, 2010. **169**(1): p. 59-62.
94. Blattner, J., et al., *Compartmentation of phosphoglycerate kinase in Trypanosoma brucei plays a critical role in parasite energy metabolism.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(20): p. 11596-11600.
95. Herbert, A., et al., *A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase.* Proceedings of the National Academy of Sciences, 1997. **94**(16): p. 8421.
96. Kypr, J., et al., *Circular dichroism and conformational polymorphism of DNA.* Nucleic acids research, 2009. **37**(6): p. 1713-1725.
97. Rusche, L.N., et al., *The two RNA ligases of the Trypanosoma brucei RNA editing complex: cloning the essential band IV gene and identifying the band V gene.* Mol Cell Biol, 2001. **21**(4): p. 979-89.
98. McManus, M.T., et al., *Identification of candidate mitochondrial RNA editing ligases from Trypanosoma brucei.* RNA, 2001. **7**(2): p. 167-75.
99. Cruz-Reyes, J., et al., *Distinct functions of two RNA ligases in active Trypanosoma brucei RNA editing complexes.* Mol Cell Biol, 2002. **22**(13): p. 4652-60.
100. Worthey, E.A., et al., *Comparative analysis of editosome proteins in trypanosomatids.* Nucleic Acids Res, 2003. **31**(22): p. 6392-408.
101. Kolev, N.G., E. Ullu, and C. Tschudi, *The emerging role of RNA-binding proteins in the life cycle of Trypanosoma brucei.* Cellular microbiology, 2014. **16**(4): p. 482-489.
102. Mugo, E. and C. Clayton, *Expression of the RNA-binding protein RBP10 promotes the bloodstream-form differentiation state in Trypanosoma brucei.* PLoS pathogens, 2017. **13**(8): p. e1006560-e1006560.

17. Appendix

Appendix 1- The comprehensive MS dataset is presented in the table below:

Accession Number	Alternate ID	Molecular Weight	(Tetracycline) peptide count	(Tetracycline+ -RNase) Peptide count	(Tetracycline+ +RNase) Peptide count
A0A3L6L4A3_9TRYP (+1)	DPX39_080056400	103 kDa	17	19	16
C9ZWW7_TRYB9	TbgDal_VIII8500	100 kDa	22	17	12
C9ZN76_TRYB9	TbgDal_IV4290	103 kDa	12	20	4
Q38B00_TRYB2	Tb10	22 kDa	8	9	5
C9ZUB9_TRYB9	TbgDal_VII8560	113 kDa	26	6	2
D0A4M0_TRYB9	TbgDal_X13100	34 kDa	5	7	21
D0A311_TRYB9 (+1)	TbgDal_X7440	15 kDa	7	14	3
Q38D54_TRYB2	Tb09	35 kDa	11	11	0
A0A3L6L0B3_9TRYP (+1)	TEF1	49 kDa	4	5	13
C9ZK34_TRYB9 (+1)	TbgDal_III1370	73 kDa	9	9	10
D0A2X8_TRYB9	TbgDal_X7090	60 kDa	7	10	9
A0A3L6KXE5_9TRYP (+2)	DPX39_100086200	21 kDa	2	5	11
A0A3L6LBZ3_9TRYP (+3)	TRRM1	50 kDa	2	13	6
D0A7F7_TRYB9 (+1)	TbgDal_XI7260	49 kDa	5	4	5
A0A3L6LHY5_9TRYP (+1)	DPX39_010020500	75 kDa	3	15	6
A0A3L6KU79_9TRYP (+3)	MERS1	44 kDa	1	4	5
A0A3L6KVM8_9TRYP (+2)	DPX39_110090300	71 kDa	6	2	10
A0A3L6L9F9_9TRYP (+4)	GAPDH	39 kDa	6	9	5
A0A3L6L476_9TRYP (+3)	PABP2	62 kDa	6	5	2
A0A3L6KYI5_9TRYP (+2)	ALD	41 kDa	10	1	3
A0A3L6L6U1_9TRYP (+2)	DPX39_050017000	28 kDa	2	5	0
A0A3L6KYM6_9TRYP (+2)	DPX39_110043000	31 kDa	6	6	3
A0A3L6L1W0_9TRYP	ZC3H40	46 kDa	5	9	5
A0A3L6LGU0_9TRYP (+4)	DPX39_030032300	25 kDa	7	3	1
A0A3L6KVN6_9TRYP (+4)	NRBD2	30 kDa	4	4	1
C9ZQP3_TRYB9	TbgDal_VI2010	83 kDa	2	12	0
A0A3L6L325_9TRYP (+2)	DPX39_100051100	94 kDa	3	6	5
A0A3L6L460_9TRYP (+2)	DPX39_080038300	38 kDa	1	6	7
Q4FKN8_TRYB2	Tb10	561 kDa	3	1	10
A0A3L6KW14_9TRYP (+2)	DPX39_100152700	66 kDa	4	5	4
A0A3L6KU64_9TRYP (+1)	DPX39_110071200	113 kDa	4	4	6

A0A3L6LCP9_9TRYP (+2)	DPX39_020028400	32 kDa	3	3	5
A0A3L6L0C4_9TRYP (+2)	KMP-11	11 kDa	4	3	1
B6SBL9_9TRYP (+2)	GRBC2	55 kDa	0	4	1
A0A3L6L6S4_9TRYP (+2)	MRB800	61 kDa	0	8	6
A0A3L6LC68_9TRYP (+4)	DPX39_040020600	49 kDa	6	5	0
A0A3L6LCU8_9TRYP (+3)	DPX39_010025300	50 kDa	2	3	5
D0A7D9_TRYB9 (+3)	TbgDal_XI7080	100 kDa	5	7	1
A0A3L6L7E7_9TRYP (+2)	DPX39_070034300	52 kDa	1	7	1
A0A3L6LCB0_9TRYP (+2)	REH2	242 kDa	1	8	0
C9ZYM9_TRYB9	TbgDal_IX6040	61 kDa	10	0	0
A0A3L6L6R3_9TRYP (+3)	DPX39_060042200	71 kDa	5	2	3
A0A3L6L9K0_9TRYP (+2)	ALBA3	21 kDa	3	3	2
Q587B0_TRYB2	Tb927	38 kDa	4	4	1
Q8WPS8_9TRYP	H25N7	98 kDa	7	2	1
A0A3L6L139_9TRYP (+2)	KH1	45 kDa	1	6	4
Q38C20_TRYB2	Tb10	66 kDa	6	3	0
A0A3L6KXV1_9TRYP (+1)	DPX39_100125000	35 kDa	0	3	2
A0A3L6KTP0_9TRYP (+2)	DPX39_000028500	121 kDa	0	10	0
A0A3L6KWG7_9TRYP (+2)	DPX39_100044400	21 kDa	4	2	2
A0A3L6LAE0_9TRYP (+2)	DPX39_050036600	58 kDa	0	5	3
A0A3L6L5A5_9TRYP (+2)	DPX39_080013100	31 kDa	3	3	2
C9ZXT5_TRYB9	TbgDal_IX3050	22 kDa	1	2	4
D0A8A1_TRYB9	TbgDal_XI10210	19 kDa	1	3	4
A0A3L6KVP9_9TRYP (+3)	GBP25	25 kDa	4	4	0
A0A3L6KQY7_9TRYP (+2)	DPX39_000068500	35 kDa	1	1	1
A0A3L6L696_9TRYP (+3)	DPX39_070060700	29 kDa	4	5	1
C9ZIW5_TRYB9	TbgDal_II400	97 kDa	0	2	2
A0A3L6KY27_9TRYP (+3)	DPX39_020027700	17 kDa	3	1	1
A0A3L6KY07_9TRYP (+2)	ZC3H34	22 kDa	1	4	1
A0A3L6KYA8_9TRYP (+1)	ZC3H39	32 kDa	3	7	0
C9ZIL9_TRYB9	TbgDal_I1820	41 kDa	0	6	1
A0A3L6L3A8_9TRYP (+1)	DRBD2	34 kDa	0	0	5
A0A3L6KWE1_9TRYP (+4)	DPX39_110026700	35 kDa	3	2	2
A0A3L6KT11_9TRYP (+2)	DPX39_000020600	15 kDa	2	1	1
A0A3L6KUM3_9TRYP (+3)	GCVL-2	50 kDa	1	0	4
A0A3L6KYP1_9TRYP (+2)	HSP60	60 kDa	0	3	5
A0A3L6L8F9_9TRYP (+2)	DPX39_040024700	17 kDa	3	5	0
D0A4T4_TRYB9 (+1)	TbgDal_X13750	21 kDa	2	3	0
A0A3L6KUH9_9TRYP (+5)	DPX39_110088600	21 kDa	6	3	0

A0A3L6L9Q4_9TRYP (+2)	DPX39_030050300	42 kDa	3	3	2
A0A3L6L2J4_9TRYP (+5)	DPX39_090086700	16 kDa	2	4	1
A0A3L6L0I1_9TRYP (+3)	DPX39_090015900	24 kDa	1	3	1
A0A3L6KWV3_9TRYP (+1)	LA	38 kDa	1	5	2
D0A5T7_TRYB9	TbgDal_XI1540	113 kDa	1	0	0
A0A3L6L1B7_9TRYP (+2)	DPX39_090072200	22 kDa	2	3	0
A0A3L6L650_9TRYP (+1)	TSR1	37 kDa	1	4	0
A0A3L6KYT1_9TRYP (+2)	RPS2	29 kDa	1	2	2
A0A3L6KU44_9TRYP (+3)	DPX39_110083800	28 kDa	3	1	1
A0A3L6KR86_9TRYP (+5)	DPX39_000070100	15 kDa	0	2	1
A0A3L6LDI3_9TRYP (+2)	DPX39_010025200	50 kDa	2	1	2
A0A3L6KWM9_9TRYP (+5)	DPX39_100055900	17 kDa	2	1	1
A0A3L6KZP5_9TRYP (+2)	DPX39_100007200	10 kDa	1	3	2
C9ZS66_TRYB9 (+3)	TbgDal_VII1740	28 kDa	3	3	0
A0A3L6L3F6_9TRYP (+4)	DPX39_070062400	18 kDa	3	2	0
A0A3L6KWP7_9TRYP (+2)	DPX39_100108500	61 kDa	0	3	2
A0A3L6KSZ0_9TRYP (+3)	DPX39_000027200	71 kDa	0	4	1
A0A3L6KY76_9TRYP (+2)	DPX39_110052600	13 kDa	1	2	0
A0A3L6L3S3_9TRYP (+2)	RPL10A	25 kDa	1	3	0
C9ZQH9_TRYB9 (+1)	TbgDal_VI1370	62 kDa	0	6	0
D0A3H4_TRYB9 (+1)	TbgDal_X9090	55 kDa	2	2	3
A0A3L6KUK6_9TRYP (+2)	DPX39_110094200	41 kDa	2	2	2
A0A3L6L586_9TRYP (+2)	DPX39_090026000	11 kDa	2	2	2
A0A3L6KRR8_9TRYP (+5)	DPX39_000068700	12 kDa	1	2	1
A0A3L6KR50_9TRYP (+3)	DPX39_000042400	17 kDa	1	1	2
A0A3L6KYA4_9TRYP (+2)	DPX39_100038500	62 kDa	1	3	1
A0A3L6KT35_9TRYP (+2)	DPX39_000056400	19 kDa	4	2	0
A0A3L6L1I3_9TRYP (+2)	RPL27A	16 kDa	3	2	0
A0A3L6KV83_9TRYP (+3)	DPX39_110135200	16 kDa	1	4	0
A0A3L6L209_9TRYP (+3)	QM	25 kDa	4	2	0
A0A3L6KWZ9_9TRYP (+3)	DPX39_110107300	17 kDa	1	0	1
A0A3L6KU69_9TRYP (+3)	CYPA	19 kDa	0	1	4
A0A3L6KZZ3_9TRYP	DPX39_100010600	20 kDa	1	3	0
A0A3L6KVX6_9TRYP (+2)	RACK1	35 kDa	0	1	2
A0A3L6L2D8_9TRYP (+2)	RPS18	18 kDa	2	2	0
A0A3L6L1F2_9TRYP (+1)	DPX39_000061500	22 kDa	2	2	0
A0A3L6KRH7_9TRYP	ALBA2	13 kDa	0	3	0
A0A3L6KTZ3_9TRYP (+2)	DPX39_110145300	53 kDa	0	4	1
A0A3L6L247_9TRYP (+2)	DPX39_100022300	39 kDa	0	1	0

C9ZUJ5_TRYB9 (+1)	TbgDal_VIII360	40 kDa	0	2	0
C9ZSN5_TRYB9 (+1)	TbgDal_VII3590	90 kDa	0	3	0
D0A7Y1_TRYB9 (+2)	TbgDal_XI9010	15 kDa	0	3	0
A0A3L6LAJ4_9TRYP (+3)	DPX39_040057500	18 kDa	1	2	2
C9ZP53_TRYB9	TbgDal_V3190	14 kDa	1	2	1
D0A0C5_TRYB9 (+1)	TbgDal_X17860	38 kDa	2	1	1
A0A3L6KZW1_9TRYP (+2)	RPS23	16 kDa	3	2	0
B3FB38_9TRYP (+1)	kPAP1	58 kDa	0	1	1
A0A3L6L221_9TRYP (+2)	DPX39_090037600	41 kDa	0	1	1
A0A3L6L2X4_9TRYP (+3)	TSR1IP	44 kDa	1	1	0
A0A3L6L4X4_9TRYP (+2)	B5R	32 kDa	0	3	0
A0A3L6KWQ4_9TRYP (+3)	gBP21	23 kDa	1	3	0
C9ZKB8_TRYB9 (+1)	TbgDal_III2200	31 kDa	0	2	0
C9ZPI5_TRYB9 (+1)	TbgDal_V4520	182 kDa	0	0	3
A0A3L6LC75_9TRYP (+3)	KPAF1	114 kDa	0	3	0
Q582X9_TRYB2	Tb927	29 kDa	0	5	0
A0A3L6L280_9TRYP	DPX39_080056300	100 kDa	1	2	1
A0A3L6L3N2_9TRYP (+3)	TRYP2	26 kDa	1	1	2
A0A3L6KZQ4_9TRYP (+2)	DPX39_100082200	48 kDa	2	2	0
A0A3L6LAR1_9TRYP (+2)	DPX39_050016500	54 kDa	0	2	2
A0A3L6KYG7_9TRYP (+1)	ZC3H46	56 kDa	0	1	2
A0A3L6L9C8_9TRYP (+2)	PBP1	60 kDa	2	1	0
A0A3L6L6W2_9TRYP (+2)	DPX39_090066300	15 kDa	2	1	0
A0A3L6KZ39_9TRYP	DPX39_100063700	60 kDa	1	2	0
C9ZQF4_TRYB9	TbgDal_VI1120	27 kDa	2	0	0
A0A3L6L4G0_9TRYP (+4)	RPS3	24 kDa	3	1	0
A0A3L6L7C5_9TRYP (+3)	DPX39_040039000	26 kDa	2	1	0
A0A3L6KRN4_9TRYP (+5)	DPX39_000039100	21 kDa	1	1	0
A0A3L6L1A6_9TRYP (+6)	glk1	56 kDa	1	1	0
M1E1E8_TRYB2		7 kDa	1	0	1
A0A3L6KYD6_9TRYP (+2)	DPX39_110009600	31 kDa	2	0	0
A0A3L6L950_9TRYP (+3)	DPX39_080011000	22 kDa	1	0	0
A0A3L6LBT2_9TRYP (+2)	DPX39_040021200	19 kDa	0	1	0
A0A3L6KYK1_9TRYP (+3)	DPX39_100091600	16 kDa	0	0	3
A0A3L6L8R3_9TRYP (+2)	DPX39_040015700	42 kDa	0	2	0
A0A3L6KWG5_9TRYP (+3)	DPX39_100035300	47 kDa	0	0	2
A0A3L6L7Q4_9TRYP (+2)	RPS14	16 kDa	1	1	1
C9ZUE2_TRYB9 (+1)	TbgDal_VII8750	22 kDa	1	1	1
C9ZXV5_TRYB9 (+10)	TbgDal_IX3250	45 kDa	0	2	1

A0A3L6KVQ5_9TRYP (+2)	ZC3H41	58 kDa	0	0	1
A0A3L6KXW7_9TRYP (+3)	DPX39_100116000	81 kDa	0	1	2
A0A3L6KWA0_9TRYP (+5)	mMDH	33 kDa	0	1	2
A0A3L6LF84_9TRYP (+2)	DPX39_040055900	25 kDa	1	1	0
Q38BK7_TRYB2	Tb10	29 kDa	1	0	1
A0A3L6L5D3_9TRYP (+3)	DPX39_070022200	17 kDa	2	0	0
A0A3L6KTK4_9TRYP (+2)	DPX39_110014600	35 kDa	0	2	0
A0A3L6KWF3_9TRYP	DPX39_000025600	13 kDa	0	0	1
A0A3L6L515_9TRYP (+2)	DPX39_070013700	11 kDa	0	0	3
C9ZXU0_TRYB9 (+2)	TbgDal_IX3100	112 kDa	0	2	0
C9ZW82_TRYB9 (+1)	TbgDal_VIII6090	44 kDa	0	0	2
A0A3L6KSA9_9TRYP (+2)	DPX39_000023100	10 kDa	0	0	2
A0A3L6L431_9TRYP (+2)	HSP10	11 kDa	0	0	2
A0A3L6KXH0_9TRYP (+1)	eIF-5A	18 kDa	0	0	2
A0A3L6KWM3_9TRYP (+3)	DPX39_110053500	102 kDa	0	2	0
A0A3L6L7L0_9TRYP (+2)	DPX39_060016000	53 kDa	0	2	0
A0A3L6L5Z6_9TRYP (+3)	DPX39_070039400	21 kDa	0	0	2
A0A3L6LDA3_9TRYP (+2)	DPX39_030017200	30 kDa	0	1	0
Q389P5_TRYB2 (+1)	Tb10	99 kDa	0	1	1
A0A3L6LC39_9TRYP (+2)	DPX39_040007100	22 kDa	1	1	0
A0A3L6L5N4_9TRYP	DPX39_100097400	5 kDa	1	1	0
D0A745_TRYB9 (+1)	TbgDal_XI6140	62 kDa	0	1	0
A0A3L6KR76_9TRYP (+5)	DPX39_000068600	24 kDa	0	0	1
A0A3L6L810_9TRYP (+1)	XRNA	159 kDa	0	1	0
A0A3L6L5Q4_9TRYP (+3)	DPX39_060027100	25 kDa	0	0	1
A0A3L6KZM1_9TRYP (+4)	DPX39_100006900	28 kDa	0	0	1
A0A3L6KWC1_9TRYP (+2)	TIM	27 kDa	0	0	2
A0A3L6L0M5_9TRYP (+2)	DRBD3	37 kDa	2	0	0
A0A3L6L7R9_9TRYP (+1)	RBP17	17 kDa	2	0	0
A0A3L6LB36_9TRYP (+2)	DPX39_030044700	62 kDa	0	0	2
C9ZQF0_TRYB9	TbgDal_VI1080	35 kDa	0	2	0
C9ZR33_TRYB9 (+1)	TbgDal_VI3410	20 kDa	0	0	2
A0A3L6L5I8_9TRYP (+2)	DPX39_060024500	29 kDa	0	2	0
A0A3L6LBE4_9TRYP (+2)	DPX39_030055400	25 kDa	0	2	0
A0A3L6LAN7_9TRYP (+4)	DPX39_030042300	69 kDa	2	0	0
A0A3L6LB24_9TRYP (+2)	DPX39_030016900	38 kDa	2	0	0
A0A3L6L0K2_9TRYP (+2)	DPX39_090058100	19 kDa	0	2	0
A0A3L6KWA9_9TRYP (+2)	DPX39_110071700	8 kDa	1	0	0
A0A3L6KTN9_9TRYP (+2)	DPX39_000056500	19 kDa	0	1	0

A0A3L6KWD8_9TRYP (+3)	RPL44	12 kDa	0	1	0
C9ZNN3_TRYB9	TbgDal_V1490	91 kDa	0	1	0
A0A3L6L5H9_9TRYP (+2)	DPX39_060025900	16 kDa	0	0	1
A0A3L6KXK6_9TRYP (+2)	DPX39_100133400	43 kDa	1	0	0
A0A3L6L9A7_9TRYP (+2)	DPX39_060025200	124 kDa	0	1	0
A0A3L6KWJ8_9TRYP (+5)	DPX39_110011000	18 kDa	1	0	0
A0A3L6L3P6_9TRYP (+2)	DPX39_080048500	19 kDa	0	1	0
A0A3L6KWY5_9TRYP (+2)	DPX39_100155100	43 kDa	0	0	1
C9ZN03_TRYB9 (+1)	TbgDal_IV3570	37 kDa	0	0	1
C9ZYE6_TRYB9 (+1)	TbgDal_IX5210	63 kDa	0	1	0
A0A3L6L442_9TRYP (+4)	DPX39_070036800	14 kDa	1	0	0
A0A3L6L8K2_9TRYP (+4)	DPX39_040080700	45 kDa	0	0	1
A0A3L6KUU3_9TRYP	DPX39_110107000	15 kDa	1	0	0
Q4FKH6_TRYB2	Tb11	105 kDa	0	1	0
A0A3L6L0N4_9TRYP (+2)	DPX39_100066000	19 kDa	0	0	1
C9ZZ43_TRYB9 (+1)	TbgDal_IX7680	83 kDa	0	0	1
A0A3L6KY55_9TRYP (+5)	DPX39_100155400	34 kDa	1	0	0
A0A3L6L452_9TRYP (+1)	DPX39_070055100	16 kDa	0	0	1
A0A3L6LAF8_9TRYP (+1)	DPX39_030025700	29 kDa	0	1	0
C9ZXX7_TRYB9	TbgDal_IX3470	71 kDa	0	1	0
D0A516_TRYB9 (+1)	TbgDal_X14580	13 kDa	0	1	0
Q8WPT5_9TRYP	H25N7	94 kDa	0	1	0
A0A3L6L9Y8_9TRYP (+2)	EIF3D	31 kDa	0	1	0
A0A3L6KV64_9TRYP (+2)	DPX39_090050300	10 kDa	1	0	0
A0A3L6L8S6_9TRYP	DPX39_040072400	103 kDa	1	0	0
A0A3L6KRQ9_9TRYP (+2)	DPX39_000057800	31 kDa	0	0	1
A0A3L6KTV9_9TRYP (+2)	DPX39_110031300	61 kDa	0	0	1
A0A3L6KUF2_9TRYP (+2)	DPX39_110047600	78 kDa	0	0	1
A0A3L6KW16_9TRYP (+3)	DPX39_110068500	10 kDa	0	0	1
A0A3L6KXA9_9TRYP (+2)	DPX39_110095600	16 kDa	0	0	1
A0A3L6L4H7_9TRYP (+5)	TDPX3	20 kDa	0	0	1
A0A3L6L5H1_9TRYP (+2)	DPX39_060021800	47 kDa	0	0	1
A0A3L6LB04_9TRYP (+2)	DPX39_030052700	42 kDa	0	0	1
A0A3L6LBT9_9TRYP (+3)	DPX39_020018600	41 kDa	0	0	1
A0A3L6LFR2_9TRYP (+6)	DPX39_010011000	47 kDa	0	0	1
C9ZIV0_TRYB9 (+2)	TbgDal_II170	21 kDa	0	0	1
D0A2T5_TRYB9 (+2)	TbgDal_X6650	15 kDa	0	0	1
D0A3S0_TRYB9 (+1)	TbgDal_X10050	56 kDa	0	0	1
D0AAA9_TRYB9 (+1)	TbgDal_XI17300	23 kDa	0	0	1

Q57WL2_TRYB2	Tb927	103 kDa	0	0	1
A0A220QKC6_TRYBB (+3)	KPAF3	117 kDa	0	1	0
A0A3L6KTM7_9TRYP (+2)	MRPL4	53 kDa	0	1	0
A0A3L6KUT1_9TRYP (+2)	DPX39_110124100	159 kDa	0	1	0
A0A3L6KY22_9TRYP (+2)	DPX39_100041300	42 kDa	0	1	0
A0A3L6KYM2_9TRYP (+2)	DPX39_100046600	158 kDa	0	1	0
A0A3L6KZ45_9TRYP (+2)	DPX39_100159700	44 kDa	0	1	0
A0A3L6KZT6_9TRYP (+2)	RBP29	41 kDa	0	1	0
A0A3L6L1E7_9TRYP (+2)	DPX39_100144200	46 kDa	0	1	0
A0A3L6L1L5_9TRYP (+2)	RBSR1	18 kDa	0	1	0
A0A3L6L2K0_9TRYP (+1)	eIF4G3	69 kDa	0	1	0
A0A3L6L2T2_9TRYP (+2)	DPX39_080028400	52 kDa	0	1	0
A0A3L6L427_9TRYP (+2)	DPX39_070071500	88 kDa	0	1	0
A0A3L6L8V9_9TRYP (+2)	DPX39_040033600	18 kDa	0	1	0
A0A3L6LCE5_9TRYP (+2)	DPX39_040076300	31 kDa	0	1	0
C9ZID5_TRYB9 (+1)	TbgDal_I1130	75 kDa	0	1	0
C9ZMU8_TRYB9 (+1)	TbgDal_IV3020	159 kDa	0	1	0
C9ZRQ6_TRYB9 (+1)	TbgDal_VII220	20 kDa	0	1	0
C9ZY15_TRYB9 (+1)	TbgDal_IX3850	54 kDa	0	1	0
A0A3L6KZX4_9TRYP (+2)	DPX39_100135100	81 kDa	0	1	0
A0A3L6KU18_9TRYP (+3)	PEX11	24 kDa	1	0	0
A0A3L6KY86_9TRYP (+2)	DPX39_100016800	58 kDa	1	0	0
A0A3L6L0V0_9TRYP (+2)	DPX39_090039400	22 kDa	1	0	0
A0A3L6L8L6_9TRYP (+2)	DPX39_050026800	52 kDa	1	0	0
D0A0M1_TRYB9 (+2)	TbgDal_X18820	34 kDa	1	0	0
Q387Q3_TRYB2	Tb10	48 kDa	1	0	0