Regulation of the Variant Surface Glycoprotein (VSG) Expression and Characterisation of the Nucleolar DExD/H box Protein Hel66 in *Trypanosoma brucei*

Regulation der Expression des variable Oberflächen-Glykoprotein (VSG) und Charakterisierung des nukleolären DExD/H box Protein Hel66 in *Trypanosoma brucei*

Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,

Section: Infection and Immunity

submitted by

Majeed Bakari Soale

from

Yapei, Ghana

Würzburg, 2021

Submitted on:

Members of the Thesis Committee

Chairperson: Prof. Dr. Keram Pfeiffer

Primary Supervisor: Prof. Dr. Markus Engstler

Supervisor (Second): Prof. Dr. Klaus Brehm

Supervisor (Third): Prof. Dr. Mark Carrington

Date of Public Defence:……..……

Date of Receipt of Certificates:……

AFFIDAVIT

I hereby confirm that my thesis entitled "Regulation of the Variant Surface Glycoprotein (VSG) Expression and Characterisation of the Nucleolar DExD/H box Protein Hel66 in *Trypanosoma brucei*" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Place, Date Signature

EIDESSTATTLICHE ERKLÄRUNG

Hiermit erkläre ich an Eides statt, die Dissertation „ Regulation der Expression des variable Oberflächen-Glykoprotein (VSG) und Charakterisierung des nukleolären DExD/H box Protein Hel66 in *Trypanosoma brucei*" eigenständig, d.h. insbesondere selbstständig und ohne Hilfsmittel eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ort, Datum Unterschrift

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SUMMARY

The variant surface glycoprotein (VSG) of African trypanosomes plays an essential role in protecting the parasites from host immune factors. These trypanosomes undergo antigenic variation resulting in the expression of a single VSG isoform out of a repertoire of around 2000 genes. The molecular mechanism central to the expression and regulation of the VSG is however not fully understood.

Gene expression in trypanosomes is unusual due to the absence of typical RNA polymerase II promoters and the polycistronic transcription of genes. The regulation of gene expression is therefore mainly post-transcriptional. Regulatory sequences, mostly present in the 3´ UTRs, often serve as key elements in the modulation of the levels of individual mRNAs. In *T. brucei* VSG genes, a 100 % conserved 16mer motif within the 3´ UTR has been shown to modulate the stability of *VSG* transcripts and hence their expression. As a stability-associated sequence element, the absence of nucleotide substitutions in the motif is however unusual. It was therefore hypothesised that the motif is involved in other essential roles/processes besides stability of the *VSG* transcripts.

In this study, it was demonstrated that the 100 % conservation of the 16mer motif is not essential for cell viability or for the maintenance of functional VSG protein levels. It was further shown that the intact motif in the active VSG 3´ UTR is neither required to promote VSG silencing during switching nor is it needed during differentiation from bloodstream forms to procyclic forms. Crosstalk between the VSG and procyclin genes during differentiation to the insect vector stage is also unaffected in cells with a mutated 16mer motif. Ectopic overexpression of a second VSG however requires the intact motif to trigger silencing and exchange of the active VSG, suggesting a role for the motif in transcriptional VSG switching. The 16mer motif therefore plays a dual role in VSG *in situ* switching and stability of *VSG* transcripts. The additional role of the 16mer in the essential process of antigenic variation appears to be the driving force for the 100 % conservation of this RNA motif.

A screen aimed at identifying candidate RNA-binding proteins interacting with the 16mer motif, led to the identification of a DExD/H box protein, Hel66. Although the protein did not appear to have a direct link to the 16mer regulation of VSG expression, the DExD/H family of proteins are important players in the process of ribosome biogenesis. This process is relatively understudied in trypanosomes and so this candidate was singled out for detailed characterisation, given that the 16mer story had reached a natural end point. Ribosome biogenesis is a major cellular process in eukaryotes involving ribosomal RNA, ribosomal

proteins and several non-ribosomal *trans*-acting protein factors. The DExD/H box proteins are the most important *trans*-acting protein factors involved in the biosynthesis of ribosomes. Several DExD/H box proteins have been directly implicated in this process in yeast. In trypanosomes, very few of this family of proteins have been characterised and therefore little is known about the specific roles they play in RNA metabolism. Here, it was shown that Hel66 is involved in rRNA processing during ribosome biogenesis. Hel66 localises to the nucleolus and depleting the protein led to a severe growth defect. Loss of the protein also resulted in a reduced rate of global translation and accumulation of rRNA processing intermediates of both the small and large ribosomal subunits. Hel66 is therefore an essential nucleolar DExD/H protein involved in rRNA processing during ribosome biogenesis. As very few protein factors involved in the processing of rRNAs have been described in trypanosomes, this finding represents an important platform for future investigation of this topic.

ZUSAMMENFASSUNG

Das variable Oberflächen-Glykoprotein ("varaint surface glycoprotein", VSG) der Afrikanischen Trypanosomen schützt den Parasiten vor Immunfaktoren des Wirtes. Trypanosomen beherrschen die antigene Variation und expremieren nur eine einzige VSG Isoform aus einem Repertoire von ungefähr 2000 Genen. Der molekulare Mechanismus der die Expression dieser VSG Gene reguliert ist nicht komplett bekannt.

Die Genexpression ist in Trypanosomen sehr ungewöhnlich. Es gibt keine typischen Promotoren für RNA Polymerase II und Gene werden polycistronisch transkribiert. Daher ist die Regulation der Genexpression hauptsächlich posttranskriptional. Die Expression individueller mRNAs wird durch regulatorische Sequenzen reguliert, die sich häufig in den 3´ UTRs befinden. In den VSG Genen von *T. brucei* moduliert ein zu 100% konserviertes 16mer Motiv in der 3´ UTR die Stabilität der *VSG* Transkripte und damit deren Expression. Für eine Sequenz, die die Stabilität der mRNA reguliert, ist das Fehlen von Nukleotid Substitutionen sehr ungewöhnlich. Es wurde deshalb spekuliert, dass das 16mer Motiv neben der Stabilisierung des VSG Transkriptes noch an anderen essentiellen Prozessen beteiligt ist.

In dieser Arbeit wurde gezeigt, dass die 100%ige Konservierung des 16mer Motives weder für das Überleben der Zellen, noch für den Erhalt der Expression des VSG Protein in funktioneller Menge notwendig ist. Außerdem wurde gezeigt dass das intakte Motiv in der 3´UTR des aktiven VSGs weder für das „VSG silencing" während des VSG Austausches („switching") noch für die Differenzierung von Blutbahnformen zu prozyklischen Formen benötigt wird. Auch die Interaktionen („crosstalk"), die während der Differenzierung zum Insekten Stadium zwischen den VSG und Prozyklin Genen stattfinden, sind in Zellen mit mutiertem 16mer Motiv noch funktionell. Die ektopische Überexpression eines zweiten VSGs benötigt allerdings das intakte Motiv, um das aktive VSG zu inaktivieren und auszutauschen: dies suggeriert eine Rolle des Motivs im transkriptionalen „VSG switching". Das 16mer Motif spielt daher eine Doppelrolle bei der Regulation der Stabilität der *VSG* Transkripte und im VSG *in situ* „switching". Letzteres, die Rolle im essentiellen Prozess der antigenen Variation, ist dabei offensichtlich die treibende Kraft hinter der 100%igen Konservierung des RNA Motives.

Eine Suche nach möglichen RNA bindenden Proteinen, die mit dem 16mer interagieren, führte zur Identifikation des DExD/H box Proteins Hel66. Obwohl das Protein wohl nicht direkt an der Regulation der VSG Expression über das 16mer beteiligt ist, spielen Mitglieder der DexD/H Proteinfamilie eine wichtige Rolle in der Biogenese von Ribosomen. Dieser Prozess ist in Trypanosomen noch nicht komplett verstanden und daher wurde das Protein für eine nähere

Analyse ausgewählt, auch weil die 16mer Story ohne weitere Kandidaten zu einem Ende gekommen war. Die Biogenese von Ribosomen ist ein wichtiger zellulärer Prozess in Eukaryoten und benötigt ribosomale RNA, ribosomale Proteine sowie einige nicht-ribosomale, *trans*-agierende Protein Faktoren. Proteine der DExD/H box Familie sind die wichtigsten trans-agierenden Proteinfaktoren, die an der Biogenese der Ribosomen beteiligt sind. In der Hefe sind mehrere DExD/H box Proteine bekannt, die eine direkte Rolle in diesem Prozess spielen. In Trypanosomen sind erst sehr wenige Proteine aus dieser Familie untersucht worden und es ist daher kaum bekannt, welche spezifische Rollen sie im RNA Metabolismus spielen. In dieser Arbeit wurde gezeigt, dass Hel66 an der rRNA Prozessierung während der Biogenese der Ribosomen beteiligt ist. Hel66 ist im Nukleolus lokalisiert und die Reduktion des Proteins durch RNAi führte zu einem schweren Wachstumsphänotyp. Reduktion von Hel66 führte auch zu einer globalen Reduktion der Translation sowie zur Akkumulation von Synthese-Zwischenstadien der rRNAs sowohl der kleinen und als auch der großen ribosomalen Untereinheit. Hel66 ist daher ein essentielles nukleoläres DExD/H Protein dass an der Prozessierung der rRNA während der Biogenese der Ribosomen beteiligt ist. Da bisher erst wenige Proteine bekannt sind, die in Trypanosomen an diesem Prozess beteiligt sind, sind diese Ergebnisse ein sehr wichtiger Ausgangspunkt für weitere Untersuchungen in der Zukunft.

1.0 INTRODUCTION

1.1 African trypanosomes

African trypanosomes are extracellular, tsetse-transmitted, single-celled protist parasites predominantly distributed in sub-Saharan Africa. They are responsible for the disease Human African Trypanosomiasis (HAT), also known as sleeping sickness, in humans and a disease complex in animals termed Animal African Trypanosomiasis (AAT). Trypanosomes are classified in the order Kinetoplastida, family Trypanosomatidae and genus *Trypanosoma*. African trypanosomes include an array of different species of parasites but only three species (namely; *Trypanosoma brucei, Trypanosoma congolense* and *Trypanosoma vivax*) account for the vast majority of disease in humans and animals (Radwanska et al., 2018). *Trypanosoma brucei* is the most-studied African trypanosome with the subspecies *Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense* being the causative agents of HAT. *T. brucei brucei, T. b. evansi, T. congolense*, *T. vivax* and several other animal infective species are responsible for AAT. Apart from African trypanosomes being medically important parasites, they present a unique architecture and features (Figure 1) that make them excellent models for studies in cell biology. The flagellum, flagellar pocket, mitochondrion, kinetoplast and nucleus are single-copy and have precise positioning within the cytoskeletal corset, and are concentrated between the posterior and centre of the cell (Matthews, 2005).

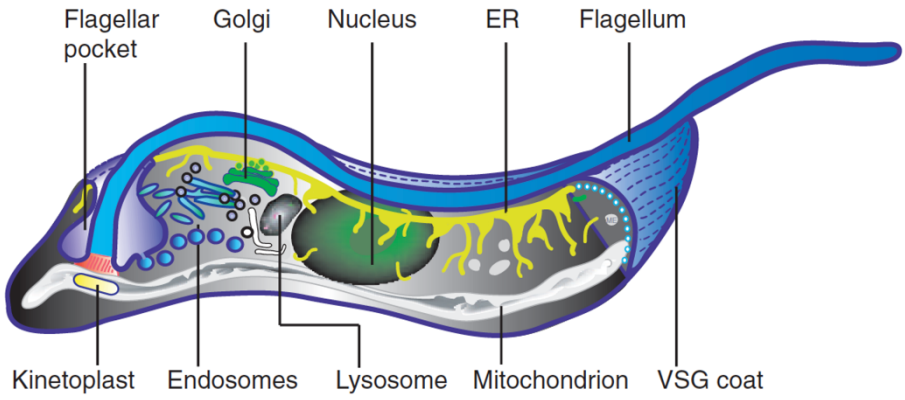

Figure 1. The cellular architecture of Trypanosoma brucei. Diagram of the ultrastructure of *T. brucei* showing an overview of all the major organelles and their location (adapted from Overath & Engstler, 2004).

1.1.1 Human African Trypanosomiasis (HAT)

Human African Trypanosomiasis is a potentially fatal neglected tropical disease caused by two subspecies of *Trypanosoma brucei* in sub-Saharan Africa. An estimated 70 million people are at risk of contracting the disease but in recent years there has been drastic reduction in reported

cases with only 977 cases reported in 2018 (Simarro et al., 2012; WHO, 2020). The disease manifests in two forms: a chronic disease prevalent in western and central Africa caused by *T. b. gambiense* and an acute disease prevalent in eastern and southern Africa caused by *T. b. rhodesiense* (Simarro et al., 2010). The disease distribution is restricted to regions where the insect vector (tsetse fly) is present. Humans are naturally resistant to infection by most African trypanosomes except *T. b. gambiense* and *T. b. rhodesiense*. Trypanolytic factor (TLF) 1 and TLF2 have been identified as the factors responsible for the innate immunity against *Trypanosoma brucei brucei* and other animal infective African trypanosomes (Hajduk et al., 1989; Raper et al., 1999). Despite differences in the composition of these two serum complexes, both TLFs contain the haptoglobin-related protein (HPR) and the trypanolytic toxin apolipoprotein L1 (APOL1). Uptake of TLF1 and 2 results in the interaction of the APOL1 component with the lysosome membrane, causing anionic pore formation that triggers cell lysis (Molina-Portela et al., 2008; Pérez-Morga et al., 2005; Vanhollebeke et al., 2007). The ability of *T. b. gambiense* and *T. b. rhodesiense* to infect humans is due to an acquired resistance to TLF1 and 2 in human serum (Pays & Vanhollebeke, 2009). *T. b. rhodesiense* expresses a protein called serum resistance-associated protein (SRA) which binds to APOL1 and prevents its membrane insertion, thus hindering the anionic pore formation and cell lysis (Xong et al., 1998; Vanhamme et al., 2003). *T. b. gambiense* expresses a *T. b. gambiense*-specific glycoprotein (TbgGP) that prevents APOL1 interaction with the lysosomal membrane, due to its ability to induce membrane stiffening upon interaction with lipids. Inactivation of the haptoglobin-haemoglobin receptor due to a L210S substitution also reduces uptake of APOL1 in this subspecies (Uzureau et al., 2013).

Clinical presentation of HAT is dependent on the parasite subspecies, the host factors and the stage of the disease. The disease generally progresses in two stages: a haemolymphatic stage (first stage), where the parasite is restricted to the blood and lymphatic system and a meningoencephalitic stage (second stage), where the parasite crosses the blood-brain barrier into the central nervous system (Büscher et al., 2017). The haemolymphatic stage is characterised by non-specific symptoms such as intermittent fevers, headaches, pruritus (itching of the skin), lymphadenopathy, hepatomegaly, skin rashes and signs of immune system mobilisation. The meningoencephalitic stage on the other hand involves neuropsychiatric symptoms including sleep disorder, tremors, antisocial or aggressive behaviour and episodes of confusion or delirium (Bouteille & Buguet, 2012; Büscher et al., 2017; Kennedy, 2008). The disease progression is acute in *rhodesiense* HAT with the second stage of the disease involving the central nervous system (CNS) occurring a few weeks after initial infection. In *gambiense*

HAT, infection is chronic and usually slow, with delayed CNS infection lasting months to years. In both cases however, untreated disease almost invariably results in death (Kennedy, 2004).

Treatment of HAT depends on the disease form and the stage of the disease. First stage *gambiense* and *rhodesiense* HAT is treated with pentamidine and suramin, respectively. For treatment of second stage disease, Nifurtimox-Eflornithine combination therapy (NECT) is used for *gambiense* HAT and melarsoprol used for *rhodesiense* HAT (Büscher et al., 2017; Priotto et al., 2009; WHO, 2020). Fexinidazole, an oral monotherapy has recently been recommended for treatment of first and non-severe second stage *gambiense* HAT (Lindner et al., 2019; WHO, 2020). This is a huge step in the treatment of AAT as fexinidazole is much easier to administer compared to the earlier medications.

1.1.2 Animal African Trypanosomiasis (AAT)

Animal African Trypanosomiasis (AAT) is one of the most devastating parasitic diseases of vertebrate animals. The disease affects both livestock and wildlife. AAT poses a major constraint to livestock production in sub-Saharan Africa causing huge economic losses (Shaw et al., 2014). There are several species of African trypanosomes responsible for AAT. The major pathogenic species responsible for most animal infections are *T. congolense, T. vivax* and *T. b. brucei*. Other trypanosomes that cause AAT include; *Trypanosoma simiae, Trypanosoma suis, Trypanosoma godfreyi, Trypanosoma brucei evansi* and *Trypanosoma equiperdum* (Finelle, 1973; Nantulya, 1990). Some trypanosomes such as *Trypanosoma congolense* and *Trypanosoma brucei brucei* are mostly transmitted cyclically by tsetse flies and are therefore present only within the tsetse fly belt. *T. b. evansi* and *T. vivax* (also tsetse-transmitted), which are transmitted mechanically by biting flies and *T. equiperdum,* which is transmitted sexually, can cause AAT outside the tsetse fly belt. Clinical manifestations of AAT are variable and are influenced by the species or strain of trypanosome and the host. Some animals may be infected with or without clinical signs. Most clinical cases are however chronic, although acute cases which often lead to death are also observed. The disease is generally characterised by intermittent fever, severe anaemia, weight loss, abortion, lymphadenopathy, reduced productivity and infertility (Desquesnes et al., 2013; Nantulya, 1990).

Vector control of animal trypanosomiasis is expensive and ineffective in some regions as some parasite species are transmitted mechanically by biting flies. Chemotherapy and chemoprophylaxis are therefore the principal approach in control of AAT (Grace et al., 2009). Six compounds namely; diminazene aceturate, homidium bromide, isometamedium chloride, quinapyramine sulphate, suramin sodium and melarsomine dihydrochloride are currently used

in the treatment and prophylaxis of pathogenic animal trypanosome infections (Giordani et al., 2016). Some of these drugs are toxic and resistance has been reported in some regions, making the control of AAT a major challenge in endemic regions (Dagnachew et al., 2015; Giordani et al., 2016; Knoppe et al., 2006).

1.2 Life cycle of *T. brucei*

The life cycle of *T. brucei* alternates between a mammalian host and an insect vector (tsetse fly). During the life cycle, the parasite undergoes multiple changes in its biology to adapt to the changing environmental conditions it encounters (reviewed in Matthews, 2005) (Figure 2). The parasite is transmitted to the mammalian host when an infected tsetse fly injects the non-dividing metacyclic form trypanosomes into the skin tissue during a blood meal. Following inoculation, the metacyclic parasites rapidly differentiate into bloodstream form (BSF) trypomastigotes. This transformation is characterised by re-initiation of cell division, change in morphology, and a change in the composition of the surface glycoprotein coat, the variant surface glycoprotein (VSG) (Graham et al., 1998). These dividing cells enter the lymphatic system and are disseminated to the bloodstream and other extracellular spaces. BSF trypomastigotes exhibit two main developmental stages — proliferating long slender (LS) forms and the cell cycle-arrested short stumpy (SS) form. The proliferating LS forms multiply in the bloodstream resulting in elevated parasitaemia. At high population densities, the accumulation of a quorum sensing factor SIF (stumpy induction factor) triggers differentiation of the LS forms into cell cycle-arrested SS forms (Reuner et al., 1997). The nature of SIF and the mechanism of signalling has remained elusive although components and molecular regulators of the SIF response pathway have been identified (Mony et al., 2014; Mony & Matthews, 2015). A recent study established that SIF is an oligopeptide signal generated by parasite-secreted peptidases and transported via a trypanosome G-protein-coupled receptor protein, TbGPR89 (Rojas et al., 2019). Transition to SS forms is accompanied by molecular changes including the upregulation of *expression site associated gene 9* (ESAG9), downregulation of the haptoglobin-haemoglobin receptor and expression of a stumpy-specific cell marker *protein associated with differentiation* (PAD1) (Barnwell et al., 2010; Dean et al., 2009; Vanhollebeke et al., 2010). It has also been reported that transition from LS to SS forms can occur via a SIF-independent pathway controlled by the transcriptional status of the VSG expression site (Batram et al., 2014; Zimmermann et al., 2017). The formation of SS forms is thought to be partly an adaptation that limits parasite burden and increases host survival, as these cells have a lifespan of only 48 – 72 hours (Seed & Wenck, 2003; Turner et al., 1995). The SS forms were also believed to be the only life cycle stage capable of infecting tsetse flies

(Rico et al., 2013) but Schuster et al. (2021) have recently shown that proliferating LS cells are just as successful as SS cells in infecting flies and completing the life cycle in tsetse flies.

The life cycle in the tsetse fly begins following ingestion of bloodstream form parasites (LS or SS) by the tsetse fly during a blood meal on an infected host. The process takes about 3 weeks to complete for *T. brucei* parasites (Rotureau & Van Den Abbeele, 2013). Once the bloodstream trypanosomes reach the fly midgut, they differentiate to proliferating procyclic trypanosomes. In these procyclic cells, the VSG is completely replaced by a surface coat composed of an array of different proteins called procyclins (Mowatt & Clayton, 1987; Richardson et al., 1988; Roditi et al., 1987). The procyclins are proteins with repeating units, predicted to contain 22 – 30 EP repeats (EP procyclin) or six GPEET repeats followed by three EP repeats (GPEET procyclin) (Acosta-Serrano et al., 1999). Both EP and GPEET procyclins are expressed by early procyclics. Late procyclics, which are characteristic of established midgut infections, however express only EP procyclins (Acosta-Serrano et al., 2001; Vassella et al., 2000). Repression of VSG synthesis and induction of procyclin expression occurs almost immediately, resulting in the replacement of the surface coat within 12 h and entry into the first cell cycle (Gruszynski et al., 2006). Procyclic cells also undergo metabolic reprogramming to adapt to the glucose-poor and amino acid-rich environment inside the fly (Smith et al., 2017). The differentiation from bloodstream to procyclic trypanosomes can also be induced *in vitro* in an appropriate medium following treatment with citrate and/or cis-aconitate and reduction of culture temperature from 37 ˚C to 27 ˚C (cold shock) (Brun & Schönenberger, 1981; Czichos et al., 1986; Engstler & Boshart, 2004; Overath et al.1986). This *in vitro* differentiation is synchronous when carried out with predominantly stumpy population (Ziegelbauer et al., 1990). Following establishment in the midgut, some procyclic cells cross the peritrophic matrix and colonise the anterior midgut as cell cycle-arrested long mesocyclic forms or proceed to the proventriculus. Once in the proventriculus, the mesocyclic forms become thinner and further develop into long proliferative epimastigote forms (Rose et al., 2020; reviewed in Rotureau & Van Den Abbeele, 2013). Ribonucleic acid (RNA) binding proteins are involved in the transformation to epimastigotes as forced expression of the RNA binding protein ALBA3 hinders this transformation (Subota et al., 2011). Overexpression of RNA binding protein RBP6 *in vitro* also results in successful differentiation of procyclics to epimastigotes and infective metacyclics (Kolev et al., 2012). The epimastigote forms in the proventriculus undergo an asymmetric division to produce long and short epimastigote daughter cells. These cells then migrate from the proventriculus to the fly salivary glands where the short epimastigotes colonise the gland, undergoing elongation after attachment to the epithelium via their flagellum (Rotureau & Van Den Abbeele, 2013).

These cells also express the *brucei* alanine-rich protein (BARP), a stage-specific surface coat for epimastigotes (Urwyler et al., 2007). The attached epimastigotes proliferate and differentiate into free-swimming cell cycle-arrested metacyclic trypanosomes in the lumen of the salivary gland. The differentiation to these mammalian infective metacyclics is characterised by detachment of pre-metacyclics from the salivary gland epithelium and acquisition of a metacyclic VSG (mVSG) (Ramey-Butler et al., 2015; Tetley & Vickerman, 1985). All of these differentiation events during the life cycle are driven by coordinated changes to gene expression.

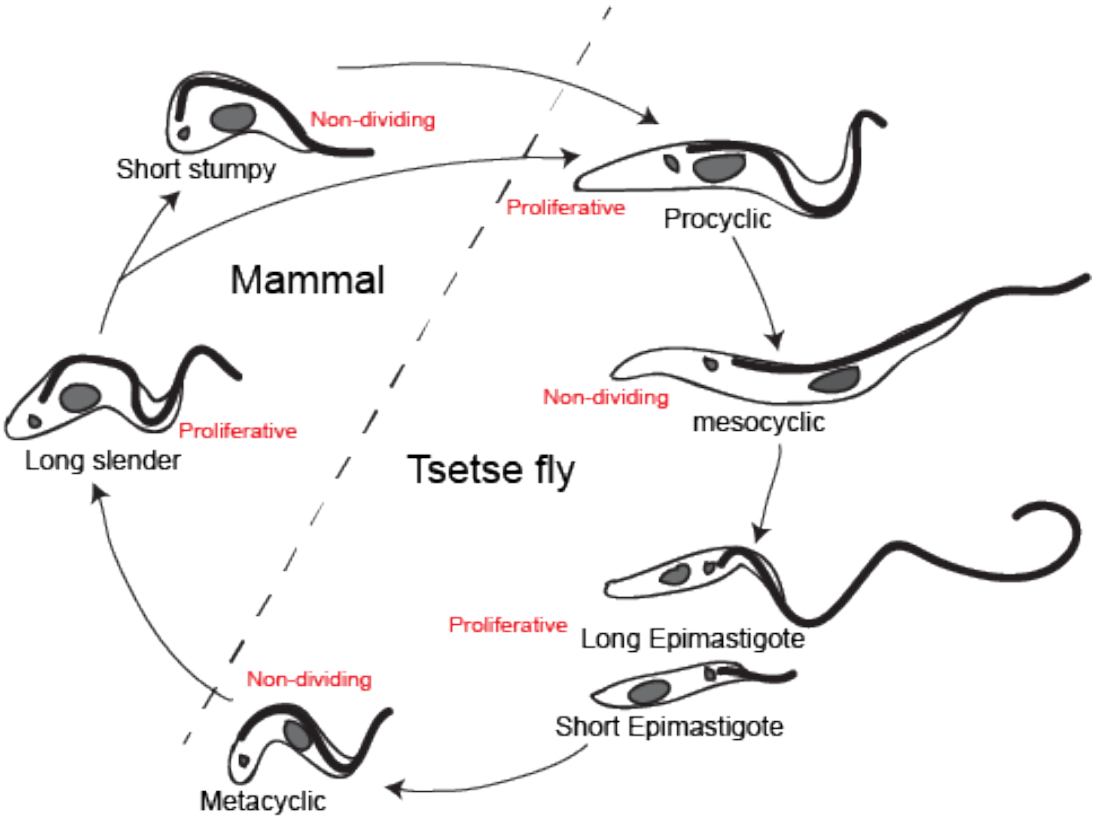

Figure 2. Life cycle of *Trypanosoma brucei*. Non-dividing metacyclic trypanosomes are inoculated into a mammalian host by the tsetse fly during a blood meal. In the mammalian host, the parasite re-enters the cell cycle and differentiates into proliferative long slender trypanosomes which are disseminated to the blood and other tissues. Some long slender cells differentiate to non-dividing short stumpy trypanosomes in response to the quorum sensing signal SIF. Uptake of both the long slender and short stumpy trypanosomes by a tsetse fly establishes an infection in the fly. The cells differentiate into proliferative procyclic trypanosomes in the tsetse fly midgut. Some procyclic cells colonise the anterior midgut as non-dividing mesocyclics which later differentiate into proliferative epimastigotes (long and short) in the proventriculus. The epimastigote cells migrate to the salivary gland where the short epimastigotes attach to the salivary gland epithelium and differentiate into free-swimming non-dividing metacyclic trypanosomes. The figure was adapted and modified from Schuster et al., (2021). Trypanosome cartoons were drawn using fluorescence images from Schuster et al. (2017) as a reference.

1.3 Genome organisation and regulation of gene expression in *Trypanosoma brucei*

Transcription initiation is a major control point for the regulation of gene expression in most organisms. Trypanosomes and related kinetoplastids however appear to have completely lost the ability to regulate transcription by RNA polymerase II (RNA pol-II) (Clayton, 2002). The genome of trypanosomes is organised into long polycistronic transcription units (PTUs) containing several protein-coding genes in a head-to-tail orientation (Daniels et al., 2010). In *T. brucei*, the genome is diploid with a haploid genome size of ~ 35 Mb, containing over 9000 predicted genes distributed among 11 megabase-sized chromosomes (Berriman et al., 2005; El-Sayed et al., 2000). The nuclear genome also contains 1–5 intermediate chromosomes and ~ 100 minichromosomes (Ersfeld, 2011) (Figure 3). Transcription by RNA pol-II initiates bi-directionally at so called divergent strand switch regions (SSR) between two transcriptionally divergent polycistronic transcription units (Martínez-Calvillo et al., 2003; Siegel et al., 2011). Uniquely, RNA polymerase I (RNA pol-I) also transcribes protein-coding genes from the polycistronic VSG expression sites and procyclin loci in bloodstream trypanosomes and procyclics, respectively (Günzl et al., 2003). These RNA pol-I transcribed loci are subject to some level of transcription control (Narayanan & Rudenko, 2013).

Individual messenger RNAs (mRNAs) are processed from polycistronic primary transcripts by a coupled *trans*-splicing and polyadenylation reaction (Lebowitz et al., 1993). In *trans*-splicing, a common 39 nucleotide (nt) mini-exon or spliced leader (SL) sequence derived from a capped ~ 140 nt SL RNA is added to the 5′ end of all mRNAs. The addition of the SL serves two purposes: it functions together with polyadenylation in dissecting polycistronic transcripts, and also provides the cap to the resulting mRNAs (Agabian, 1990; Liang et al., 2003). The process of *trans*-splicing was first discovered in trypanosomes but has since been observed in nematodes, trematodes, euglenoids and chordates (Krause & Hirsh, 1987; Rajkovic et al., 1990; Tessier et al., 1991; Vandenberghe et al., 2001). Unlike *cis*-splicing, exons in *trans*-splicing are derived from two independently-transcribed RNAs which are joined together. *Cis*- and *trans*-splicing proceed through a similar transesterification reaction but in *trans*-splicing, a Y structure is formed instead of a lariat intermediate (Murphy et al., 1986; Sutton & Boothroyd, 1986). *Trans*-splicing is the predominant splicing process in trypanosomes since most genes typically do not contain introns. *Cis*-splicing has however been observed in two intron-containing *T. brucei* genes: the *PAP* gene encoding Poly(A) polymerase and a putative gene encoding an RNA helicase (Günzl, 2010; Jaé et al., 2010; Mair et al., 2000). Trypanosomes lack canonical RNA pol-II promoters, hence transcription is constitutive and the chromatin

structure mediated through histone modifications determines the transcription start and termination sites, several of which are located within SSRs (Ouellette & Papadopoulou, 2009). A recent study aimed at elucidating features that define RNA pol-II transcription start sites (TSS) has however identified GT-rich promoters capable of driving transcription and promoting the targeted deposition of the histone variant H2A.Z (Wedel et al., 2017). The absence of individual gene transcription control due to the polycistronic arrangement of genes makes post-transcriptional regulation the predominant mechanism governing gene expression (Clayton, 2019). Possible regulatory points are *trans*-splicing, polyadenylation, nuclear export and degradation. The levels of protein expressed may also be regulated through control of translation initiation or elongation (Clayton, 2014).

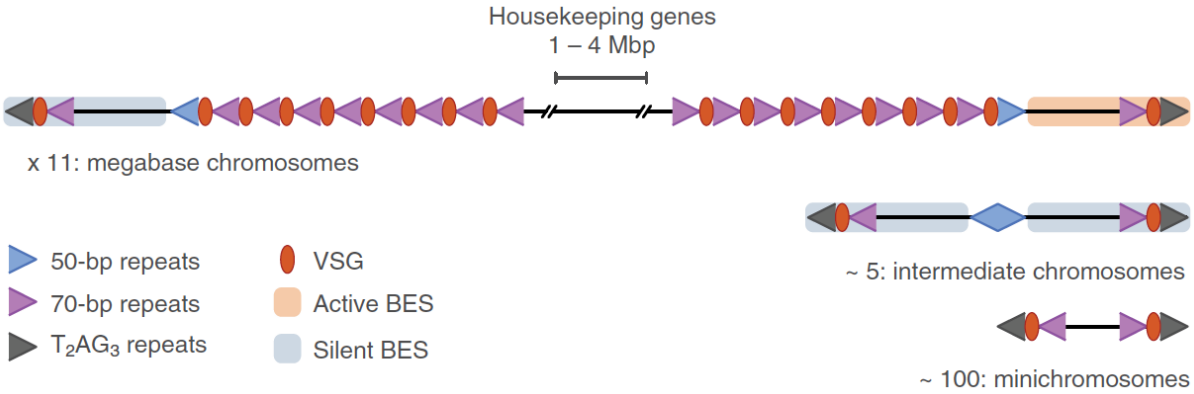

Figure 3. Genome organisation of chromosomes in *T. brucei*. The schematic shows the distribution of the VSG repertoire in the megabase, intermediate and minichromosomes. Image adapted from Glover et al. (2013).

Regulation at the stage of the coupled processes of *trans*-splicing and polyadenylation occurs through alternative splicing. In complex eukaryotes, alternative splicing provides a widespread and powerful approach to expanding the diversity of mRNA and coding capacity of the genome (Preußer et al., 2012). High throughput RNA sequencing data in trypanosomes reveals that only about 11 % of protein coding genes have unique 3´ splice sites, while the rest undergo alternative splicing using between two to three splice sites (Kolev et al., 2010; Nilsson et al., 2010; Siegel et al., 2010). The functional consequences of alternative splicing include the inhibition of translation due to skipping of start codons, inclusion or exclusion of targeting signals, decrease in mRNA stability due to inclusion or exclusion of regulatory signals and the use of alternative open reading frames.

Nuclear export of mRNAs is a key regulatory step in the expression of RNA pol-II transcribed genes in eukaryotes. Trypanosomes lack most factors involved in mRNA export in opisthokonts, raising the questions whether and how trypanosomes regulate nuclear mRNA

export (Kramer, 2021). Experimental evidence however suggests that mRNA export control is not tightly regulated in trypanosomes as dicistronic and polycistronic transcripts have been detected in the cytoplasm (Goos et al., 2019; Kramer et al., 2012).

Another means by which trypanosomes regulate gene expression is by mRNA degradation. Although codon usage contributes to constitutive mRNA decay rates (Jeacock et al., 2018; Nascimento et al., 2018), it has been found that specific sequences within the 3′ untranslated regions (UTRs) of trypanosome mRNAs determines the degradation rate and hence the mRNA abundance (Clayton & Shapira, 2007). RNA binding proteins (RBP) bind to these sequences within the 3′UTR and recruit the mRNA to either the degradation or translation machineries. Degradation of trypanosome mRNAs starts with deadenylation by the CAF1-NOT complex (Erben et al., 2013) followed by APaH-like phosphatase 1 (ALPH1) mediated decapping (Susanne Kramer, 2017), and then either a 5′–3′ degradation by XRNA or a 3′–5′ degradation by the exosome (Estévez et al., 2001).

1.3.1 *Cis*- and *trans*-acting elements in post-transcriptional control of gene expression

Untranslated regions (UTRs) of mRNA are known to play crucial roles in the post-transcriptional regulation of gene expression. Sequences within 3′ UTRs in particular have been shown to play key roles in the regulation of mRNA stability and translation efficiency in trypanosomes (Berberof et al., 1995; Erben et al., 2014; Hotz et al., 1997; Jojic et al., 2018). The major actors in post-transcriptional regulation are *trans*-acting factors such as RBPs, non-coding RNAs and metabolites and *cis*-acting elements, which are regulatory sequences usually found within the UTRs of mRNAs (Fernández-Moya et al., 2014). A single 3′ UTR may contain several *cis*-elements, allowing synergistic, antagonistic or other combinatorial regulatory interactions (Keene, 2007). Interaction between *cis*- and *trans*-acting factors results in the coordinated regulation of multiple mRNAs encoding functionally-related proteins, which are known as post-transcriptional operons or regulons (Keene, 2007).

Several *cis*- and *trans*-acting factors regulating mRNA stability have been described in complex eukaryotes (Caput et al., 1986; Shaw & Kamen, 1986). A notable example is the case of AU-rich elements (AREs), *cis*-sequences located in the UTR of short-lived mRNAs (Chen & Shyu, 1995). In trypanosomes, *cis*- and *trans*-acting elements that regulate the stability of mRNAs that change in response to developmental or environmental triggers have been predominantly investigated to understand the determinants of mRNA levels (Clayton, 2019; Nascimento et al., 2018).

Cis-elements described in *T. brucei* 3′ UTRs include: a 25-nucleotide regulatory element responsible for glycerol-induced changes in GPEET procyclin expression (Vassella et al., 2004), a 35-nucleotide stem-loop in the NT8 purine transporter gene (Fernández-Moya et al., 2014), regulatory sequences in the mRNA of *Proteins associated with differentiation 1* (MacGregor & Matthews, 2012), a 16-nucleotide regulatory element shared by several procyclin mRNAs (Furger et al., 1997; Hehl et al., 1994), a 26-nucleotide stem-loop in the 3′ UTR of *EP1* procyclin (Hotz et al., 1997), regulatory elements in the 3′ UTR of cytochrome oxidase subunits (Mayho et al., 2006), a 7-nucleotide motif in PUF9 target mRNAs (Archer et al., 2009), a 16-nucleotide motif responsible for VSG mRNA stability (Berberof et al., 1995; Ridewood et al., 2017), regulatory elements in the 3′ UTR of translationally controlled tumour protein orthologues (Jojic et al., 2018) and a 34-nucleotide element in the 3′ UTR of ESAG9 family member transcripts (Monk et al., 2013).

Although several studies have focused on the identification of *cis*-elements responsible for the regulation of developmental genes, the specific *trans*-acting factors that bind and control these *cis*-elements largely remain unknown. Recent studies have focused on identifying and characterising RBPs that function as *trans*-acting factors (Kramer & Carrington, 2011). Over 155 RBPs have been identified in *T. brucei*, including proteins identified from bioinformatic predictions and those co-purified with poly(A) RNA (Lueong et al., 2016). These proteins are ordered based on the domains involved in RNA-binding such as RNA recognition motif (RRM) domain containing proteins, CCCH zinc finger proteins and PUF proteins (Kramer & Carrington, 2011). Some RBPs which have been documented to mediate the regulation of gene expression include: RBP10, which promotes bloodstream-form state (Mugo & Clayton, 2017); RBP6, which drives procyclic trypanosome differentiation to epimastigote and metacyclic trypanosomes (Kolev et al., 2012); ZPF3, which binds to the LII *cis*-element in the 3′ UTR of *EP1* mRNA and regulates protein abundance (Paterou et al., 2006; Walrad et al., 2012); RBP24, which interacts with mRNAs encoding proteins involved in energy metabolism (Das et al., 2012); DRBD3/PTB1, which stabilises a subset of developmentally regulated mRNAs encoding membrane proteins (Estévez, 2008); PUF9, which is involved in stabilisation of mRNAs that are preferentially increased during S-phase of the cell cycle (Archer et al., 2009); ZC3H11, which stabilises transcripts involved in the heat-shock response (Droll et al., 2013) and REG9.1, which represses expression of transmission-associated surface proteins in bloodstream-form cells (Rico et al., 2017). A collection of all proteins directly bound to mRNA in bloodstream forms, and the ability of each protein to regulate expression when tethered to the 3′ UTR of a reporter RNA have been documented (Erben et al., 2014; Lueong et al., 2016).

1.4 The variant surface glycoprotein (VSG)

The persistence of trypanosomes in the mammalian host is dependent on the variant surface glycoprotein (VSG), which forms a dense cell surface coat shielding the parasite from immune recognition. The VSG is an approximately 60 kDa protein made up of an elongated N-terminal domain positioned towards the extracellular milieu and a short C-terminal domain that links the VSG to the plasma membrane via a glycosylphosphatidylinositol (GPI)-anchor (Blum et al., 1993; Ferguson et al., 1988). The VSG constitutes about 10 % of the total protein in bloodstream form *T. brucei* cells and is the most abundant protein in these cells. Each individual cell is covered with a layer of 10 million copies of a single VSG making up about 95 % of the surface coat (Cross, 1975; Grünfelder et al., 2002; Wang et al., 2010). This dense ~ 12–15 nm thick layer of VSG acts as a barrier that protects the parasite from antibodies that would otherwise bind to other immunogenic surface proteins that are invariant (Schwede et al., 2011; Schwede et al., 2015). Apart from shielding invariant surface proteins from the host immune system and prevention of complement activation, the VSG is also involved in other strategies used by African trypanosomes to evade clearance by the host immune system. Antibodies are cleared from the parasite surface at low antibody titres, with antibody-VSG complexes being rapidly endocytosed followed by subsequent proteolysis of the antibody and recycling of the VSG back to the surface (Engstler et al., 2007). At higher antibody titres, clearance of antibodies from the parasite surface is ineffective, and survival of the parasite population at this stage is dependent on an elaborate system of antigenic variation which involves switching of the single predominantly expressed VSG isoform (Schwede & Carrington, 2010).

The VSG repertoire in *T. brucei* differs from that in *T. congolense* and *T. vivax* in terms of lineage composition, sequence variation and genomic organisation, thus potentially affecting mechanisms of antigenic variation in these different trypanosome species (Silva Pereira et al., 2021). Classification of VSGs is based on specific patterns of conserved cysteine residues within the N-terminal domain (NTD) and C-terminal domain (CTD) (Carrington et al., 1991). *T. brucei* VSG NTDs are classified based on this as a- or b-type VSGs, which share 6 subgroups of cysteine-rich CTDs (Jackson et al., 2012; Marcello & Barry, 2007). *T. congolense* VSGs on the other hand are exclusively b-type, subdivided into Fam13 and Fam16 which have 15-20 CTDs associated with the specific NTD subfamilies. *T. vivax* has both a- and b-type VSGs but these do not share a conserved CTD as in *T. brucei* (Jackson et al., 2012). VSG repertoires in trypanosomes have therefore evolved to be species-specific in their phylogenetic composition (Silva Pereira et al., 2021). Unique sequence elements may thus be present in some species and

not in others as is the case of the conserved 16mer motif in *T. brucei* VSG 3′ UTR that is absent in *T. congolense* and *T. vivax*.

1.4.1 Antigenic variation and the VSG expression site

Immune evasion is crucial for the survival and persistence of infectious microorganisms in a susceptible host, whether they are bacteria, viruses or protists. Antigenic variation is one of the key strategies employed by these microbes to avoid recognition and clearance by the host immune system. The process of antigenic variation involves the ability of infecting pathogens to systematically alter the proteins exposed to the host immune system, thus presenting the host with a continually changing population that is challenging to eliminate (reviewed in Deitsch et al., 2009). Several pathogenic parasites have been shown to dedicate substantial portions of their genomes to this process, and have developed very complex mechanisms for generating and varying the expression of new antigenic determinants (Morrison et al., 2009). African trypanosomes undergo antigenic variation by continuous switches in the expression of VSGs. In *T. brucei*, individual parasites express a single VSG from a repertoire of ~ 2000 VSG genes and pseudogenes (Berriman et al., 2005; Cross et al., 2014; Marcello & Barry, 2007). Most VSGs are located in gene arrays of the subtelomeric regions of the 11 megabase chromosomes or as individual genes at subtelomeres of intermediate chromosomes and at one-third of all subtelomeres of the 100 minichromosomes (Berriman et al., 2005; Cross et al., 2014) (Figure 3). VSGs are transcribed by RNA polymerase I exclusively from specialised subtelomeric transcription units on intermediate and megabase chromosomes called expression sites (ESs). There are ~ 20 of these expression sites, 5 of which are grouped into metacyclic-form expression sites (MES) and ~ 15 as bloodstream-form expression sites (BES) (Berriman et al., 2005; Hertz-Fowler et al., 2008; Hutchinson et al., 2016). The MESs as the name suggest are activated in the metacyclic trypanosomes and are monocistronic whereas the BESs are activated in the bloodstream-form trypanosomes and are polycistronic transcription units (Figure 4). A typical BES is about 45 – 60 kb and harbours an RNA polymerase I promoter, a series of *expression site associated genes* (ESAGs), and a single VSG gene between a 70 bp repeat and the telomere (Hertz-Fowler et al., 2008). Although BESs share similarities in the general architecture, the number of ESAGs, their composition and potential to encode functional gene products is variable (McCulloch, 2004). The function of several of the ESAGs is unknown but a few have been shown to be involved in nutrient acquisition (ESAG6 and ESAG7) (Schell et al., 1991), and innate immune evasion (ESAG4 and SRA) (Salmon et al., 2012; Xong et al., 1998). The MES are organised as short (3 – 5 kb) monocistronic transcription units and generally lack the 70-bp repeats and ESAGs (Pedram & Donelson, 1999). To ensure

that only a single VSG is expressed at a time in individual cells, the active VSG is transcribed in a mutually exclusive fashion from only one of the ESs (Chaves et al, 1999).

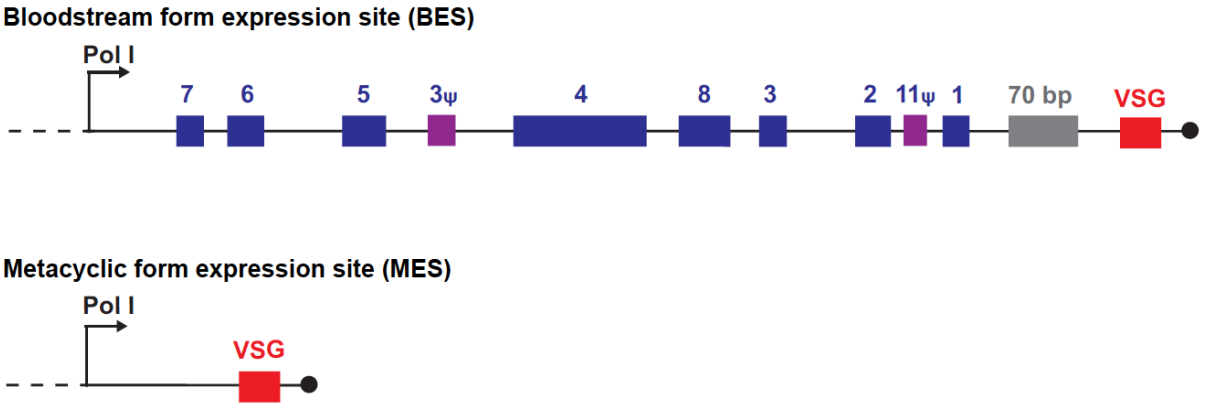

Figure 4. Organisation of VSG expression sites. Schematic of the bloodstream form (upper) and metacyclic form (lower) expression sites. The BES is a polycistronic transcription unit. Downstream of the RNA pol-I promoter (black arrow) on this ES are the expression site associated genes (ESAGs, in blue), pseudo-ESAGs (in magenta), 70 bp repeats (in grey), the VSG gene (in red) and the telomeric repeats represented as the black circle. For the monocistronic MES, downstream of the RNA pol-I promoter (black arrow) is the VSG gene (in red) and the telomeric repeats represented as the black circle. Image adapted from various sources (Borst & Ulbert, 2001; Hertz-Fowler et al., 2008; Taylor & Rudenko, 2006).

1.4.2 VSG switching

VSG switching is central to the process of antigenic variation. Switching of the expressed VSG can occur either by DNA rearrangement involving recombinational exchange of the active VSG or by transcriptional switch (*in situ* switch) which involves activation of a new ES and silencing of the previously active ES (Figure 5) (Taylor & Rudenko, 2006). Transcriptional switching occurs between VSGs occupying the ESs but as there are ~ 20 ESs, this process makes a very modest contribution to the process of antigenic variation and is therefore only effective early in infection (Borst & Ulbert, 2001; Hertz-Fowler et al., 2008). The mechanism governing transcriptional VSG switches is still not fully understand but this mode of VSG switching is obviously related to the process of monoallelic VSG expression (Pays, 2006). A few studies have shed some light on this process by identifying some of the factors involved (Batram et al., 2014; Cestari & Stuart, 2015; Figueiredo et al., 2008; Glover et al., 2016).

Three main pathways contribute to recombination-based VSG switching. These include: telomeric VSG exchange, array VSG conversion and segmental VSG conversion (Stockdale et al., 2008). The telomeric VSG exchange involves a crossover event between subtelomeres that results in replacement of the VSG in the active BES with an inactive telomeric VSG. Array gene conversion involves a duplicative conversion event where the active VSG is exchanged with a VSG from the minichromosome or subtelomeric array. Segmental VSG conversion on

the other hand involves a duplicative event using multiple intact VSGs and pseudo VSGs to generate novel VSG mosaics (da Silva et al., 2018; Hovel-Miner et al., 2016; Pedram & Donelson, 1999; Stockdale et al., 2008). The recombination-based VSG switching is thought to predominate during natural infections because it allows access to the entire VSG repertoire including the VSG pseudogenes (Robinson et al., 1999).

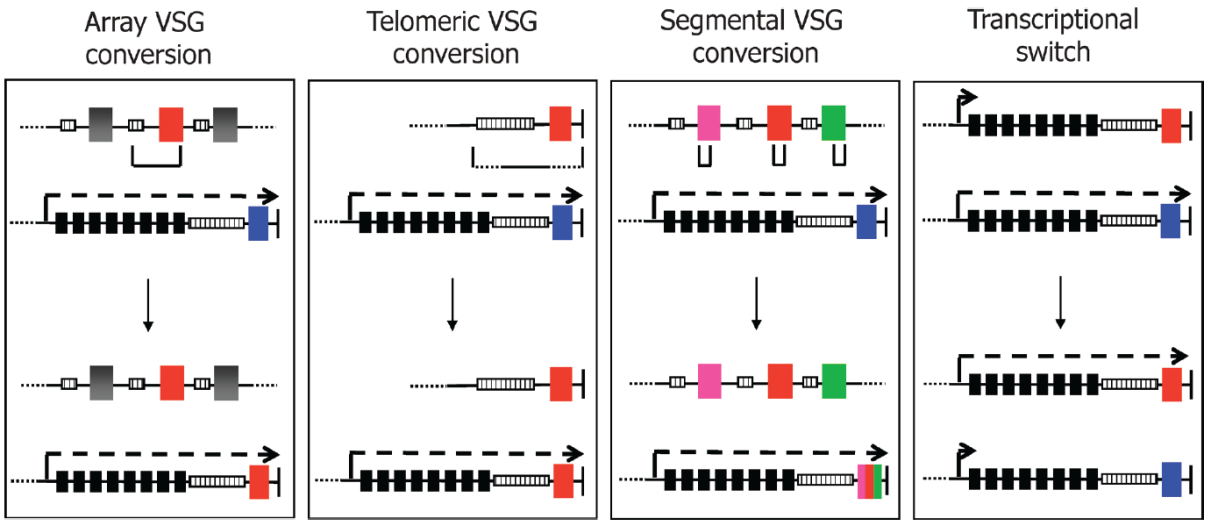

Figure 5. Mechanisms of VSG switching. There are three ways by which recombination-based VSG switching occurs. Telomeric VSG conversion/exchange happens when there is a crossover event between two subtelomeric ends resulting in exchange of the active VSG. In array VSG conversion, donor VSGs from either the subtelomeric VSG array or minichromosomes are copied into the active expression site, replacing the active VSG. In segmental VSG conversion, parts of several VSGs are copied, resulting in a novel mosaic VSG which replaces the VSG in the active BES. VSG switching can also occur by a transcriptional switch which involves the activation of a previously silent BES and silencing of the active BES. The image was adapted from Stockdale et al. (2008).

1.4.3 Allelic exclusion and factors involved in monoallelic VSG expression

Allelic exclusion describes the process by which a single allele of a gene is expressed while the other allele(s) remain silenced. This process is central to the expression of some of the largest known gene-families in protists and mammals. For instance, allelic exclusion is crucial for singular olfactory receptor expression and odour perception in mammals (Monahan & Lomvardas, 2015). Specificity of B and T cells is dependent on allelic exclusion, as this ensures the expression of only a single allele on the surface receptors of these cells (Vettermann & Schlissel, 2010). Protist parasites such as *Plasmodium*, *Giardia* and African trypanosomes also depend on allelic exclusion for efficient antigenic variation (Gargantini et al., 2016; Guizetti & Scherf, 2013; Horn, 2014). In the African trypanosome, *T. brucei*, VSG is expressed in a monoallelic fashion and this is key to the process of antigenic variation and immune evasion. Only one BES is active out of the ~ 15 BES from where a telomeric VSG can be expressed.

The active BES is associated with an extranucleolar RNA pol-I transcription site, referred to as the expression site body (ESB) (Navarro & Gull, 2001). Several molecular factors involved in transcription and regulation of the BES have been identified. These factors are either involved in the control of assembly and remodelling of the BES chromatin, the regulation of RNA pol-I recruitment to the BES promoter or are telomere-associated factors that contribute to repression of transcription or VSG switching (reviewed in Cestari & Stuart, 2018).

The active BES occupying the ESB differs in transcriptional activity compared to silent BESs which appear to be distributed in the nucleoplasm. A model was therefore proposed where the ESB represents an ES subnuclear activation site facilitating full activity of a single BES (Navarro & Gull, 2001). Transcription initiation is however not restricted to the active BES as it occurs simultaneously on multiple BESs but only elongates through the active BES (Kassem et al., 2014). This suggests that the regulation of the ES transcription does not only occur at initiation but also at the elongation and/or RNA processing steps. Studies on class I transcription initiation factor A (CITFA) provided evidence that regulation is controlled at initiation. This multi-subunit factor is enriched together with RNA pol-I at the active ES and is required for RNA pol-I recruitment and transcription activation of the ES, as loss of CITFA sub-units 1, 2 or 7 led to reduced RNA pol-I occupancy at the active ES and decreased expression of *VSG* mRNA and rRNA (Brandenburg et al., 2007; Nguyen et al., 2014; Nguyen et al., 2012). The active BES is also highly SUMOylated at the promoter through to the VSG gene compared to the silent BESs and rRNA promoters (López-Farfán et al., 2014). In addition, an HMG chromatin protein, trypanosome DNA binding protein 1 (TDP1) co-localises with RNA pol-I at the ESB and is enriched in the active ES and nucleolus (Narayanan & Rudenko, 2013). Its depletion results in 40 % – 90 % reduction in *VSG* and rRNA transcripts, and enrichment of histones H1, H2A and H3 at these RNA pol-I transcription units. VSG exclusion 1 (VEX1), a protein that associates with the active BES, was identified as an allelic exclusion regulator (Glover et al., 2016). The protein was later found to be part of a VSG exclusion (VEX) complex that is essential for maintenance of VSG allelic exclusion (Faria et al., 2019).

The chromatin structure, chromatin associated proteins, and telomere-associated factors have been reported to play a role in VSG silencing, thereby contributing to VSG allelic exclusion. For instance, the active ES is depleted of nucleosomes and deletion of histones H1, H3 or H3V results in derepression of silent ESs (Alsford & Horn, 2012; Pena et al., 2014; Povelones et al., 2012; Reynolds et al., 2016; Schulz, et al., 2016). The bromodomain factors 2 and 3 (BDF2 and BDF3) also affect VSG silencing, although inhibition or loss of these factors additionally affects the procyclins and some RNA pol-II transcribed genes (Schulz et al., 2015). The histone

H3K76 trimethyltransferase, DOT1B is additionally involved in the repression of the silent ESs (Figueiredo et al., 2008). Furthermore, a modified DNA base known as base J or hydroxymethyluracil is enriched at silent ESs and the factors responsible for its synthesis have some influence on ES silencing (Reynolds et al., 2016). The *T. brucei* silent information regulator 2-related protein 1 (SIR2RP1) and the histone deacetylases DAC1 and DAC3 have an impact on subtelomeric gene silencing but do not affect the VSG (Alsford et al., 2007; Wang et al., 2010). Factors involved in organisation of the chromatin and its structure including the histone chaperones, FACT (Denninger et al., 2010), ASF1 and CAF1 (Alsford & Horn, 2012), the ISWI chromatin remodeller (Stanne et al., 2015), nucleoplasmin-like protein (N-LP) (Narayanan et al., 2011) and the trypanosome nuclear lamins, NUP1 and NUP2 (DuBois et al., 2012; Maishman et al., 2016) are also required for VSG-ES silencing (reviewed in Duraisingh & Horn, 2016). Knockdown of the ORC1/CDC6 subunit of the origin recognition complex (ORC) and minichromosome maintenance-binding protein (MCM-BP) both result in derepression of ES VSGs (Benmerzouga et al., 2013; Kim, Park, Günzl, & Cross, 2013). Loss of SSC1, a cohesin complex component also affects ES switching (Landeira et al., 2009).

Several telomere-associated proteins including TRF2, RAP1, TIF2 and TRAF contribute to VSG silencing (Glover et al., 2016; Jehi et al., 2014a; Jehi et al., 2014b; Yang et al., 2009). A novel telomere-associated protein TelAP1 was found to influence VSG silencing kinetics during developmental differentiation (Reis et al., 2018). The inositol phosphate pathway has also been found to contribute to VSG silencing and maintenance of monoallelic transcription. Knockdown of phosphatidylinositol 5-phosphatase (PIP5Pase) or phosphatidylinositol 5-kinase (PIP5K), or overexpression of GPI-phospholipase C (GPI-PLC) resulted in derepressed silent VSG-ESs, and in the case of PIP5K, VSG switching was triggered (Cestari & Stuart, 2015).

1.4.4 VSG expression control

The VSG is essential in bloodstream trypanosomes and is the most abundant protein in this life cycle stage with its mRNA accounting for ~ 5 % of the total mRNA (Donelson & Rice-Ficht, 1985; Kraus et al., 2019). The high level of the VSG is attributed to transcription by RNA pol-I and stability of the transcripts. *VSG* mRNAs have long half-lives ranging between 1- 4.5 hours in comparison to half-lives of about 20 min for other transcripts (Ehlers et al., 1987; Fadda et al., 2014; Ridewood et al., 2017). The high abundance of VSG protein and the extremely high stability of the *VSG* transcripts compared to *ESAGs* (which are transcribed from the same PTU) suggest an additional level of control over VSG expression. Apart from factors involved in monoallelic transcription contributing to expression of the active VSG, the trypanosome cell is also able to modulate the levels of the transcribed *VSG*, as expression of an ectopic VSG results

in reduction of the levels of the endogenous VSG. Integration of a second VSG in the active expression site resulted in expression of both VSGs at a ratio of approximately 50/50 (Muñoz-Jordán et al., 1996; Smith et al., 2009). When a second VSG is also integrated at various genomic locations, there is a reduction of the endogenous *VSG* levels proportional to the ectopic *VSG* levels (Ridewood et al., 2017). In addition, overexpression of a second VSG from the rDNA spacer region under a T7 promoter resulted in rapid decrease of the endogenous VSG and a DOT1b-dependent attenuation of the active BES (Batram et al., 2014). The overexpression of the ectopic VSG produces a phenotypic plasticity as a result of varying degrees of ES-attenuation in pleomorphic trypanosomes (Zimmermann et al., 2017). A similar observation was made in monomorphic trypanosomes where the phenotypic plasticity observed depended on the active BES and the VSG isoform being overexpressed (Batram et al., 2014; Goos, Master thesis 2013; Henning, Bachelor thesis 2012; Specht, Master thesis 2013).

Little is known about the mechanisms involved in the high stability and modulation of the *VSG* mRNA levels. A common feature of all VSGs is the presence of two highly conserved motifs, the 8mer and 16mer within the 3´ UTRs. These motifs are thought to contribute to the mechanisms modulating post-transcriptional regulation of the VSGs. The 16mer motif in particular is 100 % conserved in all analysed *T. brucei VSG* transcripts and has been shown to be essential for *VSG* mRNA stability (Berberof et al., 1995; Ridewood et al., 2017). During the developmental transition from bloodstream to procyclic forms, the *VSG* mRNA becomes unstable and is rapidly lost. It has been proposed that a factor interacting with the 16mer motif is likely responsible for the maintenance of *VSG* mRNA stability and modulation of the VSG levels in the presence of more than one VSG (Ridewood et al., 2017). A recent study suggested that the 16mer is required for N6-methyladenosine (m^6A) modification of the poly (A) tails of *VSG* mRNAs (Viegas et al., 2020). This m^6A modification prevents deadenylation and promotes *VSG* mRNA stability. Another recent study has suggested that the 16mer binding factor is CFB2, an RNA binding protein which interacts with the 16mer motif, mediating *VSG* mRNA stability by recruitment of a stabilising translation-promoting complex (Nascimento et al., 2021). For modulation of *VSG* mRNA, the existence of a pathway that regulates the mRNA levels linked to functional VSG protein production has been proposed (Maudlin et al., 2020). A recent study however suggests that functional VSG protein production is dispensable for modulation of the *VSG* mRNA levels, and that endoplasmic reticulum (ER) targeting of *VSG* transcripts induces the regulation (Aroko et al., 2021).

1.5 Ribosome biogenesis and rRNA processing

Besides mRNA stability, another factor influencing the abundance of an encoded protein is the process of translation, which is carried out by ribosomes. Ribosomes are macromolecular complexes that serve as sites for protein synthesis. These complexes are large ribonucleoprotein (RNP) particles consisting of two uneven subunits of 30S and 50S (in prokaryotes) or 40S and 60S (in eukaryotes). A prokaryotic ribosome contains a 16S rRNA and 21 ribosomal proteins in the 30S subunit, while the 50S subunit has 33 ribosomal proteins as well as two rRNAs, the 5S and 23S rRNA (reviewed in Shajani et al., 2011). Eukaryotic ribosomal subunits on the other hand are more complex with the 40S small subunit (SSU) containing 33 ribosomal proteins and an 18S rRNA, whereas the 60S large subunit (LSU) consists of 5S, 5.8S and 25/28S rRNAs as well as 46 ribosomal proteins (Henras et al., 2015). In addition to the rRNAs and ribosomal proteins, the eukaryotic ribosomal subunits require > 350 specialised non-ribosomal factors for assembly (Gerhardy et al., 2014). The biosynthesis of ribosomes in eukaryotes is therefore a highly complex and energy-consuming process. All three RNA polymerases are involved in this process. RNA pol-I transcribes the precursor to the 18S and 25/28S rRNAs, RNA pol-III produces the precursor to the 5S rRNA and RNA pol-II synthesises pre-mRNAs of ribosomal proteins and other additional factors involved in the biogenesis (Henras et al., 2008). Ribosome biogenesis is a spatially-regulated process that begins in the nucleolus, continues through the nucleoplasm and ends with the assembly of mature ribosomes in the cytoplasm. Most aspects of ribosome biogenesis are conserved among eukaryotes and begins with pre-rRNA transcription and processing. The rRNA processing involves the combined action of small nucleolar RNAs and accessory protein factors such as GTPases, ATPases, helicases and endonucleases (Michaeli, 2012).

The process of ribosome biogenesis is best described in the eukaryotic model organism *Saccharomyces cerevisiae*. In this organism, the pre-rRNA is transcribed as a long 35S rRNA primary transcript containing sequences for the 18S, 5.8S and 25S rRNAs, separated by two internal transcribed spacers (ITS1 and ITS2) and flanked by two external transcribed spacers (5´ ETS and 3´ETS) (Venema & Tollervey, 1999). The 35S pre-rRNA undergoes processing involving base modifications, cleavages and trimmings to form the mature rRNAs (18S, 5.8S and 25/28S). Initial cleavages occur at the A0 and A1 sites, followed by cleavage at the A2 site which separates the small and large ribosomal subunit precursors (20S and 27SA2) and removes the 5´ ETS. The 20S pre-rRNA is processed further in the cytoplasm into the 18S rRNA. Processing of the 27SA2 pre-rRNA proceeds in the nucleus, and is either cleaved at the A3 or

B1 position followed by final cleavages at the C1 and C2 sites to produce 5.8S and 28S rRNAs (reviewed in Michaeli, 2012; Venema & Tollervey, 1999).

T. brucei diverged early from the other major eukaryotic lineages and several of its biological features, including rRNA metabolism, are unique. The processing of rRNA in this parasite therefore differs from that described in yeast and other eukaryotes (Figure 6). The large subunit of the rRNA, for instance, is cleaved into two large subunits and four smaller rRNA fragments (Campbell et al., 1987; Cordingley & Turner, 1980; Schnare et al., 1983; White et al., 1986). Processing of the rRNA in *T. brucei* begins with an initial cleavage at the B1 site of the 9.2 kb/9.6 kb pre-rRNA primary transcript which separates the processing pathway into two arms (pre-SSU and pre-LSU intermediates). The pre-SSU (~ 3.4 kb species) is further cleaved and modified to produce the mature 2.2 kb SSU. The pre-LSU (~ 5.9/5.8 kb) on the other hand is further processed into two large fragments and four smaller rRNAs which form part of the mature LSU together with 5S rRNA which is transcribed by RNA pol-III. In this arm of the processing pathway, predominant stable rRNA processing intermediates of 5.1 kb /5.0 kb and 3.9 kb are usually detected (reviewed in Michaeli, 2012; Rajan et al., 2019; Rink et al., 2019; Umaer et al., 2014).

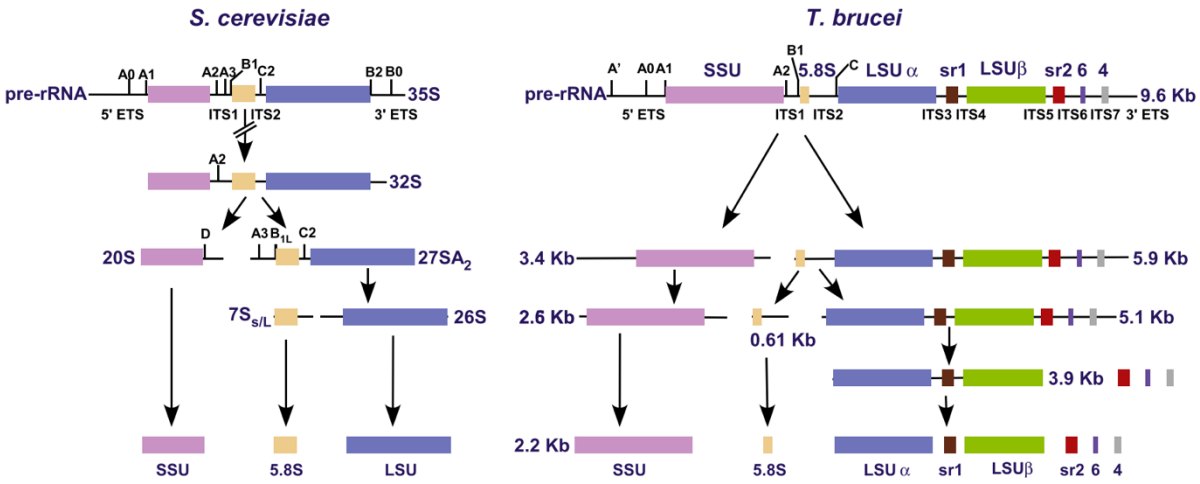

Figure 6. Comparison of the rRNA processing pathway in *Saccharomyces cerevisiae* (left) and *Trypanosoma brucei* (right). Figure adapted from Rajan et al. (2019). Note that in the *T. brucei* processing pathway, the 9.6 kb pre-rRNA, 5.9 kb intermediate and 5.1 kb intermediate are reported as 9.2 kb, 5.8 kb and 5.0 kb in some studies (Rink et al., 2019; Umaer et al., 2014).

1.6 DExD/H box helicases

The DExD/H box family of proteins constitutes a large number of proteins that are involved in many different roles in RNA metabolism. These proteins are conserved in nature and can be found in viruses, bacteria and eukaryotes. They are distinguished by the presence of a helicase core of about 8 – 9 conserved motifs including the walker B motif (motif II) which contains the characteristic 'DExD/H' sequence (where x can be any amino acid) (Fuller-Pace, 2006; Linder

& Jankowsky, 2011). The motifs within the helicase core are involved in either RNA binding (motifs Ia, Ib, IV, Iva, V and GG), NTP binding and hydrolysis (motif Q, I, II and VI) or couple NTP binding to RNA unwinding activity (Figure 7). The conserved helicase core is flanked by non-conserved N- and C- terminal extensions which are thought to confer physiological specificity through interactions with other proteins or recognition of specific RNA sequences (Fairman-Williams, Guenther, & Jankowsky, 2010). DExD/H proteins were proposed to act primarily as ATP-dependent RNA helicases based on sequence homology to DNA helicases and the ability of some members to use ATP/NTP to unwind RNA duplexes *in vitro* (Fuller-Pace, 2006). Emerging evidence however indicates that some members of this family of proteins act as chaperones that facilitate the rearrangement and formation of optimal RNA secondary structures (Jankowsky, 2011; Tanner & Linder, 2001). Some members of the DExD/H family are involved in transcriptional regulation and in several cases, their role in this process is independent of their RNA helicase activity (Fuller-Pace, 2006; Gillian & Svaren, 2004; Nakajima et al., 1997; Wilson et al., 2004).

Processing of the pre-rRNA transcripts during ribosome biogenesis requires several *trans*-acting factors including RNA helicases. In fact, RNA helicases of the DExD/H family are the most abundant *trans*-acting protein factors involved in this process (Kressler et al., 1999). In yeast, 19 out of 37 DExD/H RNA helicases have been implicated in the biosynthesis of ribosomes (Strunk & Karbstein, 2009). These RNA helicases have either been implicated in SSU biogenesis (Dhr1, Dhr2, Dbp4, Dbp8, Fal1, Rok1, Rok3), LSU biogenesis (Drs1, Dbp2, Dbp3, Dbp6, Dbp7, Dbp9, Dbp10, Mtr4, Mak5, Spb4) or both SSU and LSU biogenesis (Has1, Prp43) (Martin et al., 2013). Depletion or deletion experiments indicate that all of the RNA helicases involved in the biosynthesis of ribosomes are essential for either cell growth or viability. Homologues of yeast proteins involved in ribosome biogenesis have been identified in trypanosomes by bioinformatic analysis and purification of the processomes of the ribosomal subunits from the related kinetoplastida *Leishmania tarentolae* (Michaeli, 2012; Rajan et al., 2019). Among these proteins are homologues to DEAD box RNA helicases. There are about 51 proteins of the DExD/H family in the *T. brucei* genome (Aslett et al., 2010; Gargantini et al., 2012) but only a few of these have been experimentally characterised. In general, only a small number of proteins in *T. brucei* have been shown experimentally to be involved in ribosome biogenesis or rRNA processing. The trypanosome-specific nucleolar protein NOPP44/46 and its interacting proteins NOG1 and P34/37 have been shown to be involved in the processing of the 60S ribosomal subunit (Jensen et al., 2005; Jensen et al., 2003; Pitula et al., 2002). Depletion of the pumilio-domain protein PUF7 results in inhibition of ribosomal

RNA processing (Droll et al., 2010). Depletion of PUF7, PUF10, NRG1 (nucleolar regulator of GPEET1) and the trypanosome homologue of ERB1 in yeast also affects processing of the large ribosomal subunit (Burkard et al., 2007). TbPNO1, TbNOB1 and TbUTP10 are all involved in the processing of the 18S rRNA (Faktorová et al., 2018; Kala et al., 2017). The exoribonuclease XRNE is also involved in ribosome biogenesis as its depletion results in accumulation of the 18S and 5.8S rRNAs (Sakyiama et al., 2013). Depletion of TbRRP44, TbMTR4, TbMTR2, TbMex67 and proteins from the exosome all cause accumulation of rRNA processing intermediates (Cesaro et al,, 2019; Cristodero & Clayton, 2007; Estévez et al., 2001; Rink et al., 2019).

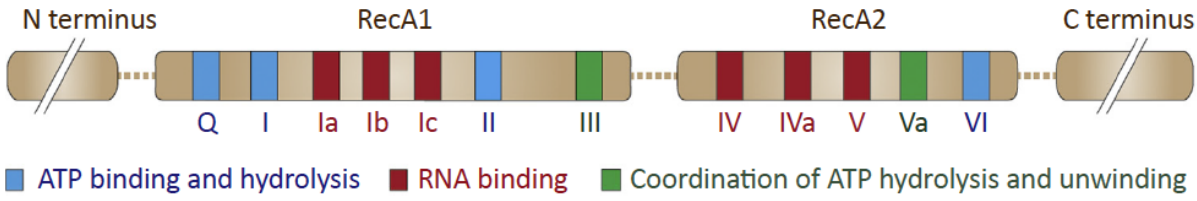

Figure 7. Schematic showing the characteristic domains of the DExD/H helicases. The role played by each motif is shown by the colour-coded legend. Figure adapted from Sloan & Bohnsack (2018).

1.7 Aims of the study

The control of VSG expression is a complex process involving multiple factors mediating expression of the *VSG* gene in the context of antigenic variation, monoallelic expression/expression site control, developmental regulation and stability of *VSG* transcripts. Although previous studies have identified some factors involved in VSG expression, the complete molecular process governing the regulation of expression is not fully understood.

Generally, regulation of gene expression in trypanosomes is post-transcriptional as the genome is organised in long polycistronic transcription units. Conserved sequences present in the UTRs, notably the 3´ UTRs, act as key elements in the modulation of individual mRNA levels. *T. brucei VSG* transcripts harbour the highly conserved 8mer and 16mer motifs within their 3´ UTRs. The 16mer motif — which is 100 % conserved in all *T. brucei* VSGs — has been shown to be essential for maintenance of *VSG* mRNA stability. Such a high conservation with no observable gaps or point mutations is however unusual for a stabilisation or destabilisation motif. Despite the unusually high conservation of this motif, no other functions besides *VSG* mRNA stability have been assigned. The general aim of this study was therefore to identify any additional roles besides mRNA stability that may be driving the 100 % conservation of the 16mer motif. The specific aims of the study were;

1. To investigate whether a 100 % conservation of the 16mer motif is necessary for functional VSG expression and cell viability.
2. To investigate the role of the intact 16mer in VSG silencing and expression site attenuation.
3. To identify and characterise interaction partners of the 16mer motif.

2.0 MATERIALS AND METHODS

2.1 Materials

2.1.1 Antibodies for Western and protein dot blots

Table 1. Antibodies used for Western blot and protein dot blot analyses

Antibody	Organism	Dilution	Source
Primary antibodies			
Anti-GFP	Mouse	1:1000	Roche, Rotkreuz (CH)
Anti-PFR1,2 L13D6	Mouse	1:20	(Kohl et al., 1999)
Anti-VSG121	Rabbit	1:2000	(Batram et al., 2014)
Anti-VSG221	Rabbit	1:5000	N. Jones, Würzburg (DE)
Anti-VSG118	Rabbit	1:10000	N. Jones, Würzburg (DE)
Anti-m^{6}A	Mouse	1:1000	Abcam
Anti-puromycin (12D10)	Mouse	1:5000	Sigma
Secondary antibodies			
Anti-mouse IgG *(IRDye680LT)*	Goat	1:10000	LI-COR Biosciences
Anti-rabbit IgG *(IRDye800CW)*	Goat	1:10000	LI-COR Biosciences
Anti-mouse (HRP)	_	1:10000	

2.1.2 Probes and dyes

Table 2. Probes used for visualisation of RNA bound to nylon membranes

Probe	Fluorophore /Label	Sequence (5′–3′)
Tubulin	5′ DY-782	ATCAAAGTACACATTGATGCGCTCCAGCTGCAGGTC
GFP	5′ DY-682	GCCGTTCTTCTGCTTGTCGGCCATGATATAA
VSG221	5′ DY-682	CAGCGTAAACAACGCACCCTTCGGTTGGTCGTCTAG
VSG121	5′ DY-682	GGCTGCGGTTACGTAGGTGTCGATGTCGAGATTAAG
VSG118	5′ DY-682	TCTTTCTCTCCTGTTGCTTGGTTTTTCTGT
16mer	5′ Biotin	UGAUAUAUUUUAACAC
Pre-18S	5′ DY-682	TCAAGTGTAAGCGCGTGATCCGCTGTGG
ITS2	5′ DY-682	ATCACTCACTACACACACGTAT
ITS3	5′ DY-682	ACGACAATCACTCACACACACATGGC
Mini exon	IRDye 700	CAAT ATAGTACAGAAACTGTTCTA ATAATAGCGTT
18S rRNA	IRDye 800	CCTTCGCTGTAGTTCGTCTTGGTGCGGTCTAAGAATTTC

2.1.3 Buffers and Solutions

T. brucei cultivation

HMI-9 medium

For 1L of HMI-9 medium, 17.66 g Iscove's modified Dulbecco's medium (IMDM), 3 g sodium bicarbonate, 136 mg hypoxanthine, 28.2 mg bathocuproine sulfonate, 14 µl β-mercaptoethanol, 39 mg thymidine, 100,000U penicillin, 100 mg streptomycin, 182 mg cysteine and 10 % (v/v) foetal calf serum (FCS) were used. Components of the medium were dissolved in filtered deionised water and the pH adjusted to 7.5. The medium was then sterilised by filtration (pore size 0.2 µm), followed by addition of the 10 % heat-inactivated FCS.

Trypanosome dilution buffer (TDB), pH 7.6

20 mM Na_2HPO_4, 2 mM NaH_2PO_4 (pH 7.7), 5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$ (pH 7.7), and 20 mM glucose in ddH_2O (filter sterilised).

2x Trypanosome freezing mix

Filter sterilised 20 % (v/v) glycerol in HMI-9 medium

DNA analyses

6x DNA loading dye

10 mM Tris-HCl (pH 7.6), 0.03 % bromophenol blue, 0.03 % xylene cyanol FF, 60 % glycerol 60 mM EDTA

1x TAE buffer

40 mM Tris, 40 mM concentrated acetic acid, 1 mM EDTA (pH 8.0)

RNA analyses

Denhardt's solution (DEN, 50x)

1 % (w/v) Ficoll 400, 1 % (w/v) polyvinylpyrolidone, 1 % (w/v) bovine serum albumin (BSA)

Glyoxal solution

75.4 % (v/v) DMSO, 21.5 % (v/v) Glyoxal (deionised), 15 mM NaH_2PO_4 (pH 6.9)

Northern blot hybridisation solution

5x SSC, 5x DEN, 0.1 % (w/v) SDS, 0.1 % (w/v) tetrasodium pyrophosphate, 0.01 % (w/v) heparin

Northern blot washing solution

0.1 % (w/v) SDS, 0.1x SSC

RNA buffer (50x)

0.5 M NaH_2PO_4 (pH 6.9)

SSC (20x)

3 M NaCl, 0.3 M trisodium citrate, (pH 7.0)

MOPS buffer (10x)

200 mM 3-[N-morpholino]propanesulfonic acid (MOPS) (free acid), 50 mM sodium acetate, 10 mM EDTA, pH adjusted to 7.0 with NaOH.

Protein analyses

Phosphate buffered saline (PBS)

10 mM Na_2HPO_4, 1.7 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, pH 7.4

Sample buffer (2x)

120 mM Tris-HCl, pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 0.004 % (w/v) bromophenol blue, 2 % (v/v) β-mercaptoethanol.

Laemmli running buffer (1x)

25 mM Tris, 192 mM glycerol, 0.1 % (w/v) SDS.

Stacking gel buffer

0.5 M Tris-HCl (pH 6.8), 0.4 % (w/v) SDS.

Stacking gel

1 ml acrylamide/bis-acrylamide 30:0.8, 1.8 ml stacking gel buffer, 4.5 ml ddH_2O, 40 µl 10 % (w/v) APS, 10 µl TEMED

Resolving gel buffer

1.5 M Tris-HCl (pH 6.8), 0.4 % (w/v) SDS

Resolving gel (12.5 %)

1 ml acrylamide/bis-acrylamide 30:0.8, 1.5 ml resolving gel buffer, 2.5 ml ddH_2O, 30 µl 10 % APS (w/v), 10 µl TEMED

Coomassie solution

10 % (w/v) 2-propanol, 5 % (v/v) glacial acetic acid, 0.05 % (w/v) Coomassie R-250 in ddH_2O

Destaining solution

10 % (v/v) 2-propanol, 5 % (v/v) glacial acetic acid in ddH_2O

2.1.4 Enzymes

All the enzymes used in this study were either obtained from ThermoFisher Scientific or New England BioLabs

2.1.5 Chemicals

Chemicals used in this study were obtained from Sigma-Aldrich (St. Louis, USA), AppliChem (Darmstadt, DE) or Carl Roth (Karlsruhe, DE) unless otherwise stated in the text.

2.1.6 Equipment and devices

Table 3. List of equipment and devices used in this study

Equipment /Device	**Manufacturer**
Airstream Class II BSC	Esco global, Changi (SG)
Binder CO_2 incubator	Binder, Tuttlingen (DE)
AMAXA Nucleofector II	Lonza, Basel (CH)
iBright FL1000 image scanner	Invitrogen, Waltham (USA)
LI-COR Clx Imager	LI-COR Biosciences (USA)
MiniFold dot blotter	GE Healthcare, Little Chalfont (UK)
Neubauer chamber	Marienfeld, Lauda-Königshofen (DE)
T-100 Thermo cycler	Bio-Rad Laboratories, Munich (DE)
Trans-Blot semi-dry transfer cell	Bio-Rad Laboratories, Munich (DE)
Tecan Infinite M200 Plate reader	Tecan, Männedorf (CH)
Gel iX imager	Intas Science Imaging Instruments, Göttingen (DE)
EM AFS2 freeze substitution system	Leica Microsytems, Wetzlar (DE)
EM HPM100 HPF-system	Leica Microsytems, Wetzlar (DE)
DMI6000B wide-field microscope	Leica, Wetzlar (DE)
Eclipse TS100 microscope	Nikon, Tokyo (JP)

2.1.7 Kits

Table 4. List of commercial kits used in this study

Kit	**Manufacturer**
High Pure PCR Template Preparation kit	Roche Diagnostics, Rotkreuz (CH)
CloneJET PCR Cloning Kit	ThermoFisher Scientific, Waltham (USA)
Amaxa Basic Parasite Nucleofector Kit 1	Lonza, Basel (CH)
NucleoSpin Plasmid Kit	Macherey-Nagel, Düren (DE)
NucleoSpin Gel and PCR clean-up kit	Macherey-Nagel, Düren (DE)
NucleoBond PC 100	Macherey-Nagel, Düren (DE)
RevertAid First Strand cDNA Synthesis Kit	ThermoFisher Scientific, Waltham (USA)

RNeasy Mini Kit	QIAGEN, Venlo (NL)
Oligotex mRNA Mini Kit	QIAGEN, Venlo (NL)
LightShift Chemiluminescent RNA EMSA Kit	ThermoFisher Scientific, Rockford (USA)

2.1.8 Oligonucleotides

Table 5. List of primers used

ID	Sequence (5' - 3')	Designed by
MBS 7	ACGGATCCCATTCCTACCTACCTAC	Majeed Bakari Soale
MBS 8	AACGAGCTCAAGCAGCAGAAACAG	Majeed Bakari Soale
MBS11	GGATCCTCCTAGCCCAAGTTCTTC	Majeed Bakari Soale
MBS12	TCTAGAGGTGCCGTGATTTATTTC	Majeed Bakari Soale
MBS13	GAATTCCGCATCTTCACTGGTGTC	Majeed Bakari Soale
MBS14	AAGCTTCAGAGGGCAGCAATTCAC	Majeed Bakari Soale
MBS17	GAAATAAATCACGGCACCGAGCTCTCGATTTACG GTTCTGAAG	Majeed Bakari Soale
MBS18	AGTGAAGATGCGGAATTCGCTAACAAATGACGCA CC	Majeed Bakari Soale
MBS21	GGGGACAAGTTTGTACAAAAAAGCAGGCTACCCA TACGATTCGACCACAG	Majeed Bakari Soale
MBS22	GGGGACCACTTTGTACAAGAAAGCTGGGTACTCA CGTTCACCTCCGTACT	Majeed Bakari Soale
MBS23a	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAGT GAGGACAAAAGCG	Majeed Bakari Soale
MBS24a	GGGGACCACTTTGTACAAGAAAGCTGGGTGCTGC CTCAGGAGATTTG	Majeed Bakari Soale
MBS25a	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTATG TGAGGGAGAGCGACG	Majeed Bakari Soale
MBS26a	GGGGACCACTTTGTACAAGAAAGCTGGGTTGCCG TGAGCTTAACTTCGT	Majeed Bakari Soale
MBS27	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTTCG GGTGCCATGGAGATT	Majeed Bakari Soale
MBS28	GGGGACCACTTTGTACAAGAAAGCTGGGTCGGCA AGTATATGGCGGCTA	Majeed Bakari Soale
MBS29	GGGGACAAGTTTGTACAAAAAAGCAGGCTCACCG CATTGGCCTTGTTC	Majeed Bakari Soale
MBS30	GGGGACCACTTTGTACAAGAAAGCTGGGTTCCGC GACCAGAAGCTCTA	Majeed Bakari Soale
MBS31a	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGAAT GGCTTGCGCGATTTA	Majeed Bakari Soale

MBS32a	GGGGACCACTTTGTACAAGAAAGCTGGGTTCGTTTACAGGTGGTGCCTC	Majeed Bakari Soale
MBS33	TAAAAGTAGCGCTTACGG	Majeed Bakari Soale
MBS34	TGCCTGCACTAACACTAC	Majeed Bakari Soale
MBS37	GGGGACAAGTTTGTACAAAAAAGCAGGCTGACCATCGTGTTTCACCC	Majeed Bakari Soale
MBS38	TATTGTGTTTCGGGAGCGACAAGGGAAGTAGAGCCTCGAAC	Majeed Bakari Soale
MBS39	GTTCGAGGCTCTACTTCCCTTGTCGCTCCCGAAACACAATA	Majeed Bakari Soale
MBS40	GGGGACCACTTTGTACAAGAAAGCTGGGTGTTCATCGTGCAGTTGTAAGCA	Majeed Bakari Soale
MBS42	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAGTGCGGCAACCTACA	Majeed Bakari Soale
MBS43	GGGGACCACTTTGTACAAGAAAGCTGGGTTAGAGCTCATTGCCACAC	Majeed Bakari Soale
MBS44	TCTGTGCAAAAGATAACCGA	Majeed Bakari Soale
MBS45	TAGAGCTCATTGCCACAC	Majeed Bakari Soale
MBS46	GGATTGTCGTCCAGGATTAG	Majeed Bakari Soale
MBS47	AAGCGGAGGTATGGTTGA	Majeed Bakari Soale
MBS48	AAGGTTTGTCTGGAACTCTG	Majeed Bakari Soale
MBS49	CGTGAGAATGATTACCGGAT	Majeed Bakari Soale
MBS50	ATCAACTTCCCAAACCCATT	Majeed Bakari Soale
MBS51	TGTGGCTTATCTCTGACAAC	Majeed Bakari Soale
MBS52	CCGAATCCCATACTGAAGAC	Majeed Bakari Soale
MBS53	TTTGAGTACCCGGCCAACAGCTCGCATGAAGTGAGGCAGAAGCTGTGGAAGTTTTCATTGAAGGAAGTCACACCACAAAATGGTACCGGGCCCCCCCTCGAG	Majeed Bakari Soale
MBS54	CTGTGATGGGTTAATTCAGCTTTTCTCTCGTTAAGTTAGTTTTTAAACGTGGCACGGGGGGCAGCCATTTTCCACATTCATTGGCGGCCGCTCTAGAACTAGTGGAT	Majeed Bakari Soale
MBS56	TGGCGGCCGCTCTAGAACTAGTGGAT	(Oberholzer et al., 2006)
MBS59	AAGCTTATGGCACCTCGTTCCCTTTA	Majeed Bakari Soale
MBS60	AGAAAGGATGTCACCAAAAAAATATAACGAAAAAATATTAAATGAATGTCCACACTTGTTTTACAAGAAACAGCAGGAAACAAACACGATGC	Majeed Bakari Soale

MBS61	GCTGTTTCTTGTAAAACAAGTGTGGACATTCATTTAATATTTTTTCGTTATATTTTTTTGGTGACATCCTTTCTAATGCCTTATTAACCATCGC	Majeed Bakari Soale
MBS62	CTCGAGCAATAACACAAGGCACCA	Majeed Bakari Soale
MBS71	AAGCTTATGGCCGTGCACAGAG	Majeed Bakari Soale
MBS72	CTCGAGCCTCCTACCTAGCTACC	Majeed Bakari Soale
MBS75	GAAAAAGAGGGGGAAAATTAAAAAAGCAAGGCCACAAATG	Majeed Bakari Soale
MBS76	TTTAATTTTCCCCCTCTTTTTC	Majeed Bakari Soale
MBS77	AAGCTTATGAACATCTACAGTTGGG	Majeed Bakari Soale
MBS78	GGATCCTTATGCGTAATCGGGCACATCGTACGGGTATGCGTAGTCTGGCACGTCGTATGGGTACGCGTAATCAGGCACATCGTAAGGGTAATTTTGTGGTGTGACTTCC	Majeed Bakari Soale
MBS79	AGCCTTTTGAGAGTTTGG	Majeed Bakari Soale
MBS80	CCAAACTCTCAAAAGGCTTCGGAGGGGCAAACAGAACGGAAA	Majeed Bakari Soale
MBS81	CACCTGATCTACGACAATG	Majeed Bakari Soale
MBS82	CATTGTCGTAGATCAGGTGG	Majeed Bakari Soale
MBS85	AAGCTTATGATTCATTCAAATAAAGTAG	Majeed Bakari Soale
MBS86	ACGTGTTAAAATATAAGTGAAAG	Majeed Bakari Soale
MBS87	CTTTCACTTATATTTTAACACGT	Majeed Bakari Soale
MBS88	CTCGAGAACCCTAACCCTAACCCTAA	Majeed Bakari Soale
C14	AATATAAGTGGAGTTTTCAAG	Christopher Batram
C15	AACTCCACTTATATTTTAACACGC	Christopher Batram
C80	AAGCTTCCTCCTACCTAGCTACCTAC	Christopher Batram
C81	GGATCCTAGAAAGTGTGACAACGTCGC	Christopher Batram
IN1	CTCGAGATGGTGAGCAAGGGCGA	Nonso Ikenga & Majeed Bakari Soale
IN2	GTCGACTTACTTGTACAGCTCGTCCA	Nonso Ikenga & Majeed Bakari Soale

HZ30	CATGGCCGTGCACAGAGC	Henriette Zimmermann
HZ31	ATTAATTAAAAAAGCAAGGCCACA	Henriette Zimmermann
HZ32	GGCGAAGCTTCCTCCTACCT	Henriette Zimmermann
HZ35	CGCCACAAGCATTCTATACGTAAA	Henriette Zimmermann
HZ49	CTACTTGAAAACTCCAACACGCAAA	Henriette Zimmermann
HZ50	CGGGTAATTTGCGTGTTGGAGTTTTCA	Henriette Zimmermann

2.1.9 Plasmids

Table 6. List of plasmids vectors used in this study

Construct name	Description/Purpose	Resistance in *E. coli*	Resistance in *T. brucei*	Reference
pbRn6 (modified)	Introduction of GFP reporter cassette downstream of the active VSG221. Used in this study to generate double-expresser and GFP expression cell lines.	Amp	Blas	Janzen et al., 2004
pLEW82v4	For introduction of an ectopic gene into the rDNA spacer region for overexpression under the control of a tetracycline-inducible T7 RNA polymerase promoter	Amp	Phleo	George Cross
pLEW100v5b1d	For introduction of an ectopic gene into the rDNA spacer region for overexpression under the control of a tetracycline-inducible rRNA polymerase promoter	Amp	Phleo	George Cross
pMOTag3H	For PCR-based HA C-terminal *in situ* tagging of genes in *T. brucei*	Amp	G418	Oberholzer et al., 2006
pMOTag3G	For PCR-based GFP C-terminal *in situ* tagging of genes in *T. brucei.* It was used in this study for generation of pMOTag3_mNG	Amp	G418	Oberholzer et al., 2006
pGL2084	For introducing RNAi target into a rDNA spacer containing an integrated	Amp	Hyg	Jones et al., 2014

	landing pad. Uses Gateway cloning and single locus targeting strategy			
pJET1.2	An intermediate vector for blunt end cloning of PCR products	Amp	-	Thermo Scientific

2.1.10 Organisms

Bacteria

Table 7. Bacterial strain used in cloning

Strain	Genotype	Application
Escherichia coli (*E. coli*) Top10	F-mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG	Molecular cloning and plasmid propagation

Trypanosomes

Table 8. Parental trypanosome cell lines

Name	Description	Selection	Reference
MITat1.2 13-90 (referred to as 13-90)	Monomorphic *Trypanosoma brucei brucei* wild-type cell line (expressing VSG221) transfected with pLew13 & pLew90 plasmids	5 µg/ml Hyg 2.5 µg/ml G418	(Wirtz et al., 1999)
MITat1.6 wild-type	Monomorphic *Trypanosoma brucei brucei* (Lister 427) wild-type cell line expressing VSG121	----------	George Cross, 1975
2T1	Monomorphic *Trypanosoma brucei brucei* wild-type cell line (expressing VSG221) transfected with the pHD1313 and ph3Ep plasmids	2.5 µg/ml Phleo 0.1 µg/ml Puro	(Alsford et al., 2005)
MITat1.5 wild-type	Monomorphic *Trypanosoma brucei brucei* (Lister 427) wild-type cell line expressing VSG118	----------	

Abbreviations: MITat1.2 (Molteno Institute Trypanozoon Antigen Type 1.2), MITat1.6 (Molteno Institute Trypanozoon Antigen Type 1.6), 2T1 (VSG221 expressing, Tagged, clone1), MITat1.5 (Molteno Institute Trypanozoon Antigen Type 1.5), BSF (bloodstream form), G418 (Geneticin), Phleo (Phleomycin), Hyg (Hygromycin), Puro (Puromycin).

2.2 Methods

2.2.1 Bacterial culture and plasmids

Cultivation of bacteria

Bacteria were cultivated either as liquid cultures in Luria Bertani (LB) medium or as solid culture on LB agar plates. LB medium was prepared by mixing pre-mixed LB-agar or LB-broth powder in an appropriate volume of distilled water. The mixture was autoclaved at 121 ˚C for 20 min, cooled down to about 55 ˚C and supplemented with 100 μg/ml ampicillin. For liquid cultures, the LB-medium with ampicillin was inoculated with bacteria and incubated at 37 ˚C on a shaker (150 – 200 rpm). Solid cultures were prepared by inoculating LB agar plates with bacteria and incubating the plates at 37 ˚C.

Preparation of chemically-competent bacteria

For the preparation of chemically competent *E. coli* Top10 cells, 100 ml of fresh LB-medium was inoculated with 1 ml bacteria suspension from an overnight culture. The culture was incubated at 37 ˚C on a shaker (150 – 200 rpm) until the optical density at 600 nm (OD600) was between 0.3 and 0.4. The bacteria were harvested by centrifugation (900x *g*, 10 min, 4 ˚C) and the pellet resuspended in 2.5 ml TSS (LB Medium containing 10 % PEG 3350, 5 % DMSO, 50 mM $MgCl_2$). Aliquots of the suspension were transferred into pre-cooled 1.5 ml microcentrifuge tubes and frozen in liquid nitrogen. Bacteria were stored at -80 ˚C until use.

Transformation of chemically competent bacteria

Chemically competent *E. coli* Top 10 cells were thawed on ice, mixed with either 10 ng of plasmid DNA or 10 μl of ligation mix and incubated on ice for 20 min. The bacterial cells were then heat-shocked at 42 ˚C for 50 s and cooled on ice for 2 min. The cells were then plated on LB agar plate containing 100 μg/ml ampicillin and incubated at 37 ˚C overnight.

Plasmid isolation from bacteria

For small amounts of plasmid DNA (usually with low purity) suitable for analytical restriction enzyme digest, plasmid DNA was extracted from 1 ml of an overnight culture using a standard alkaline lysis protocol. Higher quantities of plasmid DNA with high purity were isolated using the NucleoSpin Plasmid Kit or NucleoBond PC 100 Kit (Macherey-Nagel, DE), following the manufacturer's instructions.

2.2.2 Plasmid generation

Generation of plasmids

Table 9 contains a list of plasmids generated in this study. Restriction enzyme cloning was used for the generation of all plasmids except the RNAi constructs. For the RNAi constructs Gateway cloning involving homologous recombination using BP recombination reaction was employed. All oligonucleotides (listed in Table 8) were synthesised by Sigma Aldrich.

Generation of pbRn6M1.6.nPPT

To generate the plasmid pbRn6M1.6.nPPT, the *VSG121* ORF was amplified from plasmid pRS.121 (Batram et al., 2014) by PCR (with Phusion polymerase) using primers HZ30 and HZ31. The blunted PCR product was ligated into plasmid pJET1.2:Blas:GFP:UTR after removal of the *GFP* ORF with PacI and PaeI restriction enzymes. The fragment Blas:VSG121:UTR was then excised with BamHI and HindIII and ligated into a modified pbRn6 plasmid (Janzen et al., 2004) which was already linearised with the same enzymes. The upstream integration region of the resultant plasmid pbRn6M1.6-198 was then modified by extending the *VSG221* 3´ UTR sequence to include the native polypyrimidine tract. For this, a 452 bp sequence of part of the *VSG221* ORF and 3´ UTR was amplified from the pBSK.VSG221wt-full 3´ UTR plasmid using primers MBS7 and MBS8. The amplified fragment was cloned into pJET1.2, excised with BamHI and SacI, and then ligation into pbRn6.M1.6_198 digested with the same restriction enzymes. The resultant plasmid pbRn6M1.6.nPPT was linearised with SacI and SalI for transfection.

Generation of pbRn6M1.6Δ46-52.nPPT

The plasmid pbRn6M1.6Δ46-52.nPPT was generated in a similar manner to pbRn6M1.6.nPPT. Briefly, the *VSG121* ORF was amplified from plasmid pRS.121 as described above. The blunted PCR product was ligated into plasmid pJET1.2:Blas:GFP:Δ(46-52)UTR after removal of the *GFP* ORF with PacI and PaeI restriction enzymes. The fragment Blas:VSG121:Δ(46-52)UTR was then excised with BamHI and HindIII and ligated into a modified pbRn6 plasmid linearised with the same enzymes. The upstream integration region of the resultant plasmid was then modified by extending the *VSG221* 3´ UTR sequence to include the native polypyrimidine tract as described above. The resultant plasmid pbRn6M1.6Δ46-52.nPPT was linearised with SacI and SalI for transfection.

Generation of pbRn6M1.6TGA(46-48)ACT.nPPT

To generate the plasmid pbRn6M1.6TGA(46-48)ACT.nPPT, a mutagenesis PCR was carried out using the fusion PCR approach. Briefly, primer pairs C14/HZ35 and C15/HZ32 were used to amplify two PCR fragments from the template pbRn6M1.6_198. The two fragments were joined together in an additional PCR step using the primer pair HZ35/HZ32. The resultant PCR product was cloned into pJET1.2 to yield pJET1.2M1.6TGA(46-48)ACT. The M1.6TGA(46-48)ACT fragment was then excised with HindIII and MVa1269I and ligated into the pbRn6 backbone obtained from digestion of pbRn6GFPΔ46-52 with HindIII and MVa1269I. The upstream integration region of the resultant plasmid was then modified by extending the *VSG221* 3´ UTR sequence to include the native polypyrimidine tract as described above. The resultant plasmid pbRn6M1.6TGA(46-48)ACT.nPPT was linearised with SacI and SalI for transfection.

Generation of pbRn6M1.6Δ45-55.nPPT

For generation of the plasmid pbRn6M1.6Δ45-55.nPPT, mutagenesis PCR was carried out using the fusion PCR approach. The primer pairs HZ50/HZ35 and HZ49/HZ32 were used to amplify two PCR fragments from the template pbRn6M1.6Δ46-52. The two fragments were joined together in an additional PCR step using the primer pair HZ35/HZ32. The resultant PCR product was cloned into pJET1.2 to yield pJET1.2M1.6Δ45-55. The M1.6Δ45-55 fragment was then excised with HindIII and MVa1269I and ligated into the pbRn6 backbone linearised with HindIII and MVa1269I. The upstream integration region of the resultant plasmid was then modified by extending the *VSG221* 3´ UTR sequence to include the native polypyrimidine tract as described above. The resultant plasmid pbRn6M1.6Δ45-55.nPPT was linearised with SacI and SalI for transfection.

Generation of pbRn6M1.6Inver.8mer.nPPT

For generation of the plasmid pbRn6M1.6Inver.8mer.nPPT, a fragment of VSG121 with a 3´ UTR harbouring an inverted 8mer was excised from pES-M1.6_inv(28-35) (Batram, Master thesis 2009) using the restriction enzymes BamHI and NcoI and ligated into a pbRn6 vector backbone. The fragment was excised sequentially by first digesting with BamHI followed by blunting with Klenow fragment and then further digestion with NcoI. The excised fragment was then ligated into a pbRn6 backbone obtained from pbRn6M1.6_198 digested with HindIII (blunted) and then with NcoI. The upstream integration region of the resultant plasmid was then modified by extending the *VSG221* 3´ UTR sequence to include the native polypyrimidine tract

as described above. The resultant plasmid pbRn6M1.6Inver.8mer.nPPT was linearised with SacI and SalI for transfection.

Generation of pbRn6GFP.nPPT and pbRn6GFPΔ46-52.nPPT

For generation of plasmids pbRn6GFP.nPPT and pbRn6GFPΔ46-52.nPPT, *VSG121* ORF with the wild-type 3′ UTR or a 3′ UTR harbouring the mutation Δ46-52 within the 16mer motif were amplified from plasmids pRS.121 and pES-M1.6Δ(46-52), respectively, using the primer pair C80/C81 flanked by PacI and BamHI overhangs. The amplified fragments were cloned individually into pJet1.2 and then excised with PacI and BamHI followed by ligation into p3845 (M Carrington) linearised with the same enzymes to generate plasmids p3845.VSG1213′ UTR and p3845.VSG121Δ(46-52)3′ UTR. The *GFP* ORF coupled to VSG121 3′ UTR and VSG121Δ(46-52)3′ UTR fragments were amplified from p3845.VSG1213′ UTR and p3845.VSG121Δ(46-52)3′ UTR with primers flanked with BamHI and HindIII restriction sites. The amplified fragments were cloned into pJET1.2 and the Blas:GFP:UTR and Blas:GFP: Δ(46-52)UTR fragments excised with the restriction enzymes BamHI/HindIII and ligated into the pbRn6 plasmid backbone linearised with the same enzymes. The upstream integration regions of the resultant plasmids were then modified by extending the *VSG221* 3′ UTR sequence to include the native polypyrimidine tract as described above, giving rise to the final plasmids pbRn6GFP.nPPT and pbRn6GFPΔ46-52.nPP. The resultant plasmids were linearised with SacI and SalI for transfections.

Generation of pJET1.2_M1.2-Blas KO

pJET1.2_M1.2-Blas KO was generated by joining three PCR-amplified fragments using a fusion PCR approach. Fragment 1 (part of VSG221 ORF) was amplified from plasmid NJ199 using the primer pair MBS11/MBS12. Fragment 2 (puromycin resistance gene flanked by actin UTRs) was amplified from plasmid CJ44-Puro with the primer pair MBS17/MBS18. Fragment 3 (part of blasticidin resistance gene ORF) was amplified from plasmid pbRn6M1.6TGA(46-48)ACT.nPPT using the primer pair MBS13/MBS14. Fragments 1 and 2 were fused together in a PCR using the primer pair MBS11/MBS18. The PCR product obtained (VSG221-Puro) was then fused to fragment 3 using the primer pair MBS11/MBS14. The resultant PCR product was cloned into pJET1.2 to yield the plasmid pJET1.2_M1.2-Blas KO. The plasmid was linearised with HindIII for transfections.

Generation of pLEW82v4_EP1-eYFP+full 3′ UTR

To generate the above construct, a fusion PCR approach was used to join two PCR amplified fragments from the plasmid pGAPRONEΔLII_EP1::eYFP. Fragment 1 (EP1:eYFP with part

of the EP1 3´ UTR, flanked by HindIII and the LII region) was amplified with the primer pair MBS59/MBS60. Fragment 2 (part of the EP1 3´ UTR flanked by XhoI and LII region) was amplified using the primer pair MBS61/MBS62. The two fragments were joined in a PCR using the primer pair MBS59/MBS62. The PCR product obtained (EP1::eYFP_full3´ UTR) was cloned into pJET1.2 to obtain the plasmid pJET1.2_EP1::eYFP_full3´ UTR. The EP1::eYFP_full3´ UTR fragment was then excised from pJET1.2_EP1::eYFP_full3´ UTR using HindIII and XhoI and ligated into the pLEW82v4 vector linearised with the same restriction enzymes. The resultant plasmid (pLEW82v4_EP1-eYFP+full 3'UTR) was linearised with NotI for transfections.

Generation of pLEW82v4_M1.6N46-48

pLEW82v4_M1.6N46-48 was generated using a fusion PCR approach. Two PCR amplified fragments from pbRn6M1.6TGA(46-48)ACT.nPPT were joined together. Fragment 1 (VSG121 ORF and part of the 3´ UTR) was amplified using primers MBS71/MBS75. Fragment 2 (part of the 3´ UTR) was amplified using primers MBS72/MBS76. The two fragments were fused together in a PCR using primers MBS71/MBS72. The PCR product obtained (M1.6N46-48) was cloned into pJET1.2 to obtain the plasmid pJET1.2_M1.6N46-48. The M1.6N46-48 fragment was then excised from pJET1.2_M1.6N46-48 using HindIII and XhoI and ligated into the pLEW82v4 vector linearised with the same restriction enzymes. The resultant plasmid (pLEW82v4_ M1.6N46-48) was linearised with NotI for transfection.

Generation of pLEW82v4_M1.5N46-48

To generate pLEW82v4_M1.5N46-48, mutagenesis PCR by a fusion PCR approach was used to introduce the mutation TGA(46-48)ACT into the VSG118 3´ UTR using the plasmid NJ130 as template. Fragments 1 and 2 were amplified using the primer pairs MBS85/MBS86 and MBS87/MBS88, respectively. The two fragments were then fused with primers MBS85/MBS88. The PCR product obtained (M1.5N46-48) was cloned in pJET1.2 to obtain the plasmid pJET1.2_M1.5N46-48. The M1.6N46-48 fragment was then excised from pJET1.2_M1.5N46-48 using HindIII and XhoI and ligated into pLEW82v4 vector linearised with the same restriction enzymes. The resultant plasmid (pLEW82v4_ M1.5N46-48) was linearised with NotI for transfection.

Generation of pMOTag3_mNG

pMOTag3_mNG was generated by modification of pMOTag3G (Oberholzer et al., 2006). Briefly, the mNeonGreen ORF flanked by XhoI and SalI restriction enzymes cleavage sites was amplified from plasmid pTSARib(hygro)-mNeonGreen using primers IN1 and IN2 The amplified fragment was cloned into pJET1.2, excised with XhoI and SalI and ligated into the pMOTag vector backbone obtained by digesting pMOTag3G with the same enzymes. The resultant plasmid pMOTag3_mNG was used as a template for PCR-based C-terminal endogenous tagging of proteins.

Generation of pLEW100v5b1d_DHwt

The plasmid pLEW100v5b1d_DHwt was generated by cloning the Hel66 ORF amplified from *T. brucei* 13-90 BSF genomic DNA into pLEW100v5b1d using the HindIII and BamHI restriction enzyme sites. A silent mutation was introduced in the Hel66 ORF to remove an internal HindIII restriction enzyme cleavage site using a fusion PCR approach. Fragment 1, amplified using primers MBS77/MBS79 and fragment 2, amplified with MBS78/MBS80 were joined together in a PCR using primers MBS77/MBS78. The primer MBS77 is flanked by a HindIII restriction site whereas MBS78 is flanked by a 3xHA tag and a BamHI restriction enzyme site. The resultant PCR product (C-terminal HA-tagged Hel66 flanked by BamHI and HindIII restriction enzyme sites) was cloned into the pJET1.2 plasmid, excised with HindIII and BamHI and ligated into pLEW100v5b1d linearised with the same enzymes. The resultant plasmid (pLEW100v5b1d_DHwt) was linearised with NotI for transfection.

Generation of pLEW100v5b1d_DHatp

pLEW100v5b1d_DHatp was generated using a similar approach as described for pLEW100v5b1d_DHwt. The mutation E254Q was introduced in the Hel66 ORF amplified from pLEW100v5b1d_DHwt using the fusion PCR approach. Fragment 1 was amplified using primers MBS77/MBS81 and fragment 2 amplified with primers MBS78/MBS82. The two fragments were fused using primers MBS77/78. Subsequent steps were carried out as described for pLEW100v5b1d_DHwt. The resultant plasmid (pLEW100v5b1d_DHatp) was linearised with NotI for transfection.

Generation of RNAi constructs

To generate the RNAi constructs, the vector pGL2084 (Jones et al., 2014) was used. Target sequences for the RNAi were amplified from genomic DNA with primers carrying AttB sequences at their 5´ ends. The primers were designed to amplify between 300 bp and 500 bp of the target gene ORF sequences. The primers and target fragments were checked for

specificity to the individual genes using the Basic Local Alignment Search Tool (BLAST). In the case of pGL2084_DNAJ and pGL2084_MiRF172, primers for the target fragments were selected using RNAit (Redmond et al., 2003). The PCR-amplified fragments were then integrated into the pGL2084 vector via homologous recombination using BP recombination reactions (ThermoFisher Scientific). The resultant plasmids were linearised with AscI and used to transfect 2T1 BSF cells to generate the RNAi cell lines. The primer pairs used for each of the constructs are as follows: pGL2084_PI4P5K; MBS21/MBS22, pGL2084_Thimet OP; MBS23a/MBS24a, pGL2084_DNAJ; MBS25a/MBS26a, pGL2084_LRR; MBS27/MBS28, pGL2084_RPC40; MBS29/MBS30, pGL2084_MiRF172; MBS31a/MBS32a, pGL2084_NLP; MBS42/MBS43.

For pGL2084_Hel66, due to the conserved nature of the DExD/H box helicases sequences, fragments encoding the non-conserved N- and C- terminal extensions were fused to obtain a unique long target sequence. The primer pair MBS37/MBS38 was used for amplification of fragment 1 and MBS39/MBS40 was used for amplification of fragment 2. The two fragments were fused together by PCR using the primers MBS37/MBS40.

Table 9. List of plasmids used in this study for generation of transgenic trypanosomes

Construct name	Description/Purpose	Resistance in *E. coli*	Resistance in *T. brucei*	Made by
pbRn6M1.6.nPPT	Constitutive expression of VSG121 with wild-type VSG121 3′ UTR downstream of endogenous VSG221	Amp	Blas	MBS & HZ
pbRn6 M1.6Δ46-52.nPPT	Constitutive expression of VSG121 with a mutated VSG121 3′ UTR (deletion of nucleotides 46-52 in the 16mer motif) downstream of endogenous VSG221	Amp	Blas	MBS & HZ
pbRn6M1.6Δ45-55.nPPT	Constitutive expression of VSG121 with a mutated VSG121 3′ UTR (deletion of nucleotides 45-55 in the 16mer motif) downstream of endogenous VSG221	Amp	Blas	MBS & HZ
pbRn6 M1.6 Invert 8mer.nPPT	Constitutive expression of VSG121 with a mutated VSG121 3′ UTR (inversion of the 8mer motif) downstream of endogenous VSG221	Amp	Blas	MBS & HZ
pbRn6 M1.6 TGA(46-48)ACT.nPPT	Constitutive expression of VSG121 with a mutated VSG121 3′ UTR (substitution of nucleotides 46-48(TGA) with ACT) downstream of endogenous VSG221	Amp	Blas	MBS & HZ
pbRn6GFP.nPPT	Constitutive expression of GFP with wild-type VSG121 3′ UTR downstream of endogenous VSG221	Amp	Blas	MBS & HZ
pbRn6GFPΔ46-52.nPPT	Constitutive expression of GFP with a mutated VSG121 3′ UTR (deletion of nucleotides 45-51 in the 16mer motif) downstream of endogenous VSG221	Amp	Blas	MBS & HZ

pJET1.2_M1.2-Blas KO	For deletion of VSG221 in a double-expresser cell line expressing an ectopic VSG downstream of the endogenous VSG221	Amp	Puro	MBS
pLEW82v4.VSG221wt_full3'UTR	Targets VSG221 to the rDNA spacer region for overexpression under the control of a tetracycline-inducible T7 RNA polymerase promoter	Amp	Phleo	NJ
pLEW82v4_EP1-eYFP+full 3'UTR	Targets EP1-eYFP to the rDNA spacer region for overexpression under the control of a tetracycline-inducible T7 RNA polymerase promoter	Amp	Phleo	MBS
pLEW82v4_M1.6N46-48	Targets VSG121 with a mutated VSG121 3´ UTR (substitution of nucleotides 46-48 with ACT) to the rDNA spacer region for overexpression under the control of a tetracycline-inducible T7 RNA polymerase promoter	Amp	Phleo	MBS
pLEW82v4_M1.5N46-48	Targets VSG121 with a mutated VSG121 3´ UTR (substitution of nucleotides 46-48 with ACT) to the rDNA spacer region for overexpression under the control of a tetracycline-inducible T7 RNA polymerase promoter	Amp	Phleo	MBS
pGL2084_NLP	NLP stem-loop RNAi construct that integrates at a rDNA spacer containing an integrated landing pad	Amp	Hyg	MBS
pGL2084_LRR	LRR stem-loop RNAi construct that integrates at a rDNA spacer	Amp	Hyg	MBS

	containing an integrated landing pad			
pGL2084_MiRF172	MiRF172 stem-loop RNAi construct that integrates at a rDNA spacer containing an integrated landing pad	Amp	Hyg	MBS
pGL2084_DNAJ	DNAJ stem-loop RNAi construct that integrates at a rDNA spacer containing an integrated landing pad	Amp	Hyg	MBS
pGL2084_RPC40	RPC40 stem-loop RNAi construct that integrates at a rDNA spacer containing an integrated landing pad	Amp	Hyg	MBS
pGL2084_PI4P5K	PI4P5K stem-loop RNAi construct that integrates at a rDNA spacer containing an integrated landing pad	Amp	Hyg	MBS
pGL2084_Thimet OP	Thimet OP stem-loop RNAi construct that integrates at a rDNA spacer containing an integrated landing pad	Amp	Hyg	MBS
pGL2084_DEAD HEL	Hel66 stem-loop RNAi construct that integrates at a rDNA spacer containing an integrated landing pad	Amp	Hyg	MBS
pLEW100v5b1d_Hel66$_{WT}$	Targets wild-type Hel66 ORF to the rDNA spacer region for overexpression under the control of a tetracycline-inducible rRNA polymerase promoter	Amp	Phleo	MBS
pLEW100v5b1d_Hel66$_{ATP}$	Targets ATPase mutant Hel66 ORF to the rDNA spacer region for overexpression under the control of a tetracycline-inducible rRNA polymerase promoter	Amp	Phleo	MBS

Abbreviations: Amp (Ampicillin), Blas (Blasticidin), Puro (Puromycin), Phleo (Phleomycin), Hyg (Hygromycin), HZ (Henriette Zimmermann), MBS (Majeed Bakari Soale), NJ (Nicola Jones).

2.2.3 Trypanosome culture and analysis

Cultivation of monomorphic trypanosomes

Monomorphic BSF trypanosomes were cultivated in HMI-9 medium (Hirumi & Hirumi, 1989) supplemented with 10 % foetal calf serum (FCS) and incubated at 37 ˚C and 5 % CO_2. Cell densities were kept below 1 x 10^6 cells/ml to keep cells within the logarithmic growth phase. For this reason, the density of the trypanosome cells was checked regularly using a haemocytometer (Neubauer chamber) or Particle Coulter Counter Z2 (Beckman Coulter). The cultures were diluted with pre-warmed HMI-9 medium with appropriate antibiotics added (when necessary) for maintenance of transgenic cell lines. Harvesting of cells was carried out by centrifugation at 1,400x *g* for 10 min at 4 ˚C unless otherwise stated.

Preparation and thawing of cryo-stabilates

For a single cryo-stabilate, 4 x 10^6 cells were prepared. Briefly, 4 x 10^6 cells were harvested by centrifugation (1,400x *g* for 10 min at 4 ˚C) and the pellet resuspended in 500 µl of pre-cooled HMI-9 medium. An equal volume of pre-cooled 2x freezing mix was added and the mixture transferred into a cryotube and stored at -80 ˚C. For long-term storage (more than 1 year), stabilates were later transferred to -150 ˚C.

To revive frozen cells, stabilates were rapidly thawed in a 37 ˚C water bath. Immediately after thawing, cells were transferred into 10 ml ice-cold HMI-9 and washed by centrifugation (1,400x *g* for 10 min at 4 ˚C). The supernatant was discarded and the cells resuspended in 10 ml of pre-warmed HMI-9 medium and transferred into a culture flask. After 1 h of cultivation, the cell density was checked, the culture diluted and appropriate antibiotics added if necessary.

Generation of transgenic trypanosomes

Stable transgenic trypanosome cell lines were generated using the Amaxa Nucleofector II and the Amaxa Basic Parasite Nucleofector Kit 1 (Lonza, CH). For transfections, 3 x 10^7 BSF cells were harvested by centrifugation (1,400x *g* for 10 min at 37 ˚C). The cells were carefully resuspended in 100 µl Basic Parasite Solution I. 10 µg of linearised plasmid DNA or PCR product was then added to the cells and gently mixed. The mixture was transferred into a cuvette and electroporated using the Amaxa Nucleofector II with the pre-set program X-001. Immediately after electroporation, the cells were transferred into 30 ml pre-warmed HMI-9 medium (37 ˚C) containing antibiotics for maintenance of the parental cell line, and serially diluted (1:10, 1:50 and 1:100). Each dilution was plated in a 24-well plate (1 ml per well). The cells were then cultivated at 37 ˚C and 5 % CO_2 for 6 h or overnight (~ 16 h). After this initial incubation, 1 ml of HMI-9 medium containing antibiotics for the maintenance of the parental

cell line and 2x selection antibiotics for the transfected construct was added to each well. Isogenic clones were obtained after five to six days of incubation. A list of the transgenic trypanosomes generated in this study is shown in Table 10.

Differentiation of monomorphic bloodstream form trypanosomes to procyclic forms

Monomorphic BSF trypanosomes were grown to a density of 1.5 x 10^6 cells/ml in HMI-9 medium. 2.5 x 10^7 cells were harvested by centrifugation (1,400x *g* for 10 min at room temperature (RT)) and resuspended in 5 ml of Differentiating Trypanosome Medium (DTM) (Overath et al., 1986) containing 6 mM cis-aconitate. The parasites were then cultivated at 37 ˚C and 5 % CO_2. The cells were diluted with SDM-79 medium supplemented with 10 % heat- inactivated FCS when the cell density was above 2 x 10^7 cells/ml.

Table 10. List of transgenic trypanosomes generated in this study

Name	Description	Constructs	Selection
$221^{ES}121_{WT}$	This cell line is based on MITat1.2 13-90. VSG121 with a wild-type VSG121 3′ UTR was integrated downstream of the endogenous VSG221	pbRn6M1.6.nPPT	5 µg/ml Hyg, 2.5 µg/ml G418, 5 µg/ml Blas
$221^{ES}121_{\Delta 46-52}$	This cell line is based on MITat1.2 13-90. VSG121 with a mutated VSG121 3′ UTR (deletion of nucleotides 45-51 in the 16mer motif) was integrated downstream of the endogenous VSG221	pbRn6M1.6Δ46-52.nPPT	5 µg/ml Hyg, 2.5 µg/ml G418, 5 µg/ml Blas
$221^{ES}121_{\Delta 45-55}$	This cell line is based on MITat1.2 13-90. VSG121 with a mutated VSG121 3′ UTR (deletion of nucleotides 45-55 in the 16mer motif) was integrated downstream of the endogenous VSG221	pbRn6M1.6Δ45-55.nPPT	5 µg/ml Hyg, 2.5 µg/ml G418, 5 µg/ml Blas
$221^{ES}121_{N46-48}$	This cell line is based on MITat1.2 13-90. VSG121 with a mutated VSG121 3′ UTR (substitution of nucleotides 46-48 with ACT) was integrated downstream of the endogenous VSG221	pbRn6M1.6TGA(46-48)ACT.nPPT	5 µg/ml Hyg, 2.5 µg/ml G418, 5 µg/ml Blas
$221^{ES}121_{Inver.8mer}$	This cell line is based on MITat1.2 13-90. VSG121 with a mutated VSG121 3′ UTR (inversion of the 8mer motif) was integrated downstream of the endogenous VSG221	pbRn6M1.6Inver.8mer.nPPT	5 µg/ml Hyg, 2.5 µg/ml G418, 5 µg/ml Blas
$221^{ES}GFP_{WT}$	This cell line is based on MITat1.2 13-90. GFP with a wild-type VSG121 3′ UTR was integrated downstream of the endogenous VSG221	pbRn6GFP.nPPT	5 µg/ml Hyg, 2.5 µg/ml G418, 5 µg/ml Blas
$221^{ES}GFP_{\Delta 45-51}$	This cell line is based on MITat1.2 13-90. GFP with a mutated VSG121 3′ UTR (deletion of nucleotides 45-51 in the 16mer motif) was	pbRn6GFPΔ45-51.nPPT	5 µg/ml Hyg, 2.5 µg/ml G418, 5 µg/ml Blas

	integrated downstream of the endogenous VSG221		
$\Delta 221^{ES}121_{WT}$	This cell line is based on $221^{ES}121WT^{ES}$. The endogenous VSG221 including the VSG221 3´ UTR was deleted by replacement with a puromycin resistance gene	pJET1.2_M1.2-Blas KO	5 µg/ml Hyg, 2.5 µg/ml G418, 0.1 µg/ml Puro
$\Delta 221^{ES}121_{N46-48}$	This cell line is based on $221^{ES}121TGA(46-48)ACT^{ES}$. The endogenous VSG221 including the VSG221 3´ UTR was deleted by replacement with a puromycin resistance gene	pJET1.2_M1.2-Blas KO	5 µg/ml Hyg, 2.5 µg/ml G418, 0.1 µg/ml Puro
$\Delta 221^{ES}121_{WT}{}^{Tet}$	This cell line is based on $\Delta 221^{ES}121WT^{ES}$. VSG221 was introduced into the rDNA spacer region under the control of a tetracycline-inducible T7 RNA polymerase promotor	pLEW82v4.VSG221wt_full3'UTR	5 µg/ml Hyg, 2.5 µg/ml G418, 0.1 µg/ml Puro, 1 µg/ml Phleo
$\Delta 221^{ES}121_{N46-48}{}^{Tet}$	This cell line is based on $\Delta 221^{ES}121TGA(46-48)ACT^{ES}$. VSG221 was introduced into the rDNA spacer region under the control of a tetracycline-inducible T7 RNA polymerase promotor	pLEW82v4.VSG221wt_full 3'UTR	5 µg/ml Hyg, 2.5 µg/ml G418, 0.1 µg/ml Puro, 1 µg/ml Phleo
$\Delta 221^{ES}121_{WT}EP1^{Tet}$	This cell line is based on $\Delta 221^{ES}121WT^{ES}$. EP1-eYFP with the wild-type EP1 3´ UTR was introduced into the rDNA spacer region under the control of a tetracycline-inducible T7 RNA polymerase promotor	pLEW82v4_EP1-eYFP+full 3'UTR	5 µg/ml Hyg, 2.5 µg/ml G418, 0.1 µg/ml Puro, 1 µg/ml Phleo
$\Delta 221^{ES}121_{N46-48}EP1^{Tet}$	This cell line is based on $\Delta 221^{ES}121TGA(46-48)ACT^{ES}$. EP1-eYFP with the wild-type EP1 3´ UTR introduced into the rDNA spacer region under the control of a tetracycline-inducible T7 RNA polymerase promotor	pLEW82v4_EP1-eYFP+full 3'UTR	5 µg/ml Hyg, 2.5 µg/ml G418, 0.1 µg/ml Puro, 1 µg/ml Phleo

$221^{ES}121_{N46-48}{}^{Tet}$	This cell line is based on MITat1.2 13-90. VSG121 with a mutated VSG121 3´ UTR (substitution of nucleotides 46-48 with ACT) was introduced into the rDNA spacer region under the control of a tetracycline-inducible T7 RNA polymerase promotor	pLEW82v4_M1.6N 46-48	5 µg/ml Hyg, 2.5 µg/ml G418, 1 µg/ml Phleo
$221^{ES}118_{N46-48}{}^{Tet}$	This cell line is based on MITat1.2 13-90. VSG118 with a mutated VSG121 3´ UTR (substitution of nucleotides 46-48 with ACT) was introduced into the rDNA spacer region under the control of a tetracycline-inducible T7 RNA polymerase promotor	pLEW82v4_M1.5N 46-48	5 µg/ml Hyg, 2.5 µg/ml G418, 1 µg/ml Phleo
NLP-RNAi	This cell line is based on 2T1 cell line. RNAi target sequences of NLP (Tb927.11.13080) was integrated in a single tagged locus of the non-transcribed rRNA under a tetracycline inducible rRNA promotor	pGL2084_NLP	2.5 µg/ml Hyg, 2.5 µg/ml Phleo
LRR-RNAi	This cell line is based on 2T1 cell line. RNAi target sequences of LRR (Tb927.11.1960) was integrated in a single tagged locus of the non-transcribed rRNA under a tetracycline inducible rRNA promotor	pGL2084_LRR	2.5 µg/ml Hyg, 2.5 µg/ml Phleo
MiRF172-RNAi	This cell line is based on 2T1 cell line. RNAi target sequences of MiRF172 (Tb927.3.2050) was integrated in a single tagged locus of the non-transcribed rRNA under a tetracycline inducible rRNA promotor	pGL2084_MiRF172	2.5 µg/ml Hyg, 2.5 µg/ml Phleo
DNAJ-RNAi	This cell line is based on 2T1 cell line. RNAi target sequences of DNAJ (Tb927.10.2290) was	pGL2084_DNAJ	2.5 µg/ml Hyg, 2.5 µg/ml Phleo

	integrated in a single tagged locus of the non-transcribed rRNA under a tetracycline inducible rRNA promotor		
RPC40-RNAi	This cell line is based on 2T1 cell line. RNAi target sequences of RPC40 (Tb927.10.15370) was integrated in a single tagged locus of the non-transcribed rRNA under a tetracycline inducible rRNA promotor	pGL2084_RNAPol	2.5 µg/ml Hyg, 2.5 µg/ml Phleo
PI4P5K-RNAi	This cell line is based on 2T1 cell line. RNAi target sequences of PI4P5K (Tb927.10.4770) was integrated in a single tagged locus of the non-transcribed rRNA under a tetracycline inducible rRNA promotor	pGL2084_PI4P5K	2.5 µg/ml Hyg, 2.5 µg/ml Phleo
Thimet OP-RNAi	This cell line is based on 2T1 cell line. RNAi target sequences of TOP (Tb927.7.190) was integrated in a single tagged locus of the non-transcribed rRNA under a tetracycline inducible rRNA promotor	pGL2084_Thimet OP	2.5 µg/ml Hyg, 2.5 µg/ml Phleo
Hel66-RNAi	This cell line is based on 2T1 cell line. RNAi target sequences of Hel66 (Tb927.10.1720) was integrated in a single tagged locus of the non-transcribed rRNA under a tetracycline inducible rRNA promotor	pGL2084_DEAD HEL	2.5 µg/ml Hyg, 2.5 µg/ml Phleo
Hel66::HA-RNAi	This cell line is based on Hel66-RNAi. Hel66 was endogenously tagged at the C-terminus with hemagglutinin (HA)	PCR product from pMOTag3H	2.5 µg/ml Hyg, 2.5 µg/ml Phleo, 3.0 µg/ml G418
Hel66::mNeonGreen	This cell line is based on 2T1 cell line. Hel66 was endogenously tagged at the	PCR product from pMOTag3_mNG	2.5 µg/ml Phleo, 0.1 µg/ml Puro,

	C-terminus with hemagglutinin (HA)		3.0 µg/ml G418
$Hel66_{WT}^{Tet}$	This cell line is based on MITat1.2 13-90. Hel66 tagged at the C-terminus with HA was introduced into the rDNA spacer region under the control of a tetracycline-inducible rRNA promotor	pLEW100v5b1d_Hel66$_{WT}$	5 µg/ml Hyg, 2.5 µg/ml G418, 1 µg/ml Phleo
$Hel66_{E254Q}^{Tet}$	This cell line is based on MITat1.2 13-90. ATPase mutant (E254Q) Hel66 tagged at the C-terminus with HA was introduced into the rDNA spacer region under the control of a tetracycline-inducible T7 RNA polymerase promotor	pLEW100v5b1d_Hel66$_{ATP}$	5 µg/ml Hyg, 2.5 µg/ml G418, 1 µg/ml Phleo

Abbreviations: G418 (Geneticin), Blas (Blasticidin), Puro (Puromycin), Phleo (Phleomycin), Hyg (Hygromycin).

2.2.4 Molecular and biochemical methods

2.2.4.1 DNA analysis

Isolation of genomic DNA from trypanosomes

Genomic DNA was isolated from trypanosome cells using the High Pure PCR Template Preparation kit (Roche). The protocol for isolation of nucleic acids from mammalian whole blood, buffy coat or cultured cells was followed with slight modifications. Briefly, 2 x 10^7 cells were harvested (1,400x *g* for 10 min at RT). The cells were resuspended in 200 µl of Trypanosome Dilution Buffer (TDB) and transferred into a nuclease-free 1.5 ml microcentrifuge tube. 200 µl of binding buffer and 40 µl proteinase K were then added. The contents of the tube were mixed immediately and incubated at 70 ˚C for 10 min. Afterwards, 100 µl of isopropanol was added, mixed and the sample transferred onto the high filter column placed in a collection tube. Genomic DNA was bound to the filter column by centrifugation (8000x *g*, 1 min, RT) and washed with 500 µl of Inhibitor Removal Buffer (8000x *g*, 1 min, RT). This was followed by a second wash with 500 µl of wash Buffer. Residual wash buffer was removed from the filter column by centrifugation (20,000x *g*, 10 s, RT). Genomic DNA was eluted with 100 µl of pre-warmed elution buffer (8000x *g*, 1 min, RT). The DNA was stored at -20 ˚C.

Measurement of nucleic acid concentration

DNA and RNA concentration measurements were carried out using the Tecan Infinite M200 reader. The measurements were based on absorbance of the samples at 260 nm and 280 nm. The ratio of the absorbance at 260 nm to the absorbance at 280 nm provides information on the purity of the DNA and RNA samples.

Polymerase chain reaction (PCR)

Amplification of genomic or plasmid DNA was performed by PCR using the Phusion High-Fidelity DNA Polymerase (ThermoFisher Scientific, USA). 10 ng of plasmid DNA or up to 1 μg of genomic DNA was used as template. A standard 50 μl reaction mix contained 0.5 – 1x HF buffer, 0.2 μM forward primer, 0.2 μM reverse primers, 1 U Phusion polymerase, 100 μM dNTP mix, template DNA and Milli-Q water. The thermocycling parameters used are as follows: 98 ˚C/30s - (98 ˚C/10 s – 55 ˚C/30s – 72 ˚C/30s) 30 cycles – 72 ˚C/8 min. The annealing temperature was altered depending on the melting temperatures of the primers. The duration of elongation was also adjusted based on the length of the amplicon.

Fusion PCR as described by Heckman & Pease (2007) was used for joining two DNA fragments to form a fusion product. This usually involves a 2 step PCR. For this, four primers (two each for the amplification of DNA fragments in the first PCR) were designed. The reverse primer for fragment 1 and the forward primer for fragment 2 were designed to have overlapping homologous sequences. DNA fragments from the first PCRs were purified and equimolar amounts of the two amplicons used as template for the second PCR. For the second PCR, the forward primer for amplification of fragment 1 and the reverse primer for amplification of fragment 2 were used to join the two fragments forming the fusion product.

Integration PCR on transgenic trypanosomes was performed using the Phusion Human Specimen Direct PCR Kit (ThermoFisher Scientific). Briefly, 1 x 10^6 cells were harvested by centrifugation (1400x *g*, 10 min, 4 ˚C). After discarding the supernatant, the cell pellet was resuspended in 20 μl dilution buffer supplemented with 0.5 μl of DNA release additive. The cell suspension was mixed well and incubated at RT for 5 min followed by further incubation at 98 ˚C for 2 min. The released DNA was separated from cell debris by centrifugation (2000x *g*, 5 min, 4 ˚C), and 2 μl of the supernatant was used as template in a 25 μl or 50 μl PCR reaction.

RT-PCR

For RT-PCR, cDNA was prepared from Total RNA (refer to section on RNA analysis) and a standard PCR (as described above) carried out on the cDNA samples. PCR amplification of the cDNA was carried out using a forward primer complementary to the sliced leader sequence and a gene specific reverse primer. Since the RT-PCR was performed on RNA samples from cells that had specific proteins depleted via RNAi, a control PCR was set up with a reverse primer complementary to α-tubulin.

Purification of PCR product

PCR products were purified directly using the NucleoSpin® Gel and PCR clean up kit (Macherey-Nagel) in accordance with the manufacturer's protocol.

Agarose gel electrophoresis

DNA fragments were separated after PCR or restriction enzyme digest using agarose gel electrophoresis. Samples were mixed with DNA loading dye (2X) prior to loading onto a 0.8 % agarose gel. GeneRuler DNA ladder mix (ThermoFisher Scientific, USA) was used as a size marker. Electrophoresis was carried out at 120 V for 30 min and the gel visualised under UV light using the INTAS Gel iX imager (INTAS, Göttingen) after staining in 3 µg/ml ethidium bromide solution for 15 min. Alternatively, samples were loaded onto a 0.8 % gel pre-stained with SYBR™ Safe DNA gel stain (Invitrogen, USA). The pre-stained gels were visualised directly after electrophoresis using the iBright CL1000 imaging system (ThermoFisher Scientific, USA).

Extraction of DNA from agarose gels

After separation of DNA fragments by agarose gel electrophoresis, fragments of interest were excised using a scalpel and purified from the gel slice using the NucleoSpin® Gel and PCR clean up kit (Macherey-Nagel) following the manufacturer's instructions.

Restriction enzyme digestion

Restriction enzyme digest was used to linearise plasmid DNA, confirm the presence of fragments of interest in a plasmid and generate suitable overhangs for cloning. All restriction enzyme digests were carried out using enzymes from ThermoFisher Scientific essentially following the manufacturer's instructions.

Ligation reaction

Ligation reactions were carried out using T4 DNA ligase (Thermo Fisher Scientific) following the manufacturer's instructions. A molar ratio of 1:3 (vector: insert) was used for vector and insert ligations. The ligation reaction mix containing the required amount of vector and insert were incubated at room temperature for at least 1h.

Isopropanol precipitation

Isopropanol precipitations were carried out on linearised plasmid DNA and PCR products used for transfections. Briefly, 1 volume isopropanol and 0.1 volume sodium acetate (pH= 7.0) were added to the DNA samples and mixed well by gently inverting the tube. The samples were then centrifuged (20000x *g*, 30 min, 4 ˚C) and the supernatant discarded. The pellet was washed twice with 1 ml ice-cold 70 % ethanol (20,000 *g*, 10 min, 4 ˚C). After washing, the supernatant was removed and the pellet air-dried under sterile conditions (in a laminar flow hood). The dried DNA pellet was then dissolved in sterile nuclease-free water.

2.2.4.2 RNA analysis

Isolation of total RNA from trypanosomes

Total RNA was isolated from 1 x 10^8 trypanosome cells harvested by centrifugation (1400x *g*, 10 min, 4 ˚C). The cell pellet was resuspend in 1 ml ice-cold serum-free HMI-9 medium and transferred into an RNAse free microcentrifuge tube. The cells were centrifuged (8000x *g*, 10 min, 4 ˚C) and the supernatant removed leaving behind ~ 50 µl for resuspension of the pellet. Immediately after resuspension, the samples were frozen in liquid nitrogen and stored at -80 ˚C until the RNA extraction was performed. Total RNA was extracted using the RNeasy Mini kit (QIAGEN). Samples were placed in a heating block at 37 ˚C and vigorously mixed with a pipette in 600 µl RLT buffer supplemented with 6 µl β-mercaptoethanol. The lysate was mixed with 600 µl 70 % ethanol and applied to a column. RNA was bound to the column by centrifugation (8000x *g*, 30s, RT). The flow-through was either collected for protein precipitation or discarded. The column was then washed with 350 µl RW1 (8000x *g*, 30s, RT) and two times with 500 µl RPE buffer (8000x *g*, 30s, RT). Total RNA was eluted twice with 30 µl of RNAse free water. The concentration of the RNA samples was measured and the samples stored at -80 ˚C until use.

Isolation of Poly A+ mRNA from total RNA

Poly A+ mRNA was isolated from total RNA using the Oligotex mRNA Mini Kit following the manufacturer's protocol. Briefly, 250 µl buffer OBB and 15 µl oligotex suspension were mixed with a volume of 250 µl of 100 – 250 µg total RNA. The sample was incubated at 70 ˚C for 3 min and then placed at 25 ˚C for 10 min to allow hybridisation between the oligo dT_{30} of the oligotex particle and the poly-A tails of the mRNAs. The oligotex::mRNA complex was separated by centrifugation for 2 min at 15,000 x *g* and the supernatant discarded. 150 µl each of RNAse free water and buffer OBB were added to the oligotex::mRNA complex (pellet) and mixed by pipetting. The mixture was again incubated at 70 ˚C for 3 min and then placed at 25 ˚C for 10 min. This was followed by pelleting of the oligotex::mRNA complex by centrifugation for 2 min at 15,000 x *g* and the supernatant discarded. The oligotex::mRNA pellet was resuspended in 400 µl of buffer OW2 and transferred onto a spin column followed by centrifugation for 1 min at 15,000 x *g*. The spin column was transferred into a new RNAse-free 1.5ml microcentrifuge tube and washed with 400 µl of buffer OW2 by centrifugation for 1 min at 15,000 x *g*. The spin column was again transferred into a new RNAse-free 1.5 ml microcentrifuge tube and the poly A+ mRNA eluted twice by resuspending the resin 3 – 4 times in 20 µl of hot (70 ˚C) buffer OEB followed by centrifugation at 15,000 x *g* for 1 min. The concentration of the poly A+ mRNA samples was measured and the samples stored at -80 ˚C until use.

cDNA synthesis

cDNA synthesis was carried out using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific) following the manufacturer's instructions. 1 µg of total RNA was used as template with oligo$(dT)_{18}$ primers from the kit.

Northern blot analysis

Northern blot analyses were carried out using 4 µg or 8 µg of total RNA. The RNA was denatured with glyoxal (4 µg total RNA in 4 µl RNAse-free water + 7.2 µl of glyoxal mix) at 50 ˚C for 40 min and then resolved on a 1.5 % agarose gel (Ultra-pure Agarose; ThermoFisher Scientific) in 1x RNA buffer. The resolved RNA was transferred overnight to Hybond-N/Hybond N+ nylon membrane (GE Healthcare) by upward capillary transfer. After transfer, the RNA was UV-crosslinked (Stratalinker 1800, Stratagene, USA; autocrosslink: 1200x100 µJ/cm^2) onto the membranes and deglyoxylated by baking at 80 ˚C for 1 h. The membranes were prehybridised at 42 ˚C for 1 h in hybridisation solution. The membranes were then hybridised overnight at 42 ˚C with hybridisation solution containing 10 nM of fluorescently-labelled oligonucleotide probes. After hybridisation, the membranes were washed

three times (10 min each) with Northern wash buffer (2x SSC buffer, 0.1 % SDS), dried and the blots analysed with a LI-COR Odyssey system (LI-COR Biosciences).

RNA dot blot analysis

RNA dot blot analysis was used for quantification of highly abundant mRNAs. For this, 3 µg of total RNA was denatured with deionised glyoxal as described for Northern blot analysis. A Hybond-N membrane (GE Healthcare, UK) of required size was placed in a Minifold dot blotter (Schleicher & Schuell, DE) and washed with 200 µl of 10x SSC buffer by suction pressure using a vacuum pump until all the buffer had gone through the membrane. Each gloxal-denatured RNA sample was mixed with 40 µl of 10x SSC buffer and 45 µl loaded into the wells of the dot blotter. The suction pressure was again applied until the liquid had gone through the membrane. The membrane was washed afterwards with 200 µl 10x SSC buffer with the application of suction pressure until the buffer had completely gone through the membrane. After the wash, the membrane was removed from the dot blotter and the RNA UV-crosslinked to the membranes and deglyoxylated as described for the Northern blot analysis. Hybridisation of probes and visualisation were carried out as described for Northern blot analysis.

RNA Electrophoretic mobility shift assays (REMSA)

RNA Electrophoretic mobility shift assays (REMSA) were carried out using the Light Shift Chemiluminescent RNA EMSA kit (ThermoFisher Scientific) according to the manufacturer's instructions. Briefly, a control reaction was carried out to test the kit using the IRE/IRP system provided in the kit strictly following the instructions provided by the manufacturer. The test reactions were then carried out by incubating the recombinant protein in a 20 µl reaction mixture (1x REMSA buffer, 5 % glycerol, 2 µg tRNA) containing 10 nM biotin-labelled 16mer RNA probe and incubated for 30 min at room temperature. For competition assays, unlabelled excess 16mer RNA or IRE RNA were added. The samples were separated on a 5 % native polyacrylamide gel in 0.5x TBE (Tris-borate-EDTA) buffer. After electrophoresis, the resolved samples on the gel were transferred to a Hybond N+ nylon membrane (GE Healthcare) by wet transfer at 400 mA for 30 min. The membrane was removed after transfer and the labelled probes UV-crosslinked (autocrosslink: 1200x100 µJ/cm^2) onto the membrane. The signal from the labelled probes was detected with HRP–conjugated streptavidin using the Chemiluminescent Nucleic Acid Detection Module (ThermoFisher Scientific) according to the manufacturer's instructions.

Northwestern blot for m^{6}A detection

Northwestern blots to detect m^{6}A modification were carried out as described by Viegas et al. (2020) with slight modifications. DNase I-treated total RNA (2 µg) or poly A+ mRNA (50 ng) samples were mixed with a 5x RNA loading buffer (30.8 % formamide, 2.6 % formaldehyde, 20 % glycerol, 0.2 % bromophenol blue, 4 mM EDTA pH 8) in a ratio of 1 volume 5x loading buffer to 4 volumes of RNA sample. Samples were denatured by heating (70 ˚C for 5 min) and immediately transferred onto ice. The samples were then loaded and resolved on a 1.4 % denaturing agarose gel (1.4 % agarose, 6.3 % formaldehyde, 1x MOPS buffer) at 100 V until the samples migrated halfway across the gel. The RNA was transferred overnight to a Hybond- N+ nylon membrane (GE Healthcare) by upward capillary transfer with 20x SSC buffer. The RNA was then UV-crosslinked to the membrane (Stratalinker 2400, autocrosslink function: 1200x100 µJ/cm^2). Membranes were stained with 0.02 % methylene blue diluted in 0.3 M sodium acetate (pH 5.5) for 5 min and washed in RNase-free water. After imaging, the methylene blue was removed by incubation in de-staining solution (0.2x SSC, 1 % SDS) and washed 3 times in PBST (PBS pH 7.4 with 0.2 % Tween-20). The membranes were then blocked by incubation in 5 % skimmed milk in PBST for 1 h and then incubated overnight with 1µg/ml mouse anti-m^{6}A antibody (1:1000 in 2.5 % skimmed milk in PBST) at 4 ˚C. Membranes were washed three times in PBST and then incubated with HRP-conjugated anti-mouse IgG diluted 1:10000 in 2.5 % skimmed milk in PBST for 1 h at room temperature. Membranes were again washed three times in PBST (10 min each) and the signal developed with Western Lightning Plus-ECL, Enhanced Chemiluminescence Substrate kit (PerkinElmer). Images were acquired using the iBright FL1000 imaging system (ThermoFischer Scientific).

2.2.4.3 Protein analyses

Pull-down assays and Mass spectrometry

These experiments were carried out in collaboration with Falk Butter's laboratory at the Institute of Molecular Biology (IMB), Mainz. The DNA and RNA pull-down assays were carried out using the DNA or RNA sequence of the *T. brucei* VSG121 3´ UTR. Briefly, for the DNA pull-down, a biotinylated DNA probe containing the first 188 nucleotides of the VSG121 3´ UTR with a 22 nt spacer added to the 5´ end was used as bait. As a control, a biotinylated DNA probe containing the first 188 nucleotides of the VSG121 3´ UTR (with scrambled 8mer and 16mer motifs) and a 22 nt spacer added at the 5´ was used. For RNA pull-down, RNA containing the first 188 nucleotides of the VSG121 3´ UTR (including both the 16mer and 8mer motifs) was used as bait. Controls for this included three different RNAs. Control 1 was the first 188 nucleotides of the VSG 3´ UTR with scrambled 16mer and 8mer, control 2 was the

reverse complement of control 1 and control 3 was the first 188 nucleotides of the VSG 3´ UTR with the 16mer reversed. The biotinylated bait and control DNA or RNA were coupled to streptavidin beads and exposed to *T. brucei* cell lysate. Bound proteins were then identified by mass spectrometry.

Preparation of protein extracts from trypanosomes

Protein extracts were prepared directly from harvested trypanosome cells or precipitated from the flow-through collected from total RNA extraction. For whole cell protein lysate prepared directly from harvested cells, the cells were harvested (1400x *g*, 10 min) and washed once with TDB by centrifugation (2000x *g*, 90 s). After complete removal of the supernatant, the samples were resuspended in 1x sample buffer to a concentration of 2 x 10^5 cell equivalents/µl. The samples were then boiled at 100 ˚C for 5 min, cooled and either used immediately or stored at -20 ˚C until use.

For precipitation of proteins after RNA isolation, 5 ml ice-cold acetone was added to the flow-through collected from the RNA isolation and incubated at -20 ˚C for 30 min. The solution was then centrifugated (20000x *g*, 10 min, 4 ˚C) to separate the precipitated proteins. The supernatant was discarded and the pelleted proteins air-dried at room temperature. The pellet was resuspended in 1x sample buffer to a concentration of 2 x 10^5 cell equivalents/µl. The samples were then boiled at 100 ˚C for 5 min, cooled and either used immediately or stored at -20 ˚C until use.

Discontinuous SDS polyacrylamide gel electrophoresis

For the separation of proteins in whole cell protein lysates according to their molecular weights, discontinuous SDS polyacrylamide gel electrophoresis (SDS-PAGE) was used. Gels composed of 2.5 % stacking and 12.5 % acrylamide/bisacrylamide (37.5:1) resolving gels were used. 1 x 10^6 cell equivalents were loaded per well. PageRuler pre-stained protein ladder (ThermoFisher Scientific) was loaded in a separate lane. The electrophoresis was carried out in 1x Laemmli buffer at 120 V until the dye front left the gel.

Staining of SDS polyacrylamide gels

Proteins separated by SDS-PAGE were stained in Coomassie brilliant blue (R-250) staining solution. The gels were immersed in the staining solution for 30 mins on a rocking stage. Next, the staining solution was replaced with Coomassie destaining solution and incubated on a rocking stage until the protein bands were visible.

Western blot analyses

To analyse and quantify specific proteins, proteins were size resolved on SDS-PAGE gels and transferred onto nitrocellulose membranes (GE Healthcare Life Sciences) using a semi-dry transfer apparatus (BIO-RAD). For transfer, the blot was assembled as follows from cathode to anode: 6x Whatman paper soaked in transfer buffer, membrane, gel, 6x Whatman paper soaked in transfer buffer. The transfer was carried out at 60 mA (~ 1.1 mA/cm2) for 60–90 min. After transfer, the membranes were blocked by incubation in 5 % milk powder in PBS for 1 h at room temperature (RT) or overnight at 4 ˚C. Appropriate amounts of primary antibodies were then applied in PBS/1 % milk/0.1 % Tween-20 solution for 1 h at RT. After four washes (5 min each) with PBS/0.2 % Tween-20, IRDye-conjugated secondary antibodies were applied in PBS/1 % milk/0.1 % Tween-20 solution for 1 h at RT in the dark. Membranes were finally washed once in PBS, dried between two Whatman papers and analysed using a LI-COR Odyssey system (LI-COR).

Protein dot blot analysis

Protein dot blot was used to quantify proteins without the need for separation based on molecular weight. For this, a nitrocellulose membrane was pre-incubated in PBS, dried and fixed in a dot blotter. 3 µl of a protein sample containing 6 x 10^5 cell equivalents was loaded into each well of the dot blotter. Subsequent steps for detection of the protein with antibodies were carried out as described in the Western blot protocol above.

Surface sensing of translation (SUnSET) assay

The SUnSET assay was used to monitor global translation without the use of radioactive isotopes. The assay is based on incorporation of puromycin (an aminoacyl tRNA structural analog) into a growing polypeptide resulting in blocking of elongation and termination of translation (Goodman & Hornberger, 2013). 1 x 10^7 mid-log phase cells were treated with 10 µg/ml puromycin pre- and post RNAi induction (24 h and 48 h) for 30 min. Following this the cells were washed once with trypanosome dilution buffer (TDB) and resuspended in appropriate volumes of 1x protein sample buffer. The samples were then boiled for 5 min at 100 ˚C to prepare whole cell protein lysate. As a control for translation inhibition, cells were treated with 50 µg/ml cycloheximide for 30 min prior to treatment with puromycin. A second control where uninduced cells were neither treated with cycloheximide nor puromycin was also included. The protein samples were resolved on a 12.5 % SDS gel. After transfer onto a nitrocellulose membrane, the membrane was incubated in REVERT 700 total protein stain for 5 min and rinsed twice in wash solution. The membrane was later destained after acquisition of an image. Subsequently, blocking and antibody detection were carried out as described above

in the Western blot protocol using anti-puromycin primary antibody and IRDye 680LT-conjugated goat anti-mouse secondary antibody.

2.2.5 Light microscopy

Fluorescence microscopy to visualise endogenously-tagged genes

1 x 10^7 – 2 x 10^7 cells were harvested by centrifugation at 1400x *g* for 10 min. The cells were re-suspended in 1ml of sterile PBS and transferred into 1.5 ml microcentrifuge tubes. This was followed by centrifugation (2000x *g* for 30 s) and removal of the supernatant. The cells were then re-suspended in 500 µl PBS and fixed with 4 % FA and 0.05 % glutaraldehyde followed by incubation at 4 ˚C overnight. Fixed cells were washed twice by centrifugation (2000x *g* for 30 s) with 1ml sterile PBS. After the last wash, the supernatant was carefully removed leaving behind about 50 µl for re-suspension of cells. 4.5 µl of the cell suspension was transferred into a new 1.5 ml microcentrifuge tube and 0.5 µl of Vectashield mounting medium containing DAPI added. 3 µl of the mixture was then loaded onto a slide. Image acquisitions were performed using an automated DMI6000B wide field fluorescence microscope (Leica Microsystems, Germany), outfitted with a DFC365FX camera (pixel size 6.45 µm) and a 100x oil objective (NA 1.4). Acquired images were deconvolved using the Hyugens Essential software and further processing carried out using FIJI software (Schindelin et al., 2012).

Cell cycle analysis after DAPI staining

1 x 10^7 cells were harvested by centrifugation (1400x *g* for 10 min) as described above. In some cases the cells were fixed with 4 % FA for 15 min. After washing of cells with TDB, the supernatant was removed leaving behind ~ 50 µl for resuspension of cells. 4 µl of cells were transferred into 1.5 ml microcentrifuge tube and 1 µl of a 5 µg/ml DAPI stock solution added. After 2 – 5 min incubation, the cell cycle profile was determined by evaluating the nucleus/kinetoplast (N/K) configuration of at least 200 intact cells using a Nikon Eclipse TS100 microscope.

3.0 RESULTS

3.1 The role of the 16mer motif in cell viability and functional VSG expression

Christopher Batram conducted earlier studies on the effect of mutations in the VSG 3′ UTR on VSG expression (Batram, Diplom thesis 2009; PhD thesis 2013). In these studies, an ectopic VSG121 was introduced upstream of the endogenous VSG in 13-90 cells expressing VSG221 to generate double-expresser cells. These double-expresser cells expressed approximately 50% of both VSGs. Through mutation of the ectopic VSG121 3′ UTR, he found that deletion of the entire 3′ UTR or even parts of it harbouring the 16mer motif resulted in drastically decreased VSG levels. Some mutations involving substitution of nucleotides within the 16mer motif, 8mer motif and the region between the two motifs were however tolerated. Subsequent RNAi-mediated depletion of the endogenous VSG221 was lethal in cells with drastically decreased VSG levels but VSGs with the tolerated mutations compensated the loss of the endogenous VSG221. These experiments were however performed using double-expresser cell lines in which the ectopic VSG was placed upstream of the endogenous VSG. As VSGs are usually expressed from a telomere proximal position, it was not clear if the position of the ectopic VSG affected the outcome of the VSG 3′ UTR mutations on VSG expression. Also, since RNAi was used to deplete the endogenous VSG221, it could not be established whether the intact 16mer motif was required for expression of functional levels of the VSG.

Therefore, in this Study I generated double-expresser cell lines where the ectopic VSG121 was placed at a telomere proximal position and performed mutational analysis on the ectopic VSG121 3′ UTR. I also generated single-expresser cells by deleting the endogenous VSG221 to investigate whether cells require the 100 % conserved 16mer motif for viability and expression of functional levels of VSG.

3.1.1 Generation of stable double-expresser cells

To investigate whether 100 % conservation of the 16mer motif is essential for cell viability and expression of functional levels of the VSG, transgenic trypanosome cells expressing two VSGs from the active expression site (double-expresser cells) were generated using 13-90 cells expressing VSG221. It was hypothesised that factors binding to the telomere could interact with the VSG 3′ UTR and thus influence the regulation of VSG expression. Hence to make the double-expresser cell line $221^{ES}121_{WT}$, a modified pbRn6 construct (pbRn6M1.6.nPPT) containing an ectopic VSG121 with a wild-type 3′ UTR was integrated downstream of the endogenous VSG221, closer to the telomere (Figure 8A). The resulting $221^{ES}121_{WT}$ cells grew

slightly slower (~ 7 h doubling time) than the parental 13-90 cells (~ 6 h doubling time) (Figure 8B) and successfully expressed both VSGs. The ectopic *VSG121* mRNA was expressed at ~ 30 % of wild-type levels whereas the endogenous *VSG221* was expressed at ~ 80 % (Figure 8C). Levels of the VSG121 and VSG221 proteins were 47 % and 62 %, respectively (Figure 8D). The mRNA and protein levels of the ectopic and endogenous VSGs obtained in the $221^{ES}121_{WT}$ cell line differed from the levels in previous double-expresser cell lines where approximately equal amounts of both VSGs were obtained when an ectopic VSG was placed upstream of the endogenous VSG (Muñoz-Jordán et al., 1996; Ridewood et al., 2017; Smith et al., 2009). However, the endogenous VSG221 levels decreased at an approximately inverse amount to the expression levels of the ectopic VSG121, similar to observations made in other double-expresser cell lines (Ridewood et al., 2017; Smith et al., 2009). This suggests that the genomic location of an ectopic VSG in the expression site affects its expression levels in double-expresser cell lines. In addition, the regulation of the endogenous VSG levels to ensure production of constant amounts of VSG appears to be independent of the location of the ectopic VSG withn the active expression site.

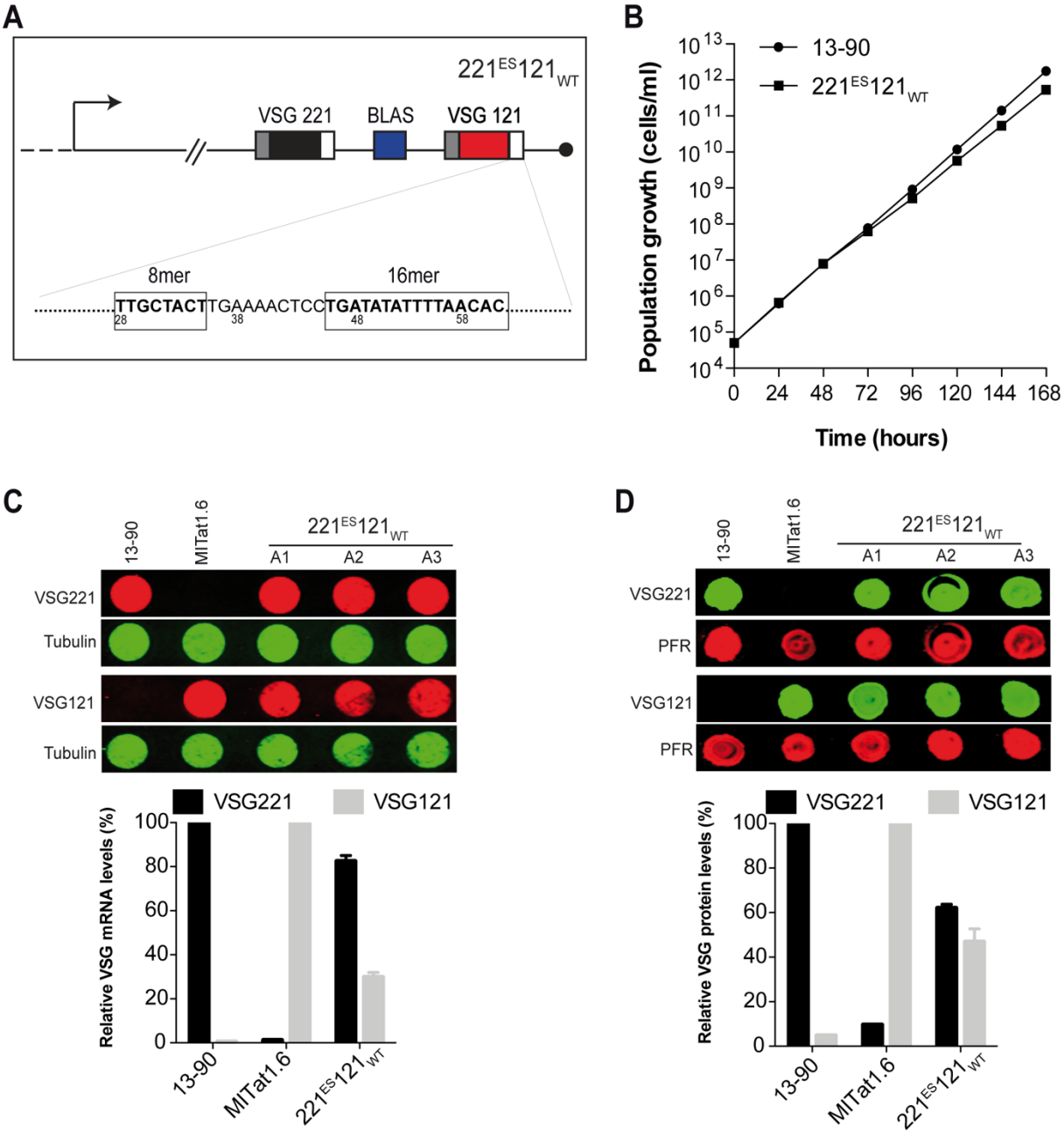

Figure 8. Generation of stable double-expresser cell line. (A) Schematic of the 221ES121$_{WT}$ double-expresser cell line. Ectopic VSG121with a wild-type *VSG121* 3′ UTR was integrated downstream of the endogenous VSG221 with a blasticin resistance cassette (BLAS) for selection. The ES promoter is indicated with an arrow and the telomere as a black circle. (B) Growth curve of the 221ES121$_{WT}$ cell line (black square). Mean ± standard error of the mean (SEM) of three independent clones is shown. The parental 13-90 cell line (black circle) served as a growth control. (C) RNA dot blot showing expression of *VSG221* and *VSG121* mRNA in three clones (A1, A2 and A3) of 221ES121$_{WT}$ cell line (upper panel). Levels of the *VSG221* mRNA in a single clone of 13-90 cells and *VSG121* mRNA in a single clone of MITat1.6 cells represent wild-type levels. *Tubulin* mRNA was used as a loading control. Quantification of *VSG* mRNA in 221ES121$_{WT}$ cells (lower panel). The levels of the *VSG* mRNA in the dot blots were normalised to the *tubulin* signal and expressed relative to levels in the parental 13-90 (VSG221) or MITat1.6 (VSG121) cells. Data are presented as means ± standard error of the mean (SEM) of three independent clones. (D) Protein dot blot showing expression of VSG221 and VSG121 protein in three clones (A1, A2 and A3) of 221ES121$_{WT}$ cell line. Levels of the VSG221 protein in 13-90 cells and VSG121 protein in MITat1.6 cells represent wild-type

levels. The paraflagellar rod proteins 1 and 2 (PFR) served as loading control. Quantification of VSG protein in $221^{ES}121_{WT}$ cells (lower panel). The levels of the VSG protein in the dot blots were normalised to the PFR signal and expressed relative to levels in the parental 13-90 (VSG221) or MITat1.6 (VSG121) cells. Data are presented as means ± standard error of the mean (SEM) of three independent clones.

3.1.2 The telomere-proximal 16mer motif does not modulate endogenous VSG expression

Under natural conditions, the expressed VSG is usually located in the telomeric region of the expression site. Therefore, having successfully generated a stable double-expresser cell line with the ectopic VSG121 at the telomere proximal position, it was necessary to investigate whether mutations in the 16mer motif that lies within the VSG121 3′ UTR, immediately upstream of the telomere, affected the expression of the endogenous VSG221. Since endogenous VSG levels are modulated in the presence of a second VSG, the ectopic *VSG121* ORF in the active expression site of the double-expresser cell line was replaced with a *GFP* ORF flanked by a VSG121 3′ UTR with either an intact (wild-type VSG121 3′ UTR) or mutated (nucleotides 46-52 of the VSG121 3′ UTR deleted) 16mer motif. If the telomere proximal 16mer is important for the expression of the endogenous VSG, a decrease in the endogenous VSG expression would be expected upon expression of the GFP with the mutated 16mer motif. To produce the GFP expressing cell lines, the pbRn6 plasmid was modified to generate the plasmids pbRn6GFP.nPPT and pbRn6GFPΔ46-52.nPPT for the intact and mutated 16mer motifs, respectively. The plasmids were then inserted downstream of the endogenous VSG221 in 13-90 BSF *T. brucei* cells to generate the cell lines $221^{ES}GFP_{WT}$ and $221^{ES}GFP_{\Delta 46\text{-}52}$ (upper: Figure 9A, upper schematic). Both cell lines expressed high levels of GFP but grew slightly slower than the parental 13-90 cells with $221^{ES}GFP_{WT}$ having a population doubling time (PDT) of 7.4 h and $221^{ES}GFP_{\Delta 46\text{-}52}$ having a PDT of 6.4 h (Figure 9A, B). Analysis of the GFP levels showed that *GFP* mRNA was 2.7-fold less in $221^{ES}GFP_{\Delta 46\text{-}52}$ cells compared to $221^{ES}GFP_{WT}$ cells (Figure 9C). The levels of the GFP protein produced were also drastically reduced (by ~ 5-fold) in the $221^{ES}GFP_{\Delta 46\text{-}52}$ cells compared to $221^{ES}GFP_{WT}$ cells (Figure 9D). This was in agreement with earlier findings that the 16mer motif is essential for the high abundance and stability of VSG mRNA (Berberof et al., 1995; Ridewood et al., 2017). The levels of the endogenous *VSG221* mRNA and VSG221 protein were however similar between the two GFP-expressing cell lines (Figure 9E, F), indicating that mutation of the telomere proximal 16mer motif did not affect the endogenous VSG221 expression. Hence, the first 16mer upstream of the telomere in the stable double-expresser cells does not appear to be important for the expression of the endogenous VSG.

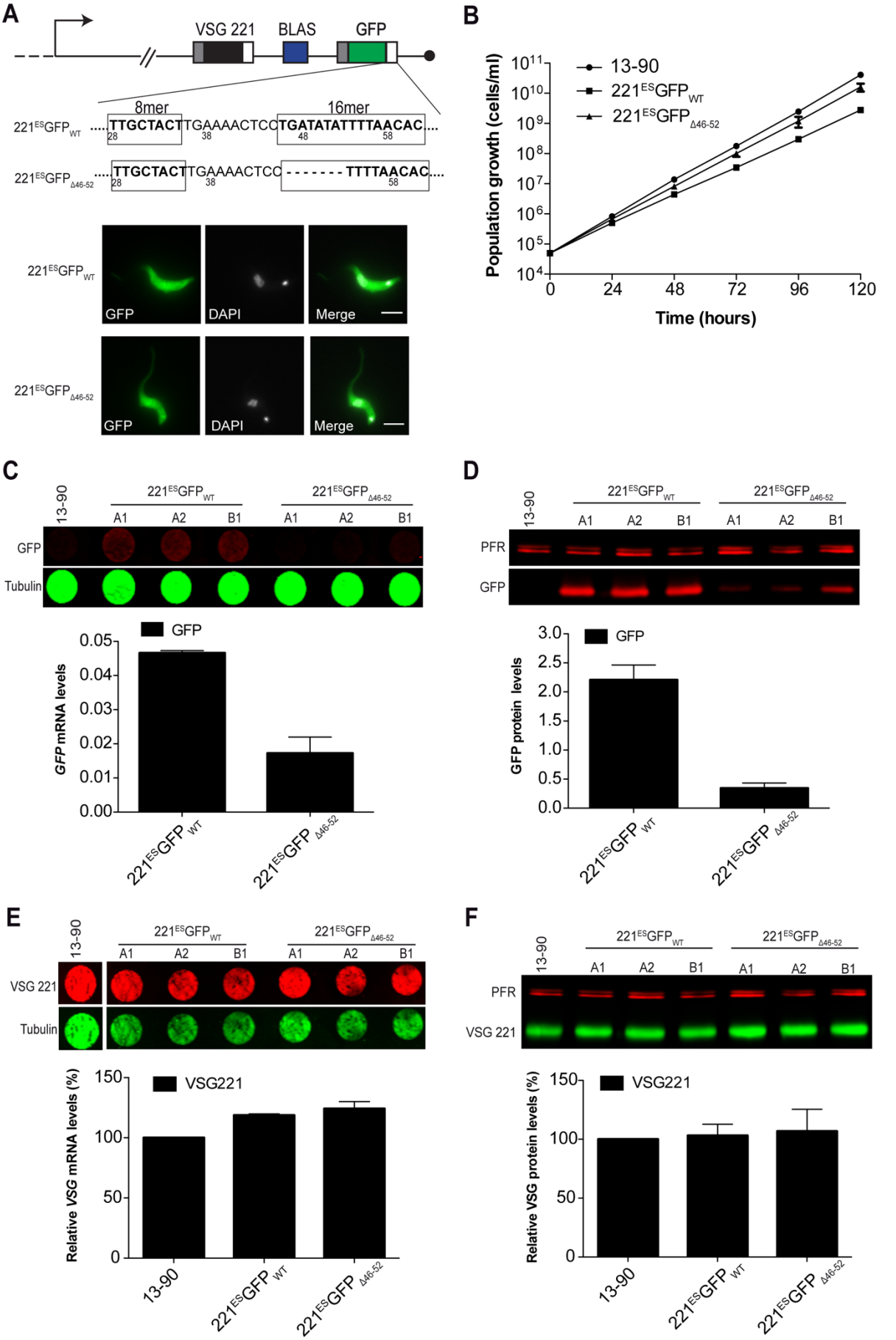

Figure 9. The telomere proximal 16mer motif does not affect the endogenous VSG221 expression. (A) Schematic and fluorescence images of the GFP reporter cell lines. The *GFP*

was either coupled to a *VSG121* 3´ UTR with an intact 16mer motif ($221^{ES}GFP_{WT}$) or a *VSG121* 3´ UTR with the first seven nucleotides of the 16mer motif deleted ($221^{ES}GFP_{\Delta 46\text{-}52}$) (upper panel). Fluorescence imaging showed high GFP expression in both cell lines (lower panel). (B) Cumulative growth curve of $221^{ES}GFP_{WT}$ and $221^{ES}GFP_{\Delta 46\text{-}52}$ cell lines. Data are averages from three independent clones with error bars representing means ± standard error of the mean (SEM). The parental 13-90 cell line (black circle) served as a growth control. (C) Quantification of *GFP* mRNA expression. The *GFP* mRNA was quantified from an RNA dot blot (upper panel) by normalising the signal from fluorescently labelled GFP probe to *tubulin* mRNA. Data are presented as means ± SEM of three independent clones. (D) Quantification of GFP protein expression. The GFP protein expression in $221^{ES}GFP_{WT}$ and $221^{ES}GFP_{\Delta 46\text{-}52}$ cell lines was quantified from a Western blot (upper panel). The paraflagellar rod proteins 1 and 2 (PFR) served as loading control and was used for normalisation. Data are presented as means ± SEM of three independent clones. (E) Quantification of *VSG221* mRNA expression. *VSG221* mRNA was quantified from an RNA dot blot (upper panel) by normalising the signal from a fluorescently labelled VSG221 probe to *tubulin* mRNA. Values of the *VSG221* mRNA are expressed relative to levels in the parental 13-90 cells and presented as means ± SEM. (F) Quantification of VSG221 protein expression. VSG protein expression was quantified from a Western blot (upper panel). PFR served as a loading control and was used for normalisation. Values of the VSG221 protein are expressed relative to levels in the parental 13-90 cells and presented as means ± SEM.

3.1.3 Mutational analysis of the 16mer and 8mer motifs

In earlier studies on the effect of mutations in the VSG 3´ UTR on VSG expression, it was predicted using bioinformatic tools that the 16mer and 8mer motifs form part of a stable stem loop structure and mutations that disrupt this structure resulted in drastic reduction in VSG expression (Batram, Diplom thesis 2009; PhD thesis 2013; Ridewood et al., 2017). These experiments were however performed using double-expresser cell lines in which the ectopic VSG was placed upstream of the endogenous VSG and the mutations were introduced in the ectopic VSG 3´ UTR which is distal to the telomere. To investigate if these mutations have a different effect when introduced in the telomere proximal VSG 3´ UTR, the ectopic VSG121 3´ UTR in the double-expresser cell line $221^{ES}121_{WT}$ was mutated. To generate the mutant cell lines modified pbRn6 constructs with different mutations within the 16mer and 8mer motifs were integrated into the active expression site downstream of the endogenous VSG221 to produce the mutant cell lines $221^{ES}121_{\Delta 46\text{-}52}$ (with the first 7 nucleotides of the 16mer deleted), $221^{ES}121_{\Delta 46\text{-}55}$ (with the first 10 nucleotides of the 16mer deleted), $221^{ES}121_{N46\text{-}48}$ (with the first 3 nucleotides of the 16mer substituted) and $221^{ES}121_{inver8mer}$ (with 8mer motif inverted) (Figure 10A). All of the mutant cell lines had population doubling times of ~ 6 h, similar to that of the parental 13-90 cells (Figure 10B). The endogenous *VSG221* mRNA remained at approximately wild-type levels although they were a little elevated in $221^{ES}121_{\Delta 46\text{-}52}$ (120 %), $221^{ES}121_{\Delta 46\text{-}55}$ (117 %) and $221^{ES}121_{N46\text{-}48}$ (111 %) cell lines and decreased in $221^{ES}121_{inver8mer}$ cell line (86 %) (Figure 10C). Less than 10 % of the ectopic *VSG121* mRNA was however transcribed in all

of the mutant double-expresser cell lines. At the protein level, VSG221 protein was expressed at or above the wild-type levels in all the cell lines (Figure 10D). VSG121 protein was less than 15 % in $221^{ES}121_{\Delta 46\text{-}52}$ (5 %), $221^{ES}121_{\Delta 46\text{-}55}$ (10 %) and $221^{ES}121_{N46\text{-}48}$ (12 %) cell lines. The $221^{ES}121_{inver8mer}$ cell line however expressed 21 % of the ectopic VSG121 protein. The RNA and protein amounts of the ectopic VSG121 expressed in the $221^{ES}121_{\Delta 46\text{-}52}$ and $221^{ES}121_{\Delta 46\text{-}55}$ cell lines were in agreement with previous studies by Christopher Batram as the mutations introduced in these cell lines have been shown to cause a dramatic decrease in VSG expression (Batram, Diplom thesis 2009). However, the cell lines $221^{ES}121_{N46\text{-}48}$ and $221^{ES}121_{inver8mer}$ produced contradictory results to the study by C. Batram, as inversion of the 8mer or substitution of the first three nucleotides were found to not significantly affect VSG expression (Batram, Diplom thesis 2009; PhD thesis 2013). Differences in the effect of the mutations could however be attributed to the location of the ectopic VSG within the expression site as it was found that the endogenous VSG was preferentially expressed in higher amounts when the ectopic VSG was inserted downstream (Figure 8).

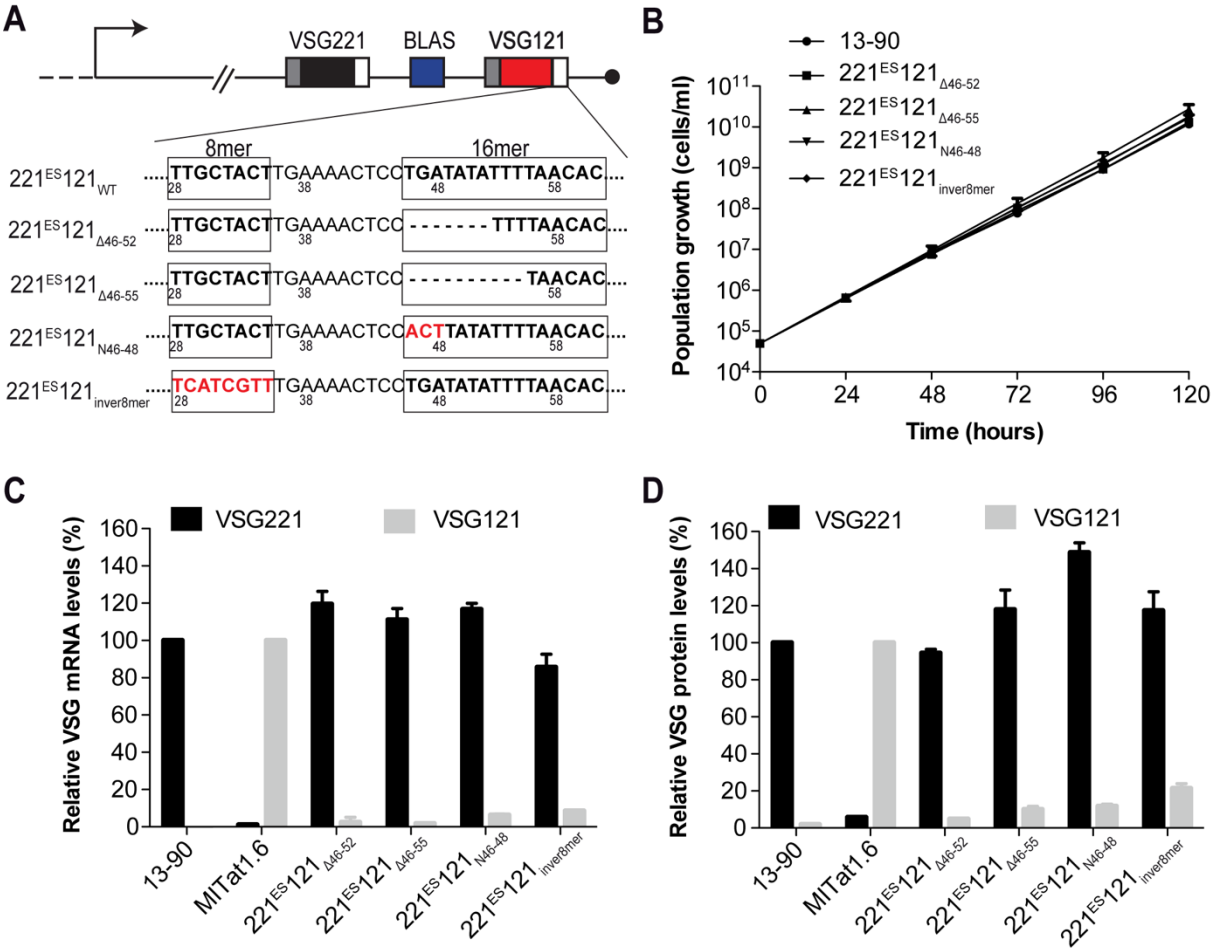

Figure 10. Mutational analysis of the ectopic VSG121 3´ UTR. (A) Schematic of the double-expresser mutant cell lines generated. Dashes (-) represent deleted nucleotides. The nucleotides in red are substitutes of the original nucleotides in the 8mer and 16mer motifs. (B) Cumulative growth curve of the mutant double-expresser cell lines. Data are averages from three

independent clones with error bars representing means ± standard error of the mean (SEM). The parental 13-90 cell line (black circle) served as a growth control. (C) Levels of *VSG221* and *VSG121* mRNA expressed in the mutant double-expresser cell lines. The quantifications were made from RNA dot blots. Amounts of the *VSG221* and *VSG121* mRNAs were normalised to *tubulin* mRNA, and expressed relative to levels in the parental 13-90 (VSG221) or MITat1.6 (VSG121) cells. Data are presented as means ± SEM of three independent clones. (D) Levels of VSG221 and VSG121 protein expressed in the mutant double-expresser cell lines. The quantifications were made from protein dot blots. Amounts of the VSG221 and VSG121 proteins were normalised to the paraflagellar rod proteins 1 and 2 (PFR), and expressed relative to levels in the parental 13-90 (VSG221) or MITat1.6 (VSG121) cells. Data are presented as means ± standard error of the mean (SEM) of three independent clones.

3.1.4 100 % conservation of the 16mer is dispensable for expression of functional VSG levels

It had been established that the 16mer motif, which is 100 % conserved in *T. brucei* VSGs, is essential for maintenance of VSG mRNA stability (Berberof et al., 1995; Ridewood et al., 2017). Results from Christopher Batram's study however demonstrated that substituting the first three nucleotides of the motif did not significantly impact VSG expression. As these experiments were performed in the presence of a second VSG using double-expresser cells, it was not established if parasites expressing a single VSG harbouring the tolerated mutation would be viable and express functional levels of the VSG. To address this, a single-expresser cell line expressing the ectopic VSG121 with the tolerated 16mer mutation was generated from the mutant double-expresser cell line $221^{ES}121_{N46\text{-}48}$. Knock-out of the endogenous VSG221 in the $221^{ES}121_{N46\text{-}48}$ cell line (using the construct pJET1.2_M1.2-Blas KO) resulted in the single-expresser cell line $\Delta221^{ES}121_{N46\text{-}48}$ (Figure 11A). As control, a single-expresser cell line with the wild-type VSG121 3´ UTR ($\Delta221^{ES}121_{WT}$) was generated from the double-expresser cell line $221^{ES}121_{WT}$ using the same knock-out approach. Both single-expresser cell lines were viable but the 16mer mutant cells ($\Delta221^{ES}121_{N46\text{-}48}$) showed reduced growth compared to $\Delta221^{ES}121_{WT}$ cells (Figure 11B). The $\Delta221^{ES}121_{WT}$ cells had a PDT of ~ 9 h while $\Delta221^{ES}121_{N46\text{-}48}$ cells had a PDT of ~ 11 h.

To investigate whether a cell cycle defect accounted for the reduced proliferation of the $\Delta221^{ES}121_{N46\text{-}48}$ cells, a cell cycle analysis was carried out. In *T. brucei*, the cell cycle is very well organised and tightly controlled. The timing of the division of the nucleus and kinetoplast DNA is different. Cells can therefore be assigned to a cell cycle stage based on the number of nuclei (N) and kinetoplasts (K) they have. A typical cell progresses from 1K1N through 1Kd1N (cells with a dividing kinetoplast) to 2K1N. The nucleus then divides giving rise to 2K2N cells. These cells then undergoe cytokinesis to produce new daughter cells with 1K1N configuration. A typical unsynchronised culture therefore contains mostly 1K1N (~ 80 %) cells and a few cells

with 2K1N (~ 12 %) and 2K2N (~ 8 %) configurations. Results from the analysis showed that the reduced growth of the $\Delta 221^{ES}121_{N46\text{-}48}$ cells was not due to a cell cycle defect, as there was no specific aberration in the cell cycle profile (Figure 11C). Analysis of the VSG expression showed that the amount of *VSG121* mRNA was 43 % higher in $\Delta 221^{ES}121_{WT}$ cells (79 %) compared to $\Delta 221^{ES}121_{N46\text{-}48}$ cells (36 %) (Figure 11D). Both cell lines however expressed similar levels of VSG121 protein comparable to wild-type amounts (Figure 11E). These data suggested that although the intact 16mer motif is required for high level VSG expression, 100 % conservation of the motif is not essential for cell viability and functional VSG expression.

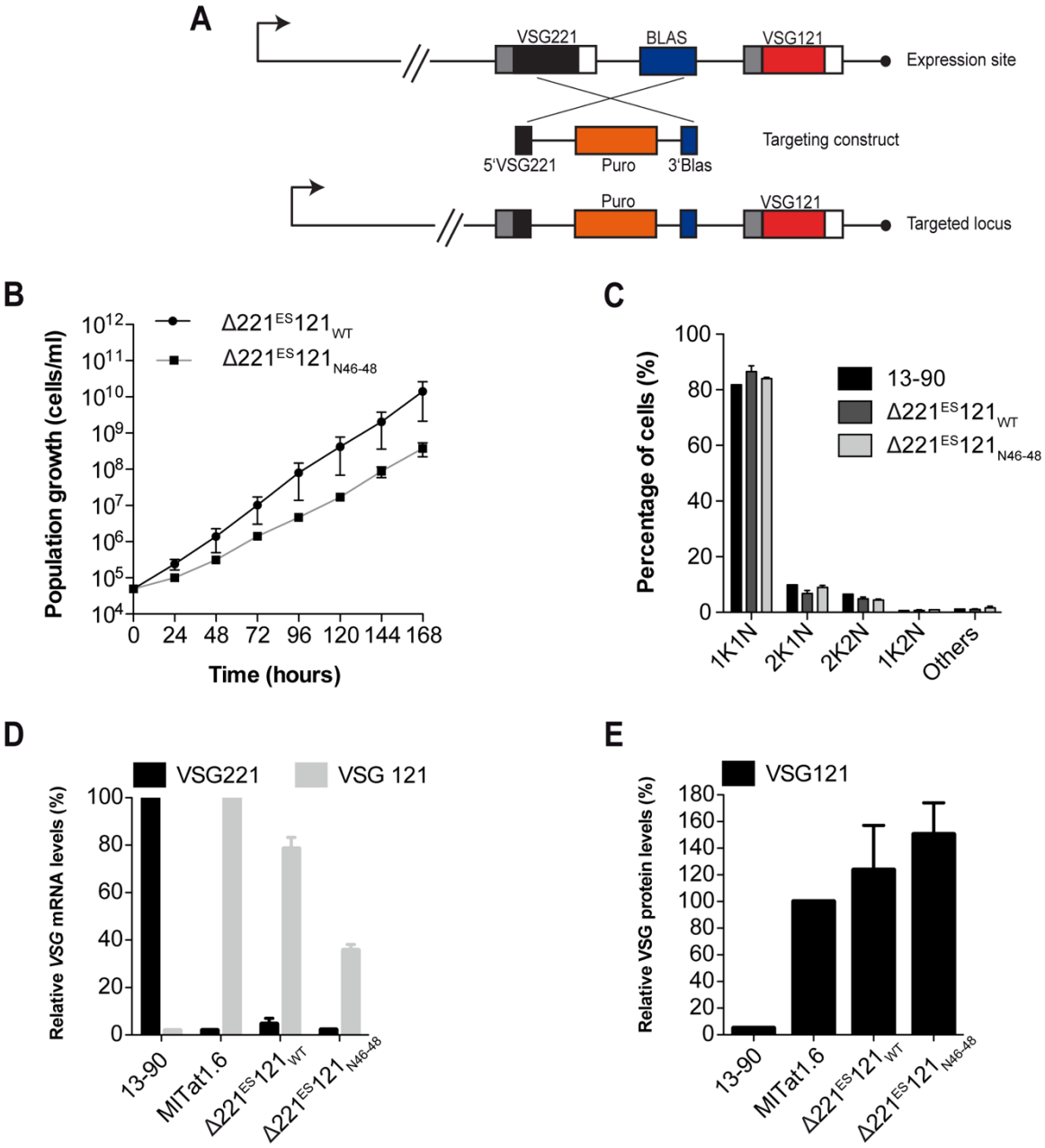

Figure 11. Mutation of the first three nucleotides of the 16mer motif supports functional VSG protein expression. (A) Schematic of the knockout strategy used in generating single-expresser cells. The endogenous *VSG221* and blasticidin resistance genes were replaced with a puromycin resistance gene. (B) Cumulative growth curve of $\Delta 221^{ES}121_{WT}$ and $\Delta 221^{ES}121_{N46\text{-}48}$

single-expresser cell lines. Data are averages from three independent clones with error bars representing means ± standard error of the mean (SEM) (C) Cell cycle analysis of $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46-48}$ single-expresser cell lines. Cells were stained with DAPI and classified to the different cell cycle stages according to the number of nuclei (N) and kinetoplasts (K). At least 200 cells were analysed for each time point. Three independent clones were analysed, and data represented as mean percentages ± SEM. The parental 13-90 cell line was used as a control. (D) Quantification of *VSG* mRNA in $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46-48}$ cells. *VSG221* and *VSG121* mRNA were quantified from RNA dot blots using fluorescently labelled probes normalised to *tubulin* mRNA (upper panel). (E) Quantification of VSG protein in $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46-48}$ cells. VSG121 protein levels were normalised to paraflagellar rod protein 1 and 2 (PFR) (upper panel). Data presented are means ± standard error of the mean (SEM) of three independent clones. For *VSG121* mRNA and protein quantifications, all values are expressed relative to levels in the parental 13-90 (VSG221) or MITat1.6 (VSG121) cells.

3.1.5 An intact 16mer is not required for N^6-methyladenosine modification of VSG transcripts

In bloodstream *T. brucei* cells, the active *VSG* mRNA is abundant and highly stable with a half-life greater than 1 h compared to less than 20 min for most other transcripts (Ehlers et al., 1987; Fadda et al., 2014). The 16mer is known to be involved in maintenance of this high *VSG* mRNA stability but the mechanism involved in stabilisation of the transcript is not fully understood. A recent study suggested that the 16mer motif is essential for N^6-methyladenosine (m^6A) modification of *VSG* poly(A) tails contributing to stabilisation of the *VSG* transcripts (Viegas et al., 2020). As the data in section 3.1.4 showed that a 100 % conservation of the 16mer motif is not essential for functional VSG protein production but is required for high transcript abundance, it was investigated whether the m^6A modification of *VSG* mRNA requires the intact 16mer motif. To test this, m^6A immunoblotting was performed using Poly(A+) RNA extracted from $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46-48}$ cells. To normalise the m^6A signal to the mRNA levels, an 'm^6A index' was computed as described (Viegas et al., 2020) by dividing the relative intensity of the m^6A signal in *VSG* transcripts by the corresponding *VSG* mRNA levels measured by quantitative dot blots (Figure 12). It was found that *VSG121* mRNA levels were 20 % less in $\Delta221^{ES}121_{N46-48}$ cells compared to $\Delta221^{ES}121_{WT}$ (Figure 12A). This difference in *VSG* mRNA levels between the two cell lines was less than the 43 % initially measured in Figure 11, pointing to possible cell adaptation as the $\Delta221^{ES}121_{N46-48}$ cells retained the mutation but expressed a higher amount of *VSG121* mRNA. The immunoblot revealed m^6A bands corresponding to the *VSG121* transcript size in both $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46-48}$ cells with m^6A intensities of 0.43 and 0.33, respectively (Figure 12B,C). The m^6A index was however similar (~ 0.5 AU) in both cell lines. This suggested that in the absence of an intact 16mer motif, m^6A modification of *VSG* mRNA still occured. 100 % conservation of the 16mer motif is therefore not required for m^6A modification of *VSG* transcripts.

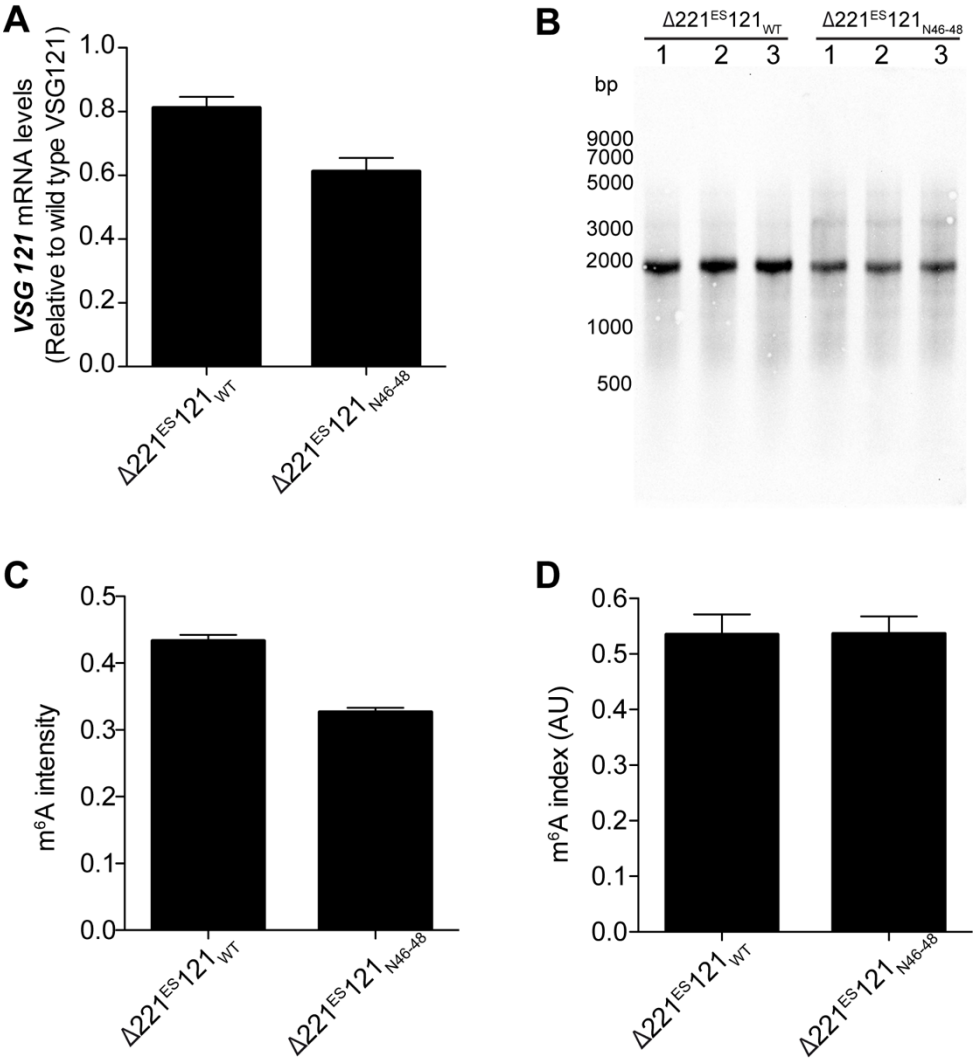

Figure 12. 100 % conservation of the 16mer motif is not required for m6A modification of VSG transcripts. (A) Quantification of *VSG* mRNA in $\Delta 221^{ES}121_{WT}$ and $\Delta 221^{ES}121_{N46\text{-}48}$ cells used in m^6A immunoblots. *VSG121* mRNA was normalised to *tubulin* mRNA and expressed relative to the *VSG121* transcript levels in MITat1.6 cells. (B) Exemplary immunoblot showing m^6A amounts in mRNA from $\Delta 221^{ES}121_{WT}$ and $\Delta 221^{ES}121_{N46\text{-}48}$ cells. (C) Intensity of m^6A signal in *VSG121* mRNA measured using Image J. The intensity of the m^6A in the VSG121 band was expressed relative to the intensity of the entire lane. (D) m^6A index in the two single-expresser cell lines. The m^6A signal was computed as the ratio of the m^6A intensity in (C) to the *VSG121* mRNA levels in (A). Data presented are means ± standard error of the mean (SEM) of three independent clones.

3.2 The role of the 16mer motif in VSG silencing and coat exchange

3.2.1 The intact 16mer is not required for VSG silencing during differentiation

A major feature of the differentiation process from bloodstream form trypanosomes to procyclic forms in the tsetse fly midgut is the replacement of the VSG coat with a procyclin coat. During this process, the highly stable active VSG becomes destabilised and degraded. The role of the 100 % conserved 16mer motif in VSG silencing during the differentiation from BSF to PCF was therefore investigated. An *in vitro* differentiation assay was performed as earlier described

using 6 mM cis-aconitate and incubation of cells at 27 °C (Engstler & Boshart, 2004; Overath et al., 1986). $\Delta 221^{ES}121_{WT}$ and $\Delta 221^{ES}121_{N46-48}$ single-expresser cells were incubated at 27 °C after treatment with 6 mM cis-aconitate and samples collected at 0, 6, 24 and 48 h after induction of differentiation. The amounts of *VSG121* mRNA and VSG121 protein were then determined over the course of differentiation. There was a decrease of *VSG121* mRNA to less than 20 % of the wild-type levels in both cell lines after 6 h of inducing differentiation (Figure 13A). The amount of *VSG121* mRNA decreased further to less than 10 % by 48 h post-induction of differentiation. The VSG121 protein also decreased drastically to ~ 25 % by 24 h and to less than 10 % by 48 h (Figure 13B). The kinetics of silencing of *VSG121* mRNA and VSG121 protein were similar in $\Delta 221^{ES}121_{WT}$ and $\Delta 221^{ES}121_{N46-48}$ cells, suggesting that the intact 16mer motif is not essential for VSG silencing during differentiation from BSF to PCF. This could also mean that crosstalk between the VSG and procyclin does not require 100 % conservation of the 16mer in the VSG 3´ UTR.

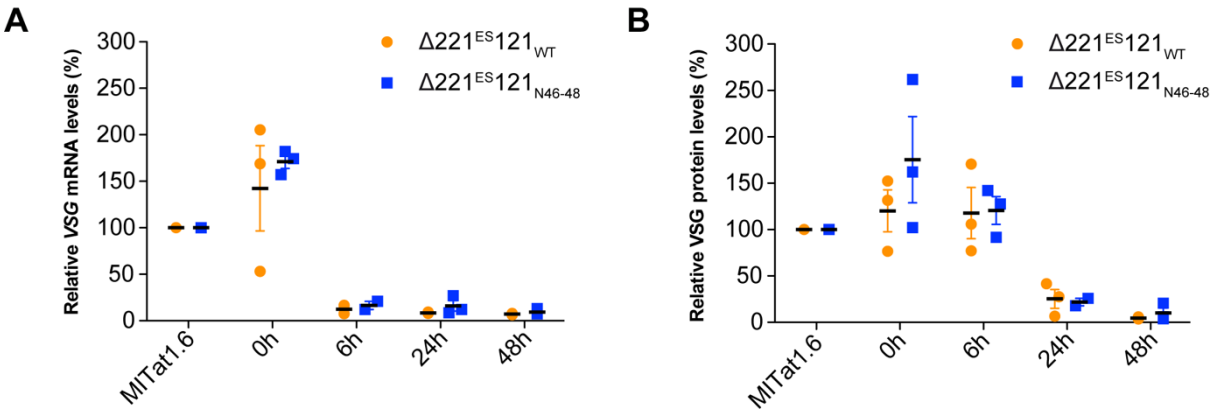

Figure 13. The active VSG is efficiently silenced in the absence of an intact 16mer motif during differentiation from BSF to PCF. (A) *VSG121* mRNA levels in $\Delta 221^{ES}121_{WT}$ and $\Delta 221^{ES}121_{N46-48}$ cells over the course of differentiation. *VSG121* mRNA was normalised to *tubulin* mRNA and expressed relative to the *VSG121* mRNA levels in MITat1.6 wild-type cell line. (B) VSG121 protein levels in $\Delta 221^{ES}121_{WT}$ and $\Delta 221^{ES}121_{N46-48}$ cells over the course of differentiation. VSG121 protein was normalised to PFR and expressed relative to the VSG121 protein levels in a MITat1.6 wild-type cell line. Values presented are means ± standard error of the mean (SEM) of three independent clones.

3.2.2 Possible crosstalk between VSG and procyclin 3´ UTRs is not dependent on an intact 16mer

Factors involved in upregulation of procyclin mRNA also downregulate the expression of VSGs (Pays, 2005). These factors responsible for the transcriptional crosstalk between the VSG and procyclin act on either the 3´ UTRs or RNA elongation (Berberof et al., 1995; Vanhamme et al., 1995). To directly check the effect of the tolerated 16mer mutation on possible crosstalk between the procyclin 3´ UTR and the ES-resident VSG 3´ UTR, an inducible overexpression system as described by Batram et al., (2014) was used. This overexpression system integrates into a transcriptionally silent rDNA spacer, and expression of the gene driven by a tetracycline-regulated T7 promoter. The system was modified by replacing the VSG gene with an EP1 gene fused to eYFP (EP1-eYFP) with a wild-type EP1 3´ UTR. The EP1-eYFP was integrated into a transcriptionally silent rDNA spacer in $\Delta 221^{ES}121_{WT}$ and $\Delta 221^{ES}121_{N46\text{-}48}$ cells to generate $\Delta 221^{ES}121_{WT}EP1^{Tet}$ and $\Delta 221^{ES}121_{N46\text{-}48}EP1^{Tet}$ cell lines, respectively (Figure 14A). Overexpression of EP1 in both $\Delta 221^{ES}121_{WT}EP1^{Tet}$ and $\Delta 221^{ES}121_{N46\text{-}48}EP1^{Tet}$ cells showed a very slight decrease in growth compared to uninduced cells (Figure 14B). A slight increase in *EP1-eYFP* mRNA and an increase in EP1-eYFP protein with no significant differences in amounts between the two cell lines observed (Figure 14C, D). There was also no significant impact on the ES-resident VSG expression at both the mRNA and protein levels. Both cell lines expressed similar levels of *VSG121* mRNA and VSG121 protein, and these amounts were not significantly different from the levels measured prior to induction of overexpression (Figure 14E, F). The similarities in VSG expression between the two cell lines possibly indicate that the crosstalk between the EP1 and VSG 3´ UTRs is not dependent on the intact 16mer motif.

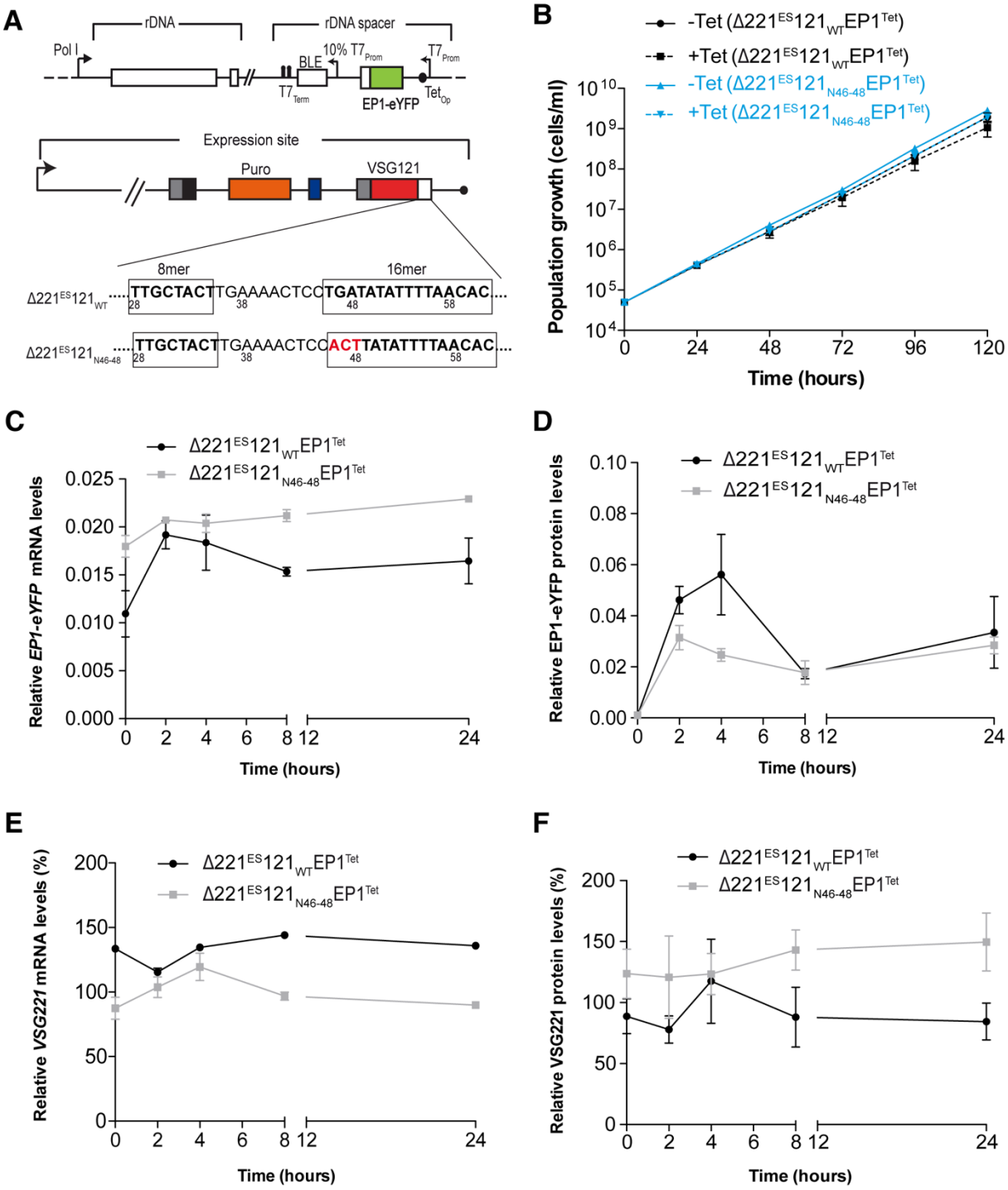

Figure 14. Inducible EP1-eYFP expression in $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46-48}$ single-expresser cell lines. (A) Illustration of the EP1-eYFP overexpression strategy used in the single-expresser cell lines. (B) Cumulative growth curve of $\Delta221^{ES}121_{WT}EP1^{Tet}$ and $\Delta221^{ES}121_{N46-48}EP1^{Tet}$ cell lines over the course of EP1-eYFP overexpression. Data are averages from three independent clones with error bars representing means ± standard error of the mean (SEM). *EP1-eYFP* mRNA (C) and EP1-eYFP protein (D) monitored in $\Delta221^{ES}121_{WT}EP1^{Tet}$ and $\Delta221^{ES}121_{N46-48}EP1^{Tet}$ cells over the course of EP1-eYFP overexpression. *EP1-eYFP* mRNA and EP1-eYFP protein were normalised to *tubulin* mRNA and PFR, respectively. Values are expressed as means ± standard error of the mean (SEM) of three independent clones. *VSG221* mRNA (E) and VSG221 protein (F) monitored in $\Delta221^{ES}121_{WT}EP1^{Tet}$ and $\Delta221^{ES}121_{N46-48}EP1^{Tet}$ cells during the course of EP1-eYFP

overexpression. The *VSG* expression levels are given relative to levels in the parental MITat1.6 cells and expressed as means ± standard error of the mean (SEM) of three clonal cells.

3.2.3 VSG silencing during switching does not require an intact 16mer in the active VSG 3´UTR

VSG switching can occur either by gene conversion, telomere crossover or *in situ* switch. During an *in situ* switching event, the active VSG expression site is silenced and the active VSG is replaced by a VSG on a previously silent expression site. This process therefore involves degradation of the previously active VSG, and hence reduced stability of the previously active VSG transcript. Since the 16mer motif is involved in stability of *VSG* mRNA (Ridewood et al., 2017), it was investigated whether the active VSG would require the intact 16mer motif for efficient VSG silencing during switching. Using an inducible overexpression system that mimics transcriptional VSG switching (Batram et al., 2014), VSG221 was integrated into the rDNA spacer in $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46\text{-}48}$ cells under a tetracycline-regulated T7 promoter to generate $\Delta221^{ES}121_{WT}221^{Tet}$ and $\Delta221^{ES}121_{N46\text{-}48}221^{Tet}$ cell lines, respectively (Figure 15A). Upon induction of overexpression with tetracycline, both cell lines maintained continuous growth. The induced $\Delta221^{ES}121_{WT}221^{Tet}$ cells however grew slightly slower than the uninduced cell population whiles the induced $\Delta221^{ES}121_{N46\text{-}48}221^{Tet}$ cells grew slightly faster in comparison to the uninduced cell population (Figure 15B). Analysis of the VSG expression levels showed that the expression site resident *VSG121* transcript and protein levels decreased with similar kinetics in both $\Delta221^{ES}121_{WT}221^{Tet}$ and $\Delta221^{ES}121_{N46\text{-}48}221^{Tet}$ cells (Figure 15C - F). Within 24 h of overexpression, the ectopic VSG221 became the most abundant VSG expressed, similar to what has been described in earlier studies (Batram et al., 2014; Zimmermann et al., 2017; Henning, Bachelor thesis 2012; Specht, Master thesis 2013). It was however observed that the $\Delta221^{ES}121_{N46\text{-}48}221^{Tet}$ cells silenced the active VSG slightly faster than the $\Delta221^{ES}121_{WT}221^{Tet}$ cells (Figure 15D, F). Absence of the 100 % conservation of the 16mer motif therefore appears to rather promote a much more rapid silencing effect within the first 8 hours of inducing overexpression. These data together suggests that the active VSG does not require an intact 16mer motif to be efficiently silenced during switching.

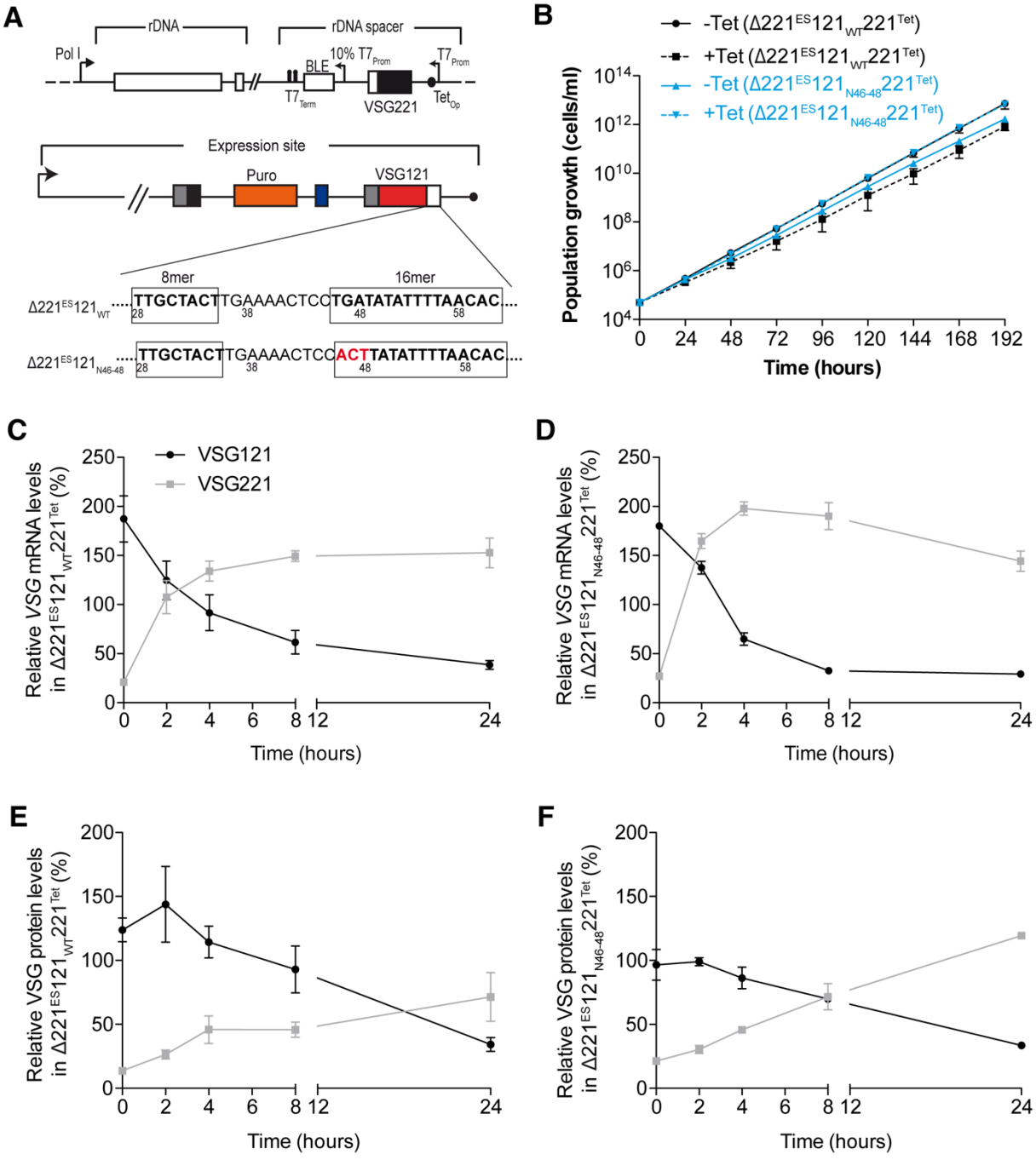

Figure 15. Inducible VSG221 expression in $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46\text{-}48}$ single-expresser cell lines. (A) Schematic of the ectopic VSG overexpression system used in the single-expresser cell lines $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46\text{-}48}$. (B) Cumulative growth curves of $\Delta221^{ES}121_{WT}221^{Tet}$ and $\Delta221^{ES}121_{N46\text{-}48}221^{Tet}$ cell lines over the course of VSG221 overexpression. *VSG* mRNA (C, D) and VSG protein (E, F) monitored in $\Delta221^{ES}121_{WT}221^{Tet}$ and $\Delta221^{ES}121_{N46\text{-}48}221^{Tet}$ cells during the course of VSG221 overexpression. *VSG* mRNA and VSG protein levels were quantified from dot blots and normalised to *tubulin* mRNA and PFR, respectively. The VSG expression levels are given relative to levels in the parental 13-90 (VSG221) or MITat1.6 (VSG121) cells, and expressed as means ± standard error of the mean (SEM) of three independent clones.

3.2.4 Overexpression of a second VSG requires an intact 16mer to trigger efficient VSG silencing

3.2.4.1 Overexpression of ectopic VSG121 with a mutated 16mer motif does not efficiently silence the endogenous VSG

Having established that the active VSG does not require a 100 % conserved 16mer motif to be efficiently silenced during an *in situ* switch, it was tested whether the previously silent VSG requires this high conservation to trigger silencing of the active VSG. Using the overexpression system mimicking an *in situ* switch, VSG121 with the tolerated 16mer mutation ($VSG121_{N46\text{-}48}$) was overexpressed from the rDNA spacer region in 13-90 cells (Figure 16A). As a control, VSG121 with the intact 16mer motif ($VSG121_{WT}$) was also overexpressed. Upon induction of overexpression of $VSG121_{WT}$, there was a significant reduction of cell proliferation after the first 24 h (Figure 16B). The cells stalled for about four days before resuming cell growth. This observation was similar to earlier observations made when VSG121 with the wild-type VSG121 3´ UTR was overexpressed (Batram et al., 2014). The growth phenotype was however different when $VSG121_{N46\text{-}48}$ was overexpressed. In these cells, although cell proliferation was reduced after 24 h, the reduction was not as pronounced as when $VSG121_{WT}$ was overexpressed (Figure 16B). Also, the cells did not stall but continued to grow, although significantly slower in comparison to the uninduced cells. This suggested that mutation of the 16mer motif produced a different effect on the cell cycle during overexpression, hence resulting in differences in the growth phenotypes between $VSG121_{WT}$ and $VSG121_{N46\text{-}48}$ overexpression.

Within the first 2 h of $VSG121_{WT}$ overexpression, there was a rapid increase in the ectopic *VSG121* mRNA to ~ 79 % of wild-type levels (Figure 16C). Between 2 to 8 h, a further increase in the mRNA levels to over 90 % was observed. The levels of *VSG121* mRNA then decreased slightly to ~ 88 % by 24 h and then to 45 % by 72h. The increase in the ectopic *VSG121* mRNA was accompanied by a decrease in the endogenous *VSG221* mRNA to ~ 73 % after 2 h with a further decrease to 32 % by 24 h. This was followed by an increase in the mRNA levels to 50 % by 72 h. The ectopic *VSG121* mRNA therefore became the predominant *VSG* mRNA expressed, with the endogenous *VSG221* mRNA decreasing to less than 35 % by 24h. At the protein level, VSG121 protein amounts increased to ~ 74 % of wild-type amounts by 24 h whereas VSG221 protein levels decreased to ~ 40 %. This was followed by a decrease in VSG121 and VSG221 to 53 % and 26 %, respectively, by 72 h. The endogenous VSG221 was therefore efficiently silenced at both the mRNA and protein levels, and the ectopic VSG121 protein became the predominantly expressed VSG. The kinetics of expression of the two VSGs observed here was similar to what was reported by Batram et al. (2014).

Overexpression of VSG121$_{N46-48}$ also resulted in a rapid increase in *VSG121* mRNA to ~ 90 % of wild-type levels within 2 h (Figure 16C). The mRNA levels increased further to ~ 120 % between 2 to 8 h, and then decreased to ~ 90 % by 24 h with a further decrease to 60 % by 72 h. The endogenous *VSG221* mRNA decreased to 55 to 60 % within the first 24 h with a further decrease to ~ 40 % by 72 h. This observation differed from the kinetics of silencing observed upon overexpression of VSG121$_{WT}$, where the endogenous *VSG* decreased to less than 35 % within the first 24 h. At the protein level, the ectopic VSG121 protein was not efficiently expressed, as only about half of the levels measured for VSG121$_{WT}$ overexpression was detected after 24 h (Figure 16D). The endogenous VSG221 protein was therefore not efficiently silenced and remained abundantly expressed at 59-80 %. These data together suggested that 100 % conservation of the 16mer is necessary to trigger efficient VSG silencing and coat exchange.

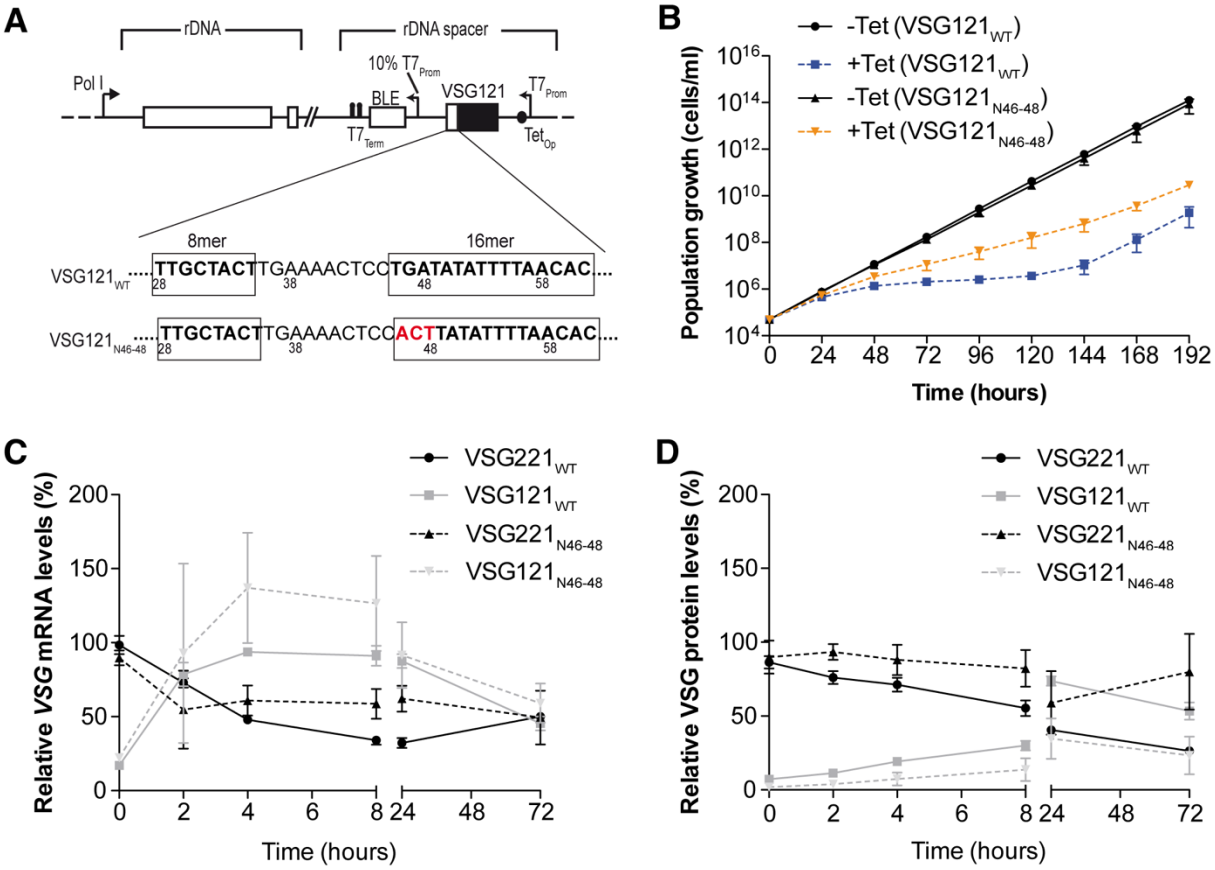

Figure 16. Inducible overexpression of VSG121 with an intact 16mer motif (VSG121$_{WT}$) and VSG121 with a mutated 16mer motif (VSG121$_{N46-48}$). (A) Schematic of the VSG121 overexpression strategy. (B) Cumulative growth curves after induction of ectopic VSG121$_{WT}$ and VSG121$_{N46-48}$ overexpression. *VSG mRNA* (C) and VSG protein (D) amounts monitored during the course of VSG121$_{WT}$ and VSG121$_{N46-48}$ overexpression. *VSG mRNA* and VSG protein levels were quantified from dot blots and normalised to *tubulin* mRNA and PFR, respectively. Values were expressed relative to VSG121 levels in MITat1.6 cells or relative to VSG221 levels in 13-90 cells, and are presented as means ± SEM of three independent clones.

3.2.4.2 Overexpression of ectopic VSG118 with a mutated 16mer motif causes distinct phenotypes

VSG118 with a mutant 16mer ($VSG118_{N46-48}$) was overexpressed from the rDNA spacer to rule out the possibility that the observed failure of $VSG121_{N46-48}$ to trigger efficient VSG silencing and coat exchange is VSG-specific (Figure 17A). As a control, VSG118 with the intact 16mer motif ($VSG118_{WT}$) was overexpressed. Ectopic overexpression of $VSG118_{WT}$ produced clones with different growth phenotypes. One set of clones ($VSG118_{WT}$-Fast) grew continuously but slightly slower than the uninduced cells, while the other set ($VSG118_{WT}$-Slow) drastically slowed down in growth between 24 and 120 hours (Figure 17B; Henning, Bachelor thesis 2012; Specht, Master thesis 2013). Overexpression of $VSG118_{N46-48}$ also produced clonal cells with two distinct growth phenotypes. One set of clones ($VSG118_{N46-48}$-Fast) grew slightly faster than the uninduced cells while the other set ($VSG118_{N46-48}$-Slow) grew slower. Mutation of the 16mer motif in VSG118 affected the growth phenotype as overexpression of $VSG118_{N46-48}$ produced clonal cell lines that grew better than those obtained from $VSG118_{WT}$ overexpression. It was also observed that silencing of the endogenous VSG was fast and efficient in the fast-growing clones compared to the slow-growing clones upon overexpression of $VSG118_{N46-48}$ (Figure 17C - F). *VSG221* mRNA in the fast growing clones decreased to 36 % within the first 6 h of induction and decreased further to 32 % after 24 h (Figure 17C). This was also reflected in the protein amounts as VSG221 protein decreased to 31 % by 24 h (Figure 17E). In the slow clones however, *VSG221* mRNA was still above 50 % after 6h of induction and only decreased to 42 % by 24 h (Figure 17D). The VSG221 protein levels were at ~ 50 % by 24 h (Figure 17F). This was in contrast to results obtained upon overexpression of $VSG118_{WT}$ and other wild-type VSGs where VSG silencing and coat exchange occurs efficiently and with similar kinetics regardless of the growth phenotype (Figure 17C – F) (Batram et al., 2014; Henning, Bachelor thesis 2012; Specht, Master thesis 2013; Zimmermann et al., 2017). *VSG221* mRNA levels were less than 25 % after 24 h of $VSG118_{WT}$ overexpression in both fast-growers and slow-growers (Figure 17C, D). The protein levels also decreased to less than 35 % regardless of the growth phenotype (Figure 17E, F). It was previously shown that the ES activity influences the growth phenotype obtained upon overexpression of an ectopic VSG (Zimmermann et al., 2017). Slower growing cells generally have a strongly attenuated ES compared to fast continuously growing cells. As the active VSG221 was silenced more efficiently in the slower growing cells compared to the fast-growing cells, the data suggest that the role of the intact 16mer in triggering efficient ES-resident VSG silencing is somewhat dependent on the degree of attenuation of the ES.

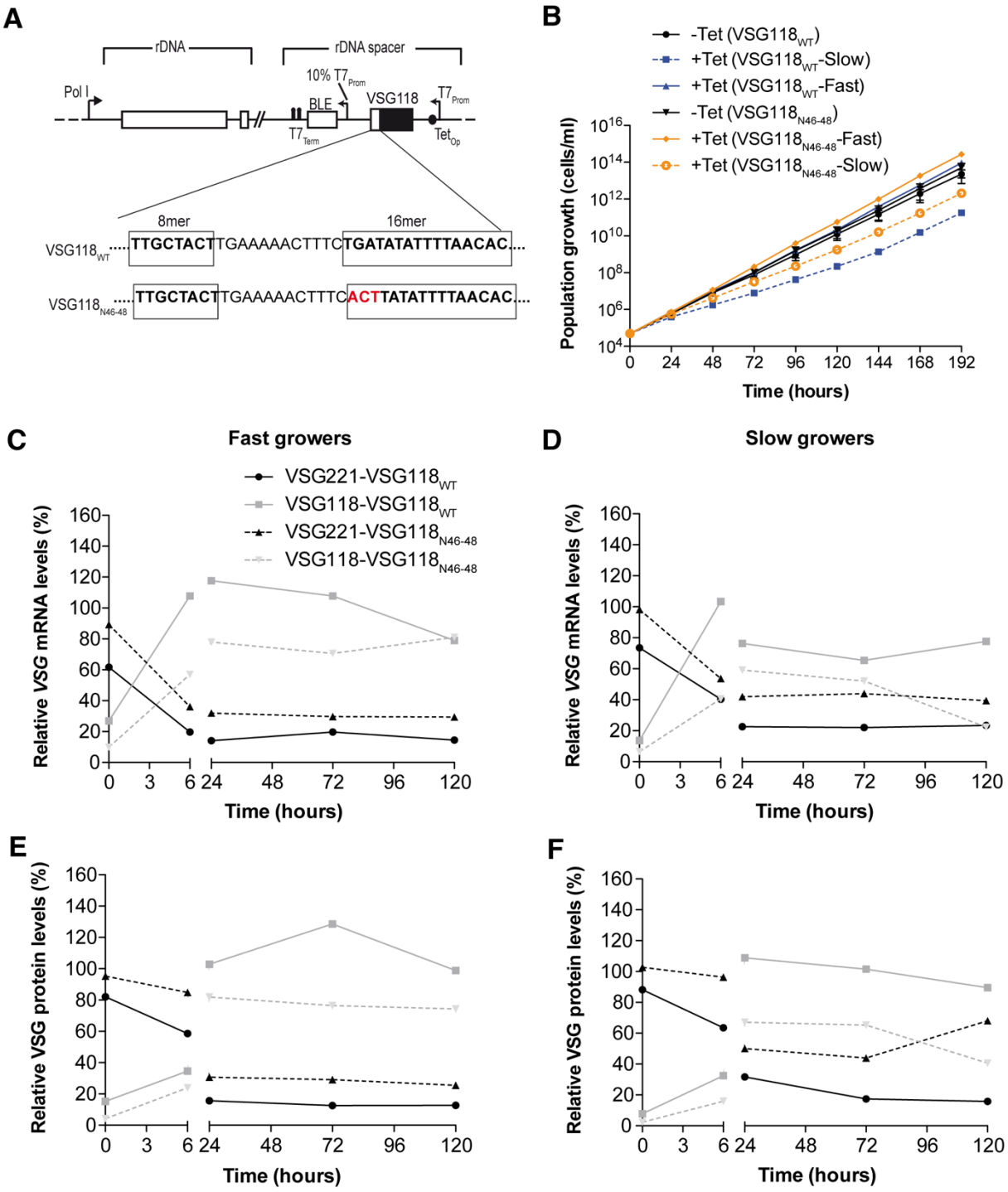

Figure 17. Inducible overexpression of VSG118 with an intact 16mer motif (VSG118$_{WT}$) and VSG118 with a mutated 16mer motif (VSG118$_{N46-48}$). (A) Schematic of the VSG118 overexpression strategy. (B) Cumulative growth curve after induction of ectopic VSG118$_{WT}$ and VSG118$_{N46-48}$ overexpression. Two distinct clonal populations (fast growers and slow growers) were observed. The plot for uninduced (-Tet) cells in VSG118$_{N46-48}$ overexpression is from four independent clones and plots for the induced (+Tet) cells are from two clones each. For VSG118$_{WT}$ three clonal cell lines (2 slow growers and 1 fast grower) were analysed. Values for the uninduced cells are represented as means ± SEM. *VSG* mRNA levels in fast growing clones (C) and slow growing clones (D) during ectopic overexpression of VSG118$_{WT}$ and VSG118$_{N46-48}$. VSG protein levels in fast growing clones (E) and slow growing clones (F)

during ectopic overexpression of VSG118$_{WT}$ and VSG118$_{N46-48}$. *VSG* mRNA and VSG protein levels were quantified from dot blots and normalised to *tubulin* mRNA and PFR, respectively. Values were expressed relative to VSG118 levels in MITat1.5 wild-type cells or relative to VSG221 levels in 13-90 cells, and are presented as means.

3.3 Identification of *trans*-acting factors interacting with the 16mer motif

3.3.1 Potential interaction partners of the 16mer DNA

In order to determine the functional significance of the 16mer motif, it was necessary to identify interacting partners of the motif. It was therefore hypothesised that the high conservation of the motif could be due to the motif acting as the recognition site for a DNA-binding domain and thereby influencing gene expression levels. To identify, potential interaction partners of the motif, a pull-down screen was carried out on a whole cell lysate using the VSG121 3´ UTR DNA sequence with either the intact 16mer motif (as bait) or a completely scrambled 16mer sequence (as control). The results were analysed by mass spectrometry. These experiments were carried out in collaboration with the group of Falk Butter (Institute of Molecular Biology, Mainz). The results of the pull-down assay are shown in the volcano plot in Figure 18.

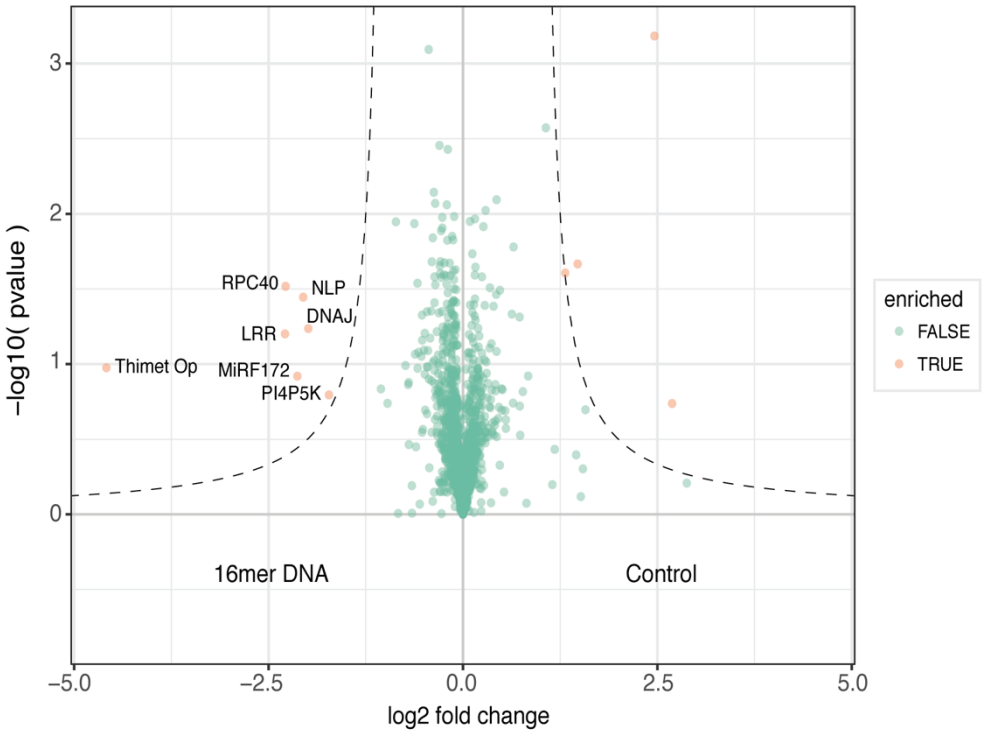

Figure 18. Volcano plot displaying interaction partners of the 16mer DNA. A pull-down assay was carried out on a whole cell lysate using a VSG121 3´ UTR DNA sequence with either the intact 16mer motif or scrambled 16mer motif (control). Eluted proteins from the pull-down were analysed by mass spectroscopy. Significantly enriched proteins are shown in orange dots.

The x-axis of the volcano plot shows the log2 fold difference between the mean of peptide counts found in the intact 16mer DNA pull-down and the mean of peptide counts found in the control (scrambled 16mer DNA). The y-axis showing the -log10(p-value) indicates the statistical significance of enrichment of the identified proteins. Statistically, seven proteins were significantly enriched in the intact 16mer DNA pull-down. They were Tb927.11.1960 (leucine rich repeat protein (LRR)), Tb927.7.190 (thimet oligopeptidase (Thimet Op)), Tb927.10.4770 (phosphatidylinositol-4-phosphate 5-kinase (PI4P5K)), Tb927.10.2290 (Chaperone protein DNAJ (DNAJ)), Tb927.11.13080 (protein with nuclear localisation (NLP)), Tb927.10.15370 (DNA-directed RNA polymerases I and III subunit RPAC1 (RPC40)) and Tb927.3.2050 (Minicircle replication factor 172 ((MiRF172)).

3.3.1.1 Effect of loss of potential 16mer DNA binding proteins on cell growth, cell cycle and VSG expression

Since the 16mer motif is essential for VSG expression, it was predicted that proteins binding to this motif and regulating gene expression at the DNA level would be essential. Therefore as a preliminary screen, the effect of the loss of the candidate 16mer DNA binding proteins on the phenotype of cells and VSG expression was investigated. Inducible RNAi cell lines were generated for each protein in order to knockdown each of the candidate proteins. The growth phenotype and cell cycle profile of the cells were then analysed upon induction of RNAi with tetracycline. It was observed that upon induction of RNAi, PI4P5K and Thimet Op did not impact the growth of the cells (Figure 19B, C). These two proteins are therefore probably not essential in bloodstream form *T. brucei* cells. For these two proteins, there was also no change or aberration in the cell cycle profile upon RNAi-mediated depletion of the proteins (Figure 20B, C). Knockdown of NLP, MiRF172, LRR and DNAJ and RPC40 all resulted in reduced cell proliferation (Figure 19A, D - G). Loss of RPC40, LRR, DNAJ and NLP caused a drastic reduction in growth whereas a slight reduction in growth was observed in MiRF172 only after 72 h of RNAi induction.

Cell cycle analysis showed that upon depletion of NLP, a slight decrease in cells with 2K2N configuration and a slight increase in the percentage of cells with aberrant cell cycle configuration (others) occurred (Figure 20A). Depletion of both LRR and DNAJ resulted in a decrease in cells with the cell cycle configuration 1K1N, 1Kd1N and 2K1N, and an increase in cells with an aberrant cell cycle configuration (Figure 20E, F). Loss of RPC40 also resulted in a decrease in the number of cells with 1K1N and 1Kd1N and an increase in cells with an aberrant cell cycle configuration (Figure 20G). Following depletion of MiRF172, a decrease in 2K1N cells and an increase in 0K1N cells was observed after 24 h (Figure 20D). The growth

phenotype and cell cycle profile obtained upon depletion of MiRF172 was similar to observations reported by Amodeo et al., (2018).

With analysis of the growth phenotype and cell cycle after depletion of the candidate 16mer DNA binding proteins, it was found that NLP, MiRF172, LRR and DNAJ and RPC40 are all essential proteins in bloodstream *T. brucei* cells. While this study was ongoing, MiRF172 was characterised as a protein involved in the reattachment of replicated minicircles to the kDNA disc (Amodeo et al., 2018). RPC40 on the other hand is annotated as an RNA polymerases I and III subunit and is therefore most likely essential for RNA pol-I and III transcription. This is evident in the observation that its depletion affected global protein levels (Appendix: Figure 31). These two proteins were therefore eliminated from further analysis of the impact of the loss of the essential candidate 16mer DNA binding proteins on VSG expression.

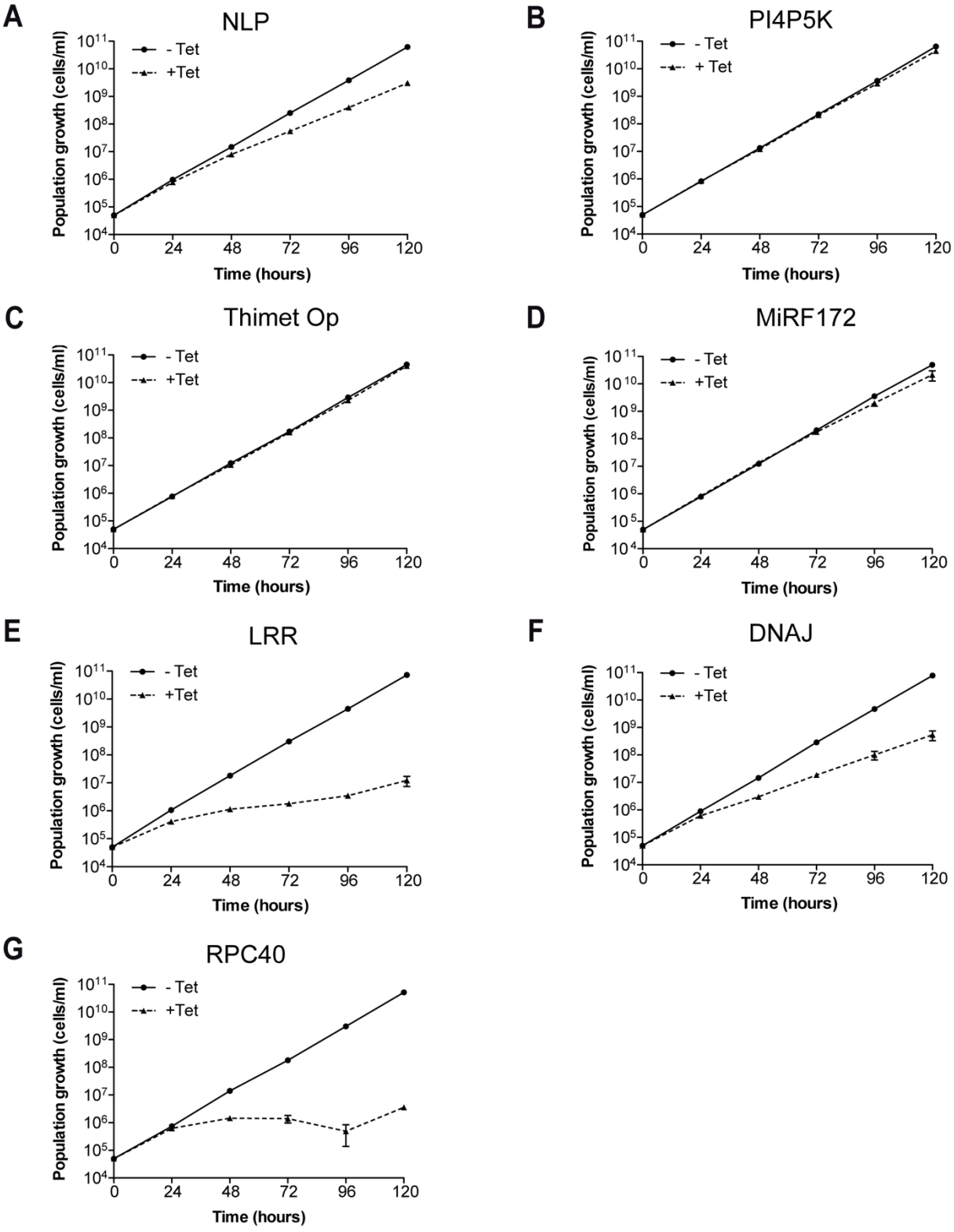

Figure 19. Effect of knock-down of candidate 16mer DNA binding proteins on growth of BSF *T. brucei* cells. Cumulative growth curves upon induction of RNAi against NLP (A), PI4P5K (B), Thimet Op (C), MiRF172 (D), LRR (E), DNAJ (F) and RPC40 (G). Data are averages from at least two independent clones with error bars representing means ± standard error of the mean (SEM).

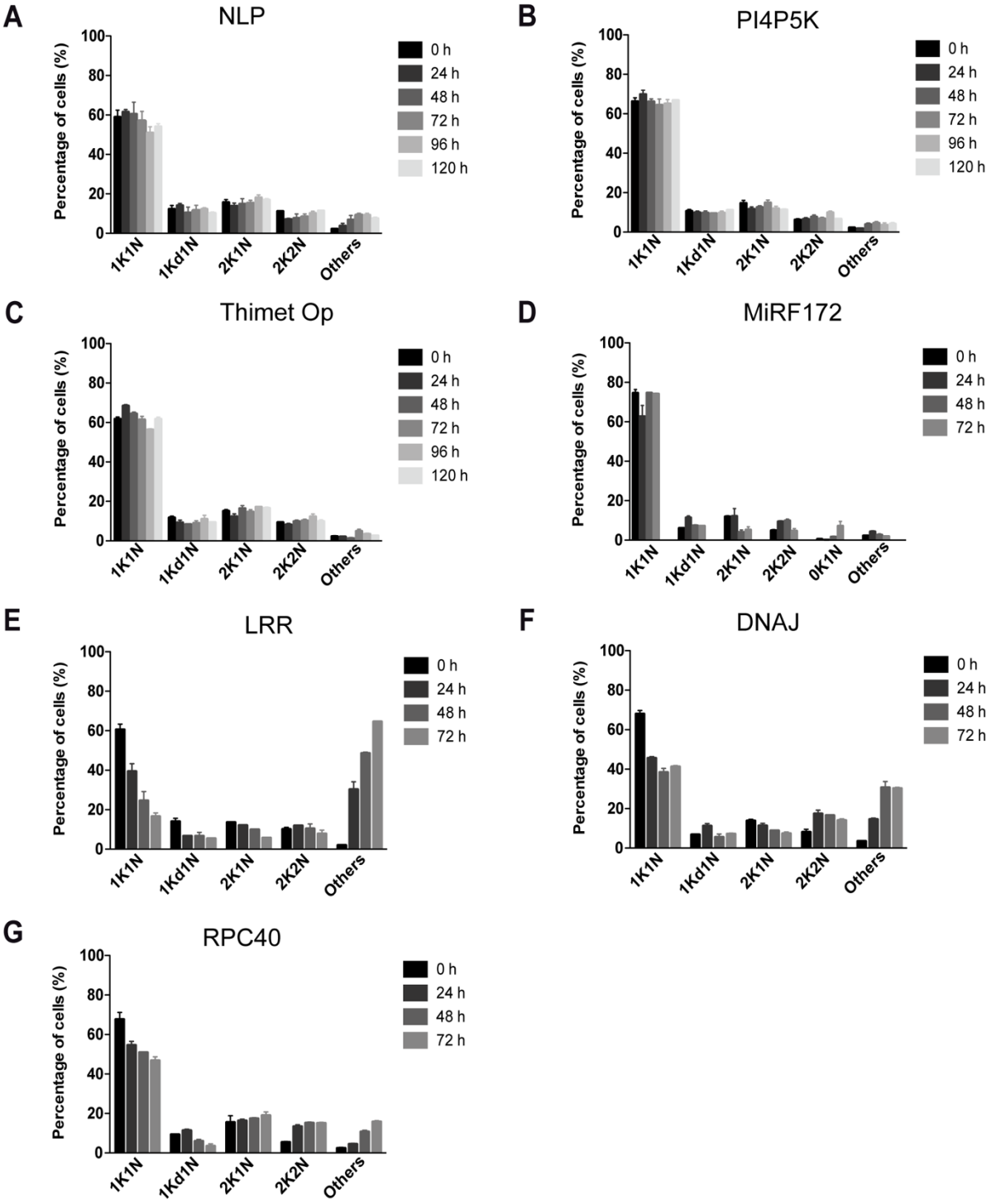

Figure 20. Effect of knock-down of candidate 16mer DNA binding proteins on cell cycle distributions of BSF *T. brucei* cells. Cell cycle distribution upon induction of RNAi against NLP (A), PI4P5K (B), Thimet Op (C), MiRF172 (D), LRR (E), DNAJ (F) and RPC40 (G). Cells were stained with DAPI and classified to the different cell cycle stages according to the number of nuclei (N) and kinetoplasts (K) present. At least 200 cells were analysed for each time point. At least two independent clones were analysed, and data presented as mean percentages ± SEM.

Analysis of the VSG221 expression upon depletion of NLP, LRR and DNAJ showed that high levels of the VSG221 were still present after 48 h of RNAi induction (Figure 21). Analysis of the overall phenotypic effect was therefore not sufficient for evaluating the functions of the potential 16mer DNA binding proteins. The absence of a drastic reduction in the VSG levels upon depletion of the proteins could also indicate that these proteins are not directly involved in the regulation of VSG expression at the DNA level.

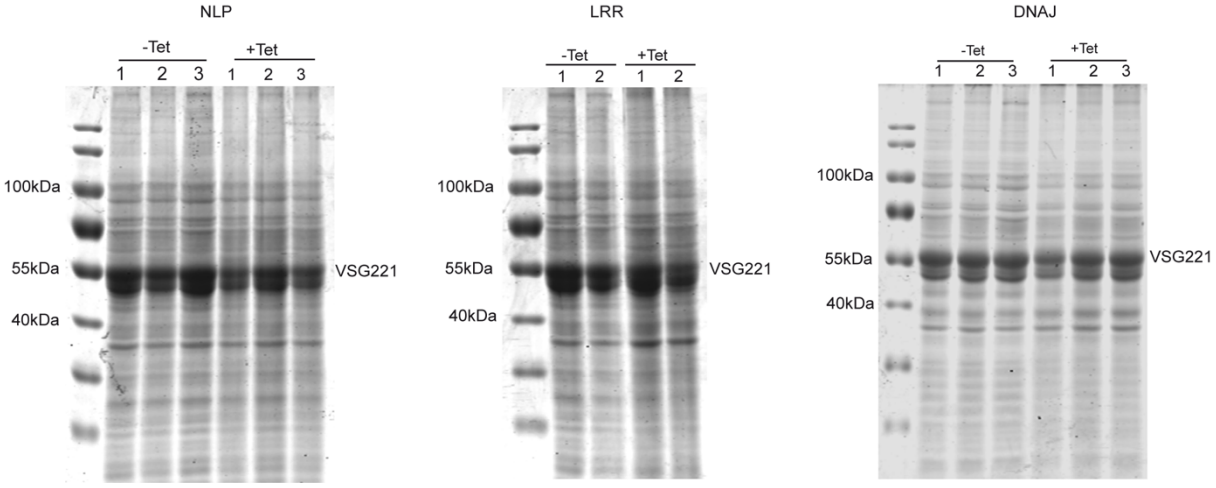

Figure 21. Effect of knock-down of NLP, LRR and DNAJ on VSG expression. Protein samples prepared from cells pre- and 48 h post-induction of RNAi were resolved on SDS-polyacrylamide gels and stained with Coomassie. The VSG221 monomer is an ~ 55 kDa protein that is always the major band in Coomassie stained gels of BSF *T. brucei* cells expressing this VSG.

3.3.2 Potential interaction partners of the 16mer RNA

RNA binding proteins can modulate the expression and abundance of specific transcripts through their interaction with specific motifs within 3′ UTRs. Therefore, besides the identification of potential 16mer DNA binding proteins, candidate proteins interacting with the 16mer RNA were also identified using pull-down assays. The pulldown assay was carried out using biotinylated RNA of the VSG121 3′ UTR with the intact 16mer and 8mer motifs. The RNA coupled to streptavidin beads was used to pull-down proteins from whole cell lysate and the proteins were analysed by mass spectrometry. As controls, biotinylated RNA of the VSG121 3′ UTR with either scrambled 16mer and 8mer motifs (control I) or a reversed 16mer motif (control III) and a reverse complement of control I (control II) were used. In all, fourteen proteins were enriched when the VSG121 3′ UTR with intact 16mer and 8mer was used as bait versus the three controls (Table 11). Tb927.10.1720 was the only protein enriched in the bait versus all three control RNAs (the scrambled 16mer and 8mer, the reverse complement of the scrambled 16mer and 8mer and the reverse complement of the 16mer) (Figure 22; Table 11). Tb927.9.1350 was also enriched in the bait versus controls I and III. The rest of the proteins

were only enriched in the bait versus one control. Tb927.10.1720 therefore appears to be the most likely 16mer RNA binding protein. Since the function of this protein was unknown, I proceeded with characterising the protein. Data on its characterisation are presented and described in section 3.4. The experiments for the RNA pull-down and mass spectrometry analysis were carried out in collaboration with the group of Falk Butter (Institute of Molecular Biology, Mainz).

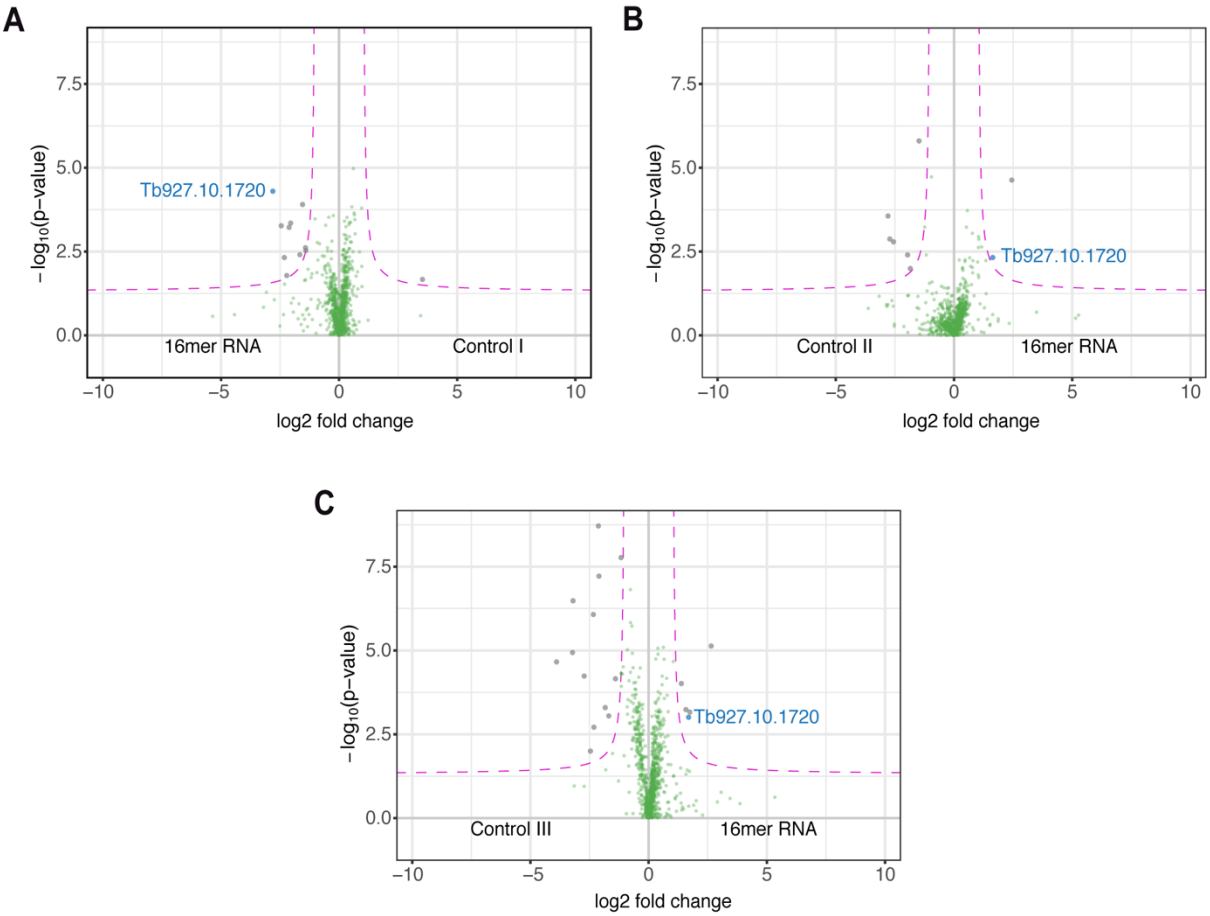

Figure 22. Identification of candidate 16mer RNA binding proteins. The first 188 bp stretch of VSG121 3´UTR harbouring intact 16mer and 8mer motifs was *in vitro* transcribed and used as bait. The results of the mass spectrometry analysis of the bait-interacting proteins are shown in volcano plots. Enriched proteins (in grey dots) in bait versus control I (first 188 nucleotides of the VSG121 3´ UTR with scrambled 16mer and 8mer) (A), control II (reverse complement of control I) (B) and control III (first 188 nucleotides of the VSG121 3´ UTR with reverse complement of 16mer) (C). The figure is adapted from Bakari-Soale et al. (2021).

Table 11. Proteins enriched in pull-down with VSG121 3´ UTR RNA with intact 16mer motif (Bait) versus controls

Protein ID	Control I	Control II	Control III	Enriched count	Annotation in TriTrypDB
Tb927.10.1720	+	+	+	3	ATP-dependent DEAD/H RNA helicase, putative
Tb927.7.6290	+			1	Kinesin, putative
Tb927.4.4860	+			1	Amino acid transporter 8, putative
Tb927.3.1680	+			1	Eukaryotic translation initiation factor 3 subunit F
Tb927.9.1350	+		+	2	Hypothetical protein, conserved
Tb927.6.700	+			1	Alanyl-tRNA synthetase, putative
Tb927.4.2840	+			1	Hypothetical protein, conserved
Tb927.3.2490	+			1	Hypothetical protein, conserved
Tb927.10.10880	+			1	ATP-binding cassette sub-family F member 1, putative
Tb927.5.1120	+			1	Phage tail fibre repeat, putative
Tb927.10.14390		+		1	FACT complex subunit POB3
Tb927.9.1600			+	1	Hypothetical protein, conserved
Tb927.10.3810			+	1	Nucleoporin NUP65
Tb927.5.1940			+	1	Hypothetical protein, conserved

3.4 Characterisation of the DExD/H box helicase Hel66

3.4.1 Bioinformatic analysis of Hel66

Tb927.10.1720 encodes an approximately 66 kDa protein consisting of 596 amino acids and is conserved in kinetoplastids. The protein is annotated as a putative ATP-dependent DEAD/H RNA helicase in the TriTrypDB, I therefore referred to the protein as Hel66 (helicase, 66 kDa) in this study. Bioinformatic analysis was carried out on Hel66 to confirm the presence of the characteristic motifs of the DEAD box protein family (Figure 23A). A multiple sequence alignment of Hel66 to previously-characterised DEAD box proteins in *Trypanosoma brucei*, *Saccharomyces cerevisiae*, *Leishmania braziliensis* and *Homo sapiens* showed that the motifs characteristic of the DEAD box proteins are highly conserved in Hel66 (Figure 23B). Motif II/Walker B motif (DEAD box) in Hel66 however has a valine (V) residue in place of alanine (A). Amino acid substitutions at this position have been observed in some members of this class of proteins sometimes designated as DExD helicases (Tanner & Linder, 2001). Like most proteins belonging to the DExD/H box family, the N- and C- terminal ends of Hel66 are less conserved. These regions confer functional specificity to members of this protein family. BLAST (Basic Local Alignment Search Tool) analysis was carried out to identify homologues of Hel66. Outside the Kinetoplastida, the closest BLAST hit to Hel66 is DDX21 in the duck *Asarcornis scutulata* (E-value: 2×10^{-33}, percentage identity: 27.51 %, Query cover: 69 %). A reverse BLAST of DDX21 however does not return Hel66 as the top hit. This suggests that Hel66 is a kinetoplastid specific protein.

3.4.2 Hel66 does not interact with the 16mer motif

As Hel66 was identified as the most likely interaction partner of the 16mer motif from the RNA pull-down assay and mass spectrometry analysis, I wanted to confirm that Hel66 directly binds to the 16mer motif. For this, an RNA electrophoretic mobility shift assay (RNA EMSA) was used to investigate the interaction in vitro. Together with Nonso J. Ikenga, recombinant GST-tagged Hel66 (GST-Hel66) was expressed and purified from ArcticExpress *E. coli* cells (Ikenga, Master thesis 2020). I carried out RNA EMSA using the purified recombinant GST-Hel66 and biotinylated 16mer RNA. When the 16mer RNA was mixed with the GST-Hel66, no band shift was detected (Figure 24A). The 16mer RNA was only detected as free probe. To check whether the concentration of GST-Hel66 used was affecting binding to the 16mer RNA, the experiment was repeated with varying concentrations of the protein but again no band shift was detected (Figure 24B). A band shift was however always observed when the positive control from the RNA EMSA kit was used (Figure 24C), suggesting functional interaction

between the IRE (iron-response element) RNA and IRP (iron-response protein). This suggested that the absence of a band shift when GST-Hel66 was mixed with 16mer RNA is a result of no functional interaction or binding between the two components.

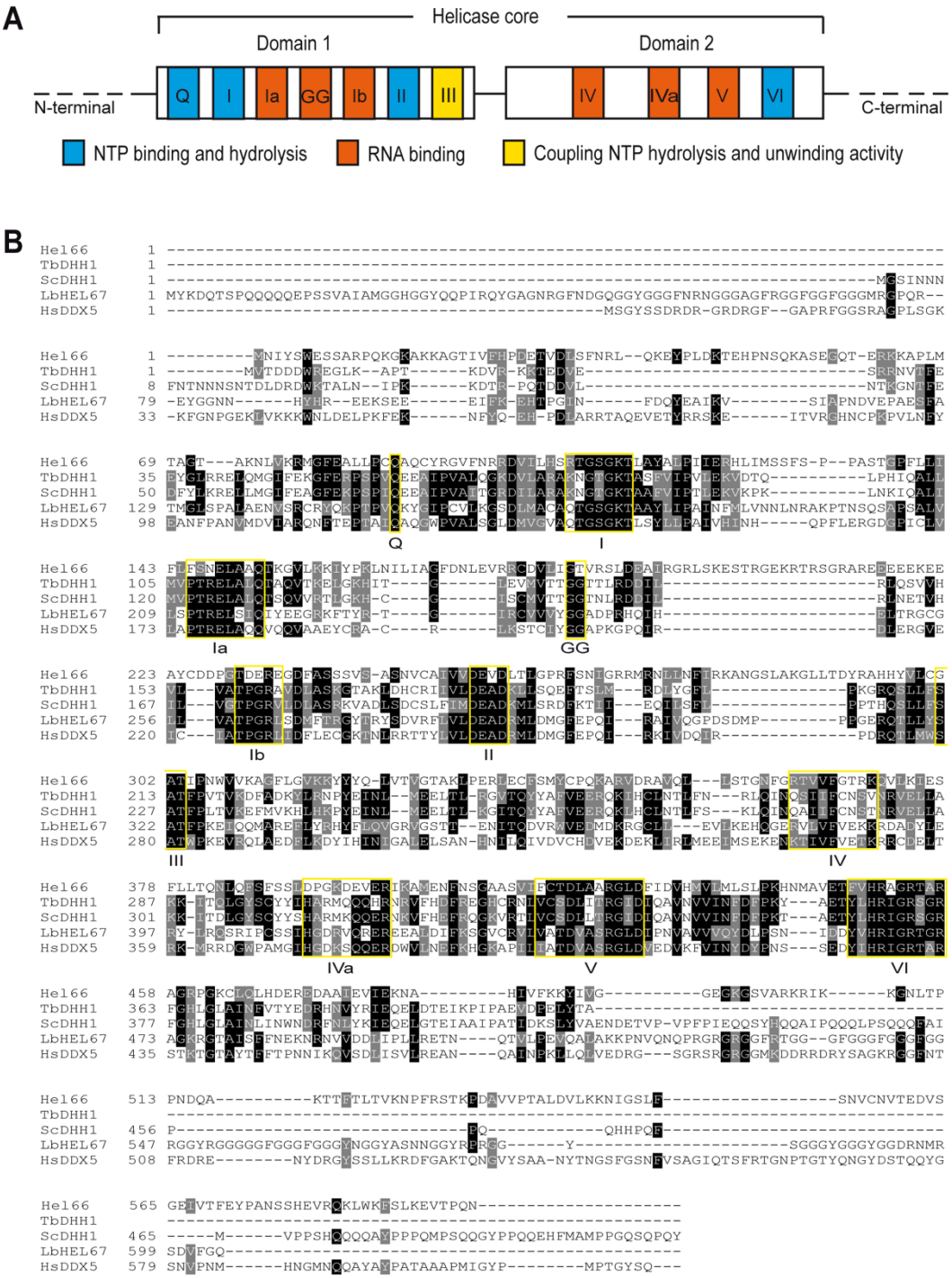

Figure 23. Bioinformatic analysis to identify conserved motifs in Hel66. (A) Illustration of the characteristic motifs present in proteins of the DExD/H box family. (B) Multiple sequence alignment of Hel66 to some previously-characterised DEAD box helicase proteins in *Trypanosoma brucei* (TbDHH1), *Saccharomyces cerevisiae* (ScDHH1), *Leishmania braziliensis* (LbHEL67) and *Homo sapiens* (HsDDX5). The conserved motifs of the DExD/H proteins are shown in yellow boxes. Identical and similar amino acid residues are shaded in black and grey, respectively. The multiple sequence alignment was performed with the online

tool, Clustal Omega (https://www.ebi.ac.uk/Tools/msa/ clustalo/). The output file from the alignment was formatted for visualisation with BoxShade (https://embnet.vital-it.ch/software/BOX_form.html). The figure is adapted from Bakari-Soale et al. (2021).

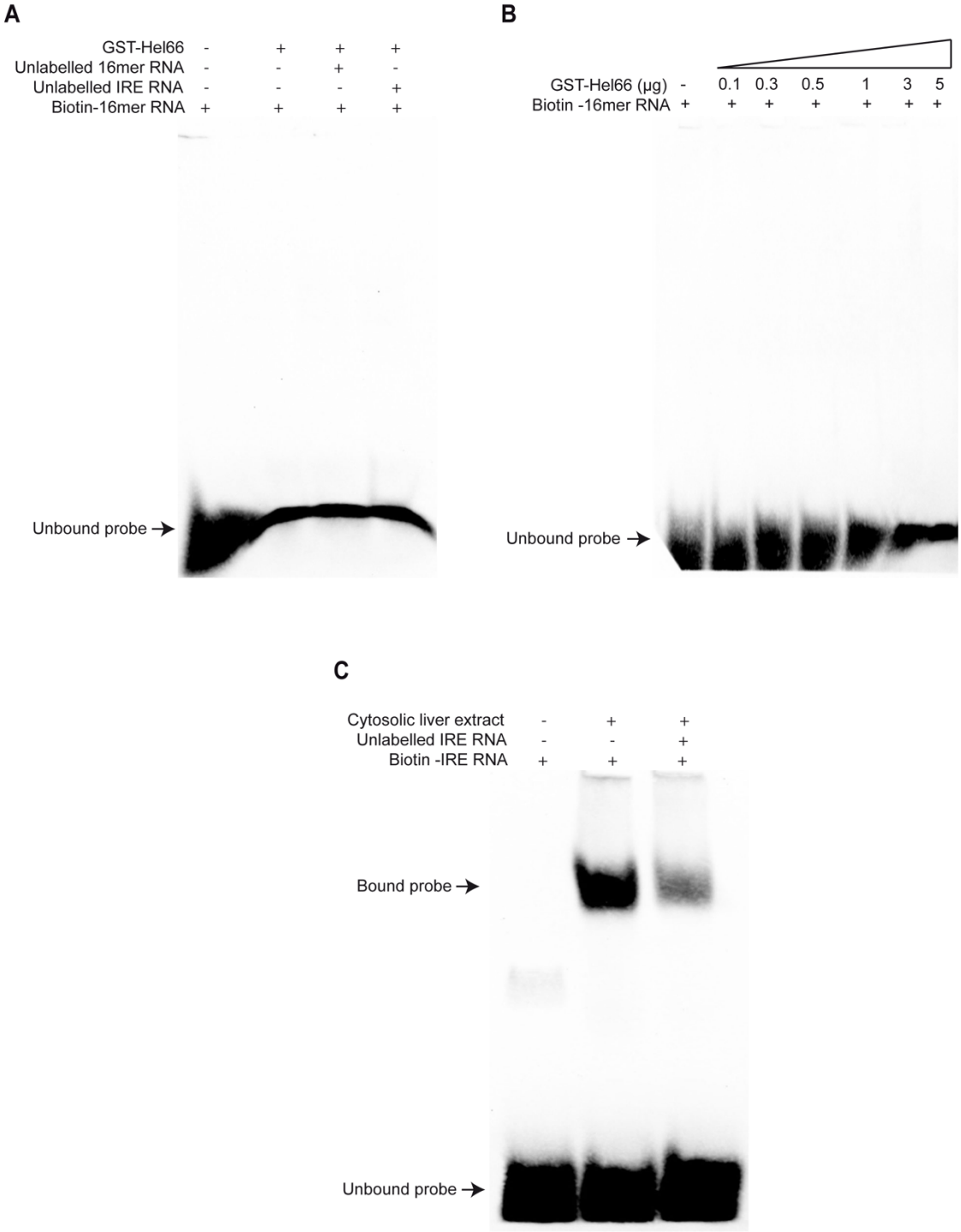

Figure 24. RNA EMSA to investigate direct interaction between GST-Hel66 and 16mer RNA. (A) No interaction between 16mer RNA and GST-Hel66. RNA EMSA with 16mer RNA and GST-Hel66. 5 µg of GST-Hel66 was mixed with 10 nM of biotinylated 16mer RNA in a reaction mix of 20 µl and incubated at room temperature for 30 min. A 200-fold excess of either unlabelled 16mer RNA or an unrelated RNA (IRE RNA) was added for competition assays. (B) No interaction between 16mer RNA and varying concentrations of GST-Hel66. RNA EMSA with the 16mer RNA and different concentrations of GST-Hel66 protein. Varying

concentrations of GST-Hel66 were mixed with 10 nM of biotinylated 16mer RNA in a reaction mix of 20 µl and incubated at room temperature for 30 min. (C) Positive control from the RNA EMSA kit showing functional binding of IRP (iron-response protein) to the IRE (iron-response element). The figure is adapted from Bakari-Soale et al. (2021).

3.4.3 Localisation of Hel66 in BSF trypanosomes

The subcellular localisation of a protein influences the function of the protein in the cell. Therefore, in the characterisation of proteins with unknown functions, it is necessary to determine the subcellular localisation since this can provide vital information about protein function. The subcellular localisation of Hel66 was therefore determined as a first step in the functional characterisation of the protein. Hel66 was previously identified in the nuclear proteome of procyclic *T. brucei* parasites (Goos et al., 2017). The genome wide localisation project TrypTag found that Hel66 localises specifically to the nucleolus in procyclic trypanosomes (Dean et al., 2017). As these previous studies focused on procyclic form trypanosomes and this particular study involved bloodstream form cells, the subcellular localisation of Hel66 in the bloodstream form *T. brucei* cells was determined. A fluorescent-tagging approach was used to avoid the need for generation of Hel66 specific antibodies. Hel66 was thus tagged *in situ* at the C-terminus with mNeonGreen. The endogenously tagged protein was found to localise to the nucleolus, the region of the nucleus with less intense DAPI staining (Figure 25A, B). The subcellular localisation of Hel66 in the bloodstream trypanosomes was in agreement with earlier observations made in the procyclic trypanosomes. The localisation of the protein is therefore not stage-specific, a possible indication that the protein performs a similar function in both bloodstream and procyclic trypanosomes. Also, since the nucleolus plays a key role in transcription and processing of rRNAs for assembly of ribosomes, Hel66 may be involved in either of these processes.

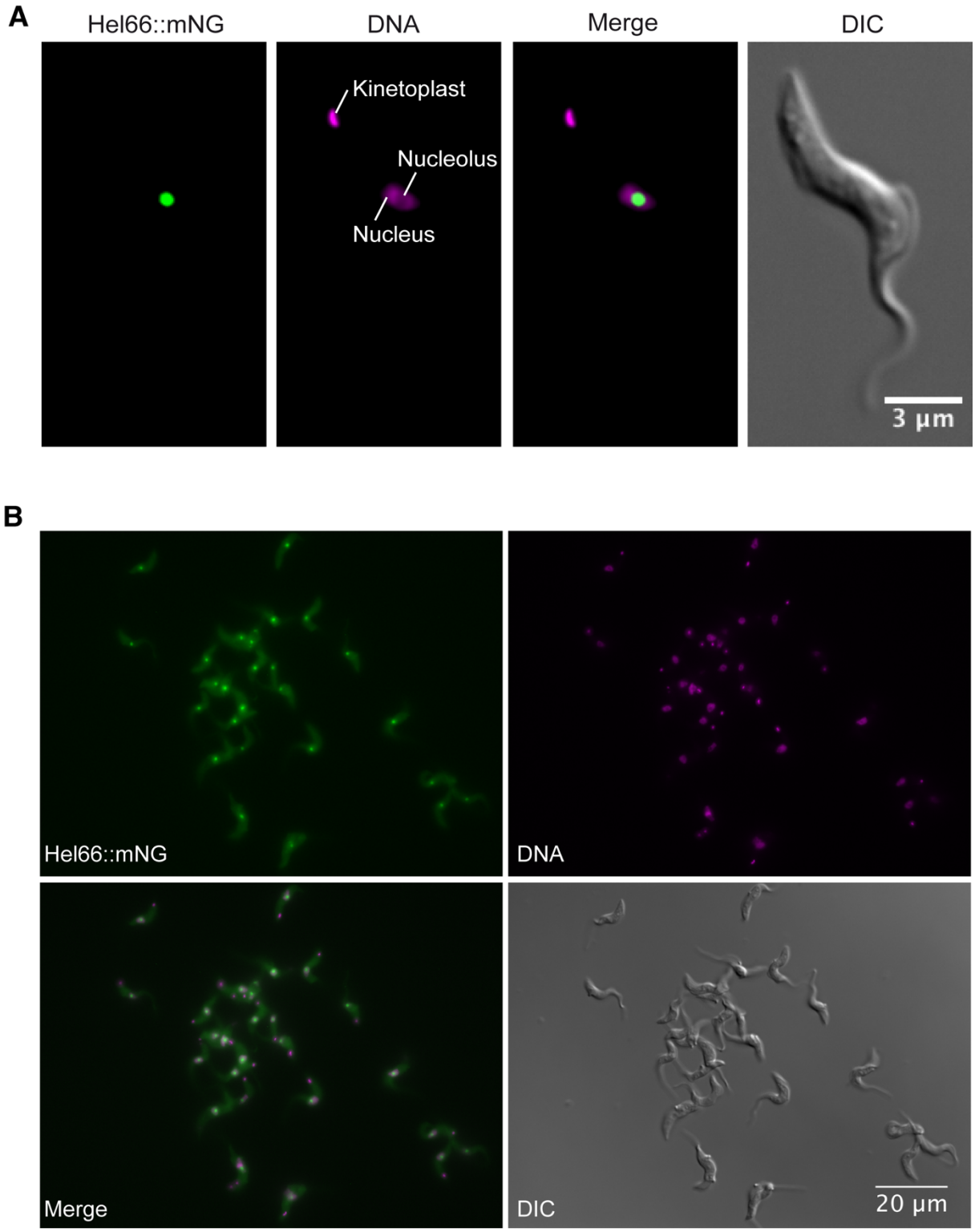

Figure 25. Hel66 localises to the nucleolus in BSF *T. brucei* cells. (A) A BSF *T.brucei* cell expressing endogenously tagged Hel66 (Hel66::mNeonGreen). (B) Population of cells expressing Hel66::mNG. The trypanosome cells with endogenously tagged Hel66 (Hel66::mNeonGreen) were fixed, DAPI-stained and imaged. The signal from the tagged Hel66 was concentrated in the nucleolus, the region of the nucleus with less intense DAPI stain. Images in this figure are adapted from Bakari-Soale et al. (2021).

3.4.4 RNAi-mediated depletion of Hel66 affects cell growth and cell cycle profile

To further investigate the function of Hel66, the protein was depleted via RNAi in a cell line expressing endogenous C-terminal HA-tagged Hel66 (Hel66::HA). The effect of the loss of protein on cell proliferation and the cell cycle was then analysed. Upon RNAi induction, Hel66::HA was efficiently depleted as the amount of the protein decreased to about 20 % of the levels in uninduced cells by 24 h with a further decrease to less than 10 % by 48 h of RNAi induction (Figure 26A,B). Silencing of Hel66 expression resulted in slower growth of cells with cells stalling after 24 h of RNAi induction (Figure 26C). The growth phenotype was investigated for a possible role of the protein in cell cycle control. Using DAPI staining, an analysis of the nuclei and kinetoplast (NK) configuration revealed a small increase in the number of cells in the pre-cytokinesis stage (2K2N) and cells having aberrant NK configuration after 48 h of RNAi induction (Figure 26D). In the population of cells with aberrant NK configuration, cells with multiple nuclei and in some cases multiple kinetoplasts were observed (Figure 26E). The changes in the NK configurations, especially the accumulation of 2K2N cells, are suggestive of a defect in cytokinesis.

3.4.5 Loss of Hel66 has no impact on VSG expression

Although the RNA EMSA experiments (section 3.4.2) failed to demonstrate a direct interaction between the 16mer RNA and Hel66, a possible role of Hel66 in VSG expression was investigated since the protein was identified from a pull-down assay using the VSG 3´ UTR as bait. The effect of Hel66 depletion on *VSG* mRNA and VSG protein levels was analysed. No significant changes in *VSG221* mRNA and VSG221 protein were observed after 48 h of Hel66 depletion (Figure 27A,B). The VSG mRNA and protein levels remained at approximately wild-type amounts. The data therefore suggests that Hel66 has no direct effect on VSG expression. The interaction leading to the identification of the protein from the pull-down and mass spectrometry analysis was therefore probably unspecific.

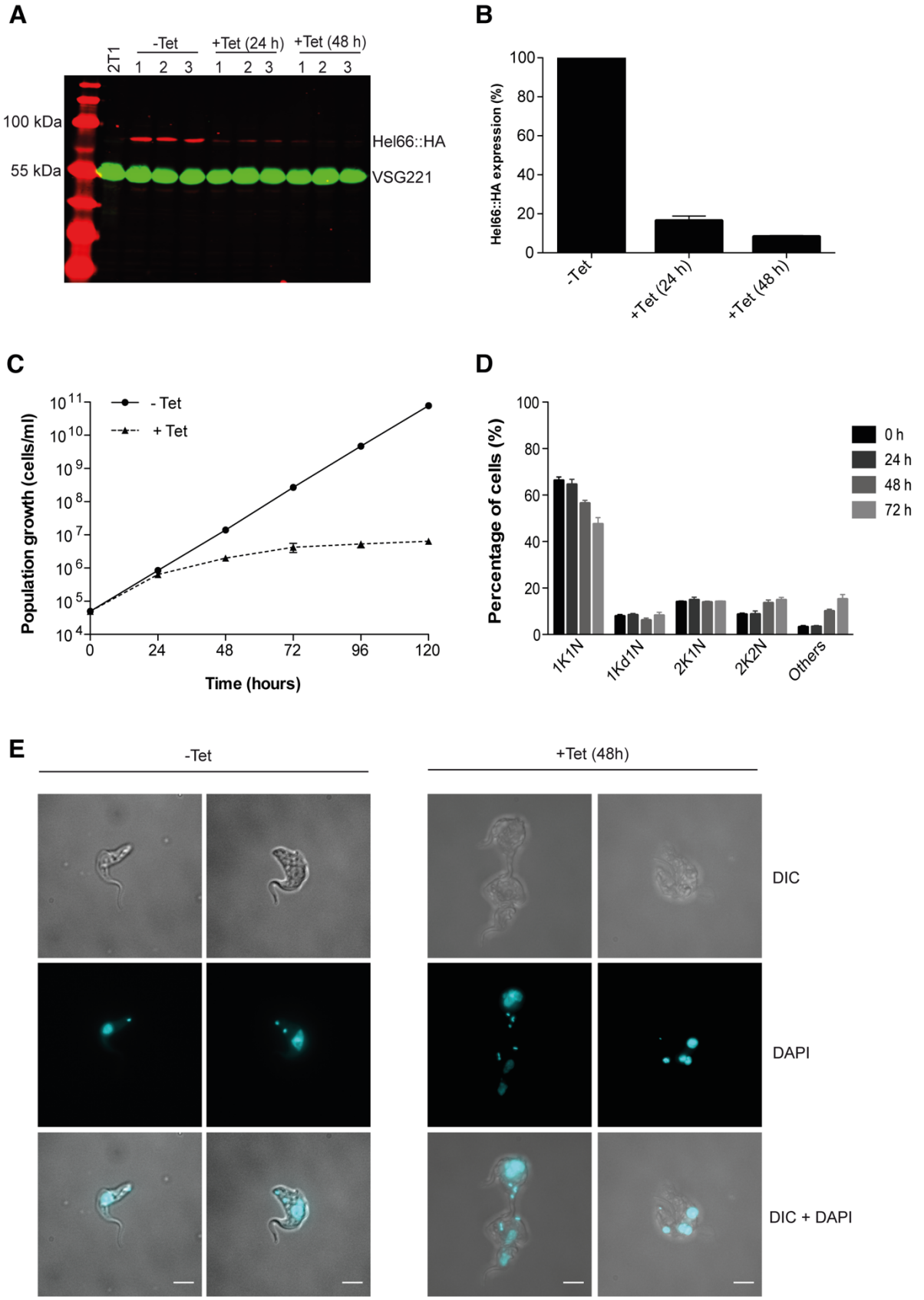

Figure 26. Hel66 is an essential protein in BSF *T. brucei*. (A) Western blot showing the extent of Hel66::HA depletion following RNAi induction in 3 independent clones (1, 2, 3). 2T1 is the parental cell line. Detection of VSG221 served as a loading control (B) Amounts of Hel66::HA protein quantified from the Western blot in (A). The signal intensity of Hel66::HA was normalised to the VSG221 signal and expressed relative to the -Tet controls. The data

shown are means ± standard error of the mean (SEM) of three independent clones. (C) Cumulative growth curve of uninduced (-Tet) and induced (+Tet) Hel66-RNAi cells. Data are means ± standard error of the mean (SEM) of three independent clones. (D) Cell cycle analysis over the course of Hel66 depletion. Cells were stained with DAPI and classified to the different cell cycle stages according to the number of nuclei (N) and kinetoplasts (K). At least 200 cells were analysed for each time point and data presented as mean percentages ± SEM of three independent clones. (E) Images showing examples of cells with aberrant nuclei and kinetoplast configurations and morphology following RNAi-mediated depletion of Hel66 (right); uninduced cells are shown as controls (left). Scale bar = 5µm. The figure is adapted from Bakari-Soale et al. (2021).

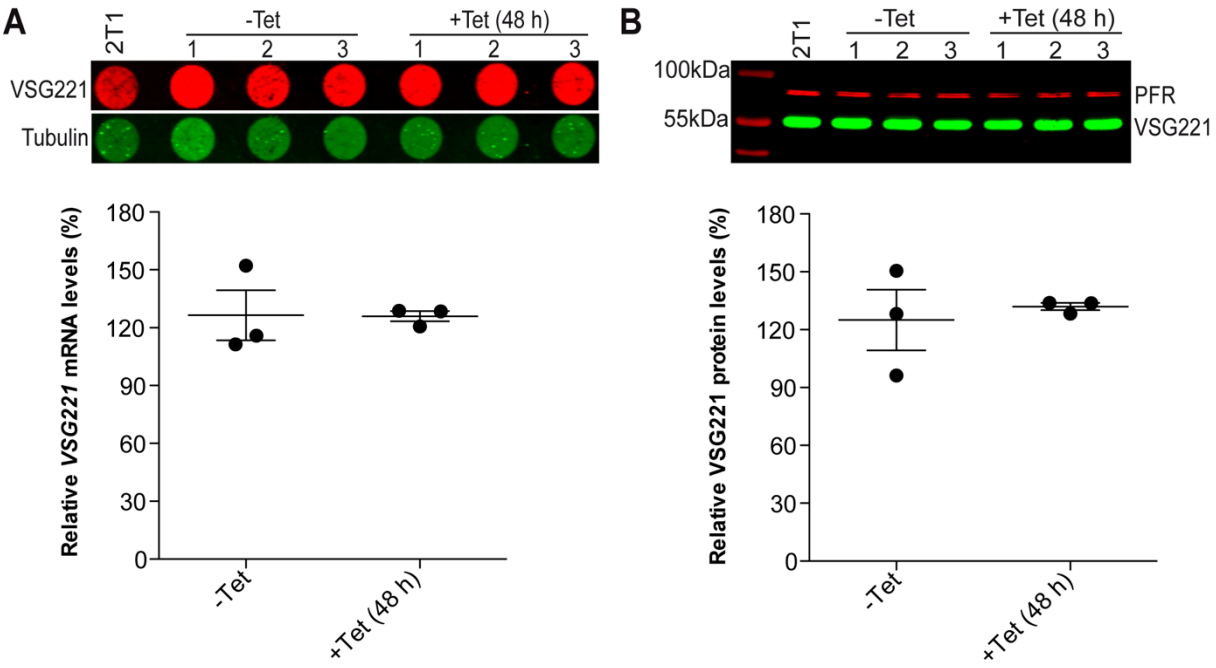

Figure 27. *VSG221* mRNA and VSG221 protein levels following depletion of Hel66. (A) *VSG221* mRNA levels in uninduced (-Tet) and induced (+Tet 48 h) Hel66-RNAi cells. *Tubulin* mRNA was used as an internal control (B) VSG221 protein levels in uninduced (-Tet) and induced (+Tet 48 h) Hel66-RNAi cells. The paraflagellar rod proteins 1 and 2 (PFR) were used as an internal control. 2T1 represents the parental cell line and 1, 2 and 3 are different clones of Hel66-RNAi cells. *VSG221* mRNA and VSG221 protein levels are expressed relative to VSG221 amounts in the parental 2T1 cell line. Data are presented as dot plots with the standard error of the mean (SEM). The long horizontal lines represent the average (mean) amounts.

3.4.6 Loss of Hel66 causes accumulation of rRNA processing intermediates

Several members of the DExD/H family of proteins have been shown to be involved in the process of ribosome biogenesis (Boisvert et al., 2012; Cristodero & Clayton, 2007; Wild et al., 2010; Zhang et al., 2011). The localisation of Hel66 to the nucleolus, the primary site of ribosome biogenesis, suggested a similar function of this protein. This led me to probe whether the protein is essential for the biosynthesis of ribosomes. The effect of Hel66 depletion on rRNA processing, an early event in ribosome biogenesis was therefore investigated. Like in other eukaryotes, ribosome biogenesis in trypanosomes is complex and the process starts in the

nucleolus with transcription of a 9.2 kb precursor transcript. The precursor transcript is then cleaved into 3.4 kb and 5.8 kb intermediates. The 3.4 kb intermediate, which is the precursor for the 18S rRNA, is further processed into the 40S subunit. The 5.8 kb intermediate (precursor for the LSU and smaller rRNAs) is cleaved into 0.61 kb (5.8S precursor) and 5.0 kb (25/28S precursor) intermediates which are further processed into components of the 60S subunit (Figure 28A).

The rRNA processing intermediates were detected by Northern blotting using probes (pre-18S, ITS2 and ITS3) that hybridise to specific regions of the *T. brucei* pre-rRNA as previously described (Jensen et al., 2003; Sakyiama et al., 2013; Umaer et al., 2014). The pre-18S probe can detect the 9.2kb precursor transcript and the 3.4 kb and 2.6 kb intermediates of the 18S rRNA. However, in this study only the 3.4 kb intermediate was detected with this probe, as the 9.2kb precursor and 2.6 kb intermediate are less stable and thus not abundant enough to be detected (Figure 28B). Within 24 h of Hel66 RNAi induction there was a 1.4-fold increase in intensity of the 3.5 kb intermediate with a 1.5-fold increase observed at 48 h (Figure 28B).

ITS2 and ITS3 probes can both detect the 9.2 kb precursor and the 5.8 kb intermediate of the LSU precursor. Additionally, the ITS2 probe detects the 0.61 kb intermediate and the ITS3 probe detects the 5.0 kb intermediate. The unstable 9.2 kb precursor was undetectable but the 5.8 kb and 0.61 kb intermediates were abundant enough for detection with the ITS2 probe (Figure 28C). Upon RNAi depletion of Hel66, hybridisation of the ITS2 probe revealed a 1.4-fold accumulation of the 5.8 kb intermediate and a 1.5-fold decrease in the intensity of the 0.61 kb intermediate after 24 h. A recovery of the intensity of the 0.61 kb intermediate was observed after 48 h of induction with the appearance of a slightly larger RNA intermediate (marked with an asterisk). This indicated that loss of Hel66 caused the accumulation of the 5.8 kb intermediate, a decrease in the 0.61 kb intermediate and the accumulation of a previously unknown precursor of the 0.61 kb intermediate (Figure 28C).

The 5.8 kb and 5.0 kb intermediates were also abundant enough for detection with the fluorescent ITS3 probe (Figure 28D). Hybridisation of the ITS3 probe revealed an accumulation of the 5.8 kb intermediate and a decrease in the intensity of the downstream 5.0 kb intermediate (Figure 28D), indicating an inhibition of the processing step from the 5.8 kb intermediate to the 5.0 kb intermediate.

The data indicate the involvement of Hel66 in the processing of ribosomal RNAs of both the small and large ribosomal subunits. Hel66 may be playing a general function in rRNA processing as loss of the protein resulted in the inhibition of more than one processing step.

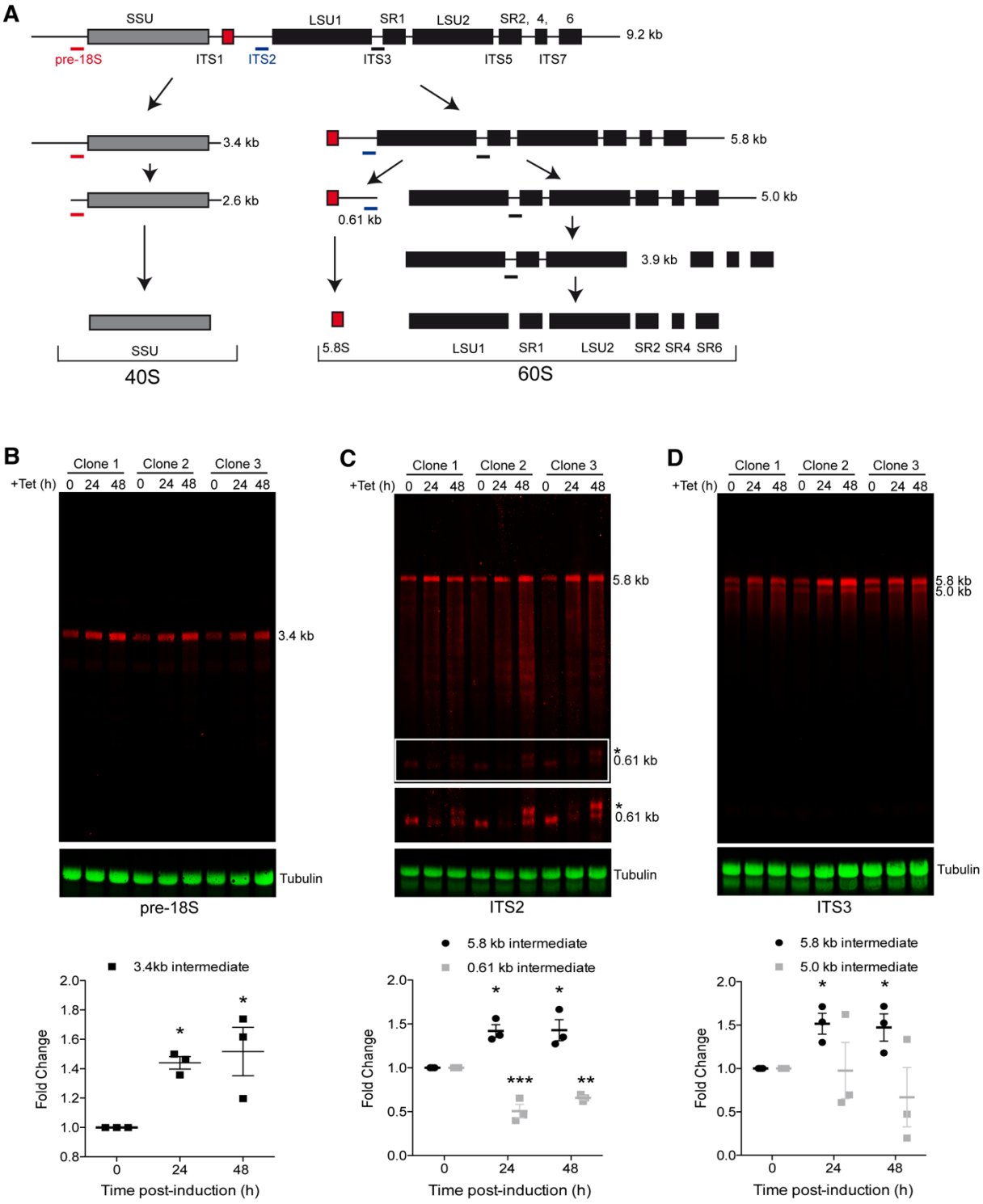

Figure 28. RNAi-mediated depletion of Hel66 results in accumulation of rRNA processing intermediates. (A) Illustration of the rRNA processing pathway in trypanosomes. The annealing sites of probes used in Northern blots are shown as thick lines in red (Pre-18S), blue (ITS2) and black (ITS3). The illustration is based on stable processing intermediates observed in previous studies (Rink et al., 2019; Umaer et al., 2014; White et al., 1986). (B-D) Northern blot analysis showing the accumulation of rRNA processing intermediates after induction of Hel66 RNAi for 0, 24 and 48 hours. Total RNA from the cells was probed with pre-18S (B), ITS2 (C) and ITS3 (D) and for *tubulin* (loading control). The experiment was performed with three independent clones. The averages of the three replicates normalised to *tubulin* with error bars representing standard error of the mean are shown below. The contrast of the blot underneath in (C) was increased to make the 0.61 kb band and its precursor (area in white box)

more visible. The previously unknown processing intermediate of the 5.8 S rRNA is marked with an asterisk. Statistical significance was determined by one-way ANOVA and Dunnette's Post-hoc test. Differences to the uninduced cells were considered statistically significant when the P-value was less than 0.05 (* represents $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$). The figure is adapted from Bakari-Soale et al., (2021).

3.4.7 Depletion of Hel66 results in decreased global translation and accumulation of mRNA

The negative impact of Hel66 depletion on rRNA processing could affect the formation of mature rRNAs and consequently protein synthesis. In order to investigate the effect of Hel66 depletion on translation, the SUnSET assay was used. This assay involves exposing cells to very small quantities of puromycin over a short period. Puromycin which is a structural analogue of tyrosyl transfer RNA is incorporated into growing polypeptide chains, thereby terminating translation (Goodman & Hornberger, 2013). The incorporation of puromycin in the nascent polypeptides therefore directly reflects the rate of global translation *in vivo*. Uninduced (-Tet) and induced (+Tet 24 h and +Tet 48 h) Hel66-RNAi cells were incubated with puromycin and the puromycin incorporation detected by Western blot using an anti-puromycin antibody. Global translation was reduced to 60 % of levels in uninduced cells (-Tet) after 24 h of RNAi induction (Figure 29A, B). A further decrease of translation to 40 % compared to the uninduced cells was observed after 48 h of RNAi induction. There was however no detectable signal without puromycin treatment but a drastic decrease in translation to < 5 % when cells were treated with the translational inhibitor cycloheximide (CHX) prior to puromycin treatment (Figure 29A, B). The data therefore indicates that depletion of Hel66 caused a decrease in global translation.

The decreased global translation could result in accumulation of mRNA if the mRNA synthesis remains unaffected. To investigate this, total RNA was isolated over a time-course of Hel66 RNAi induction and analysed by Northern blot. An oligo antisense to the spliced leader RNA which is trans-spliced onto all trypanosome mRNAs was used to probe for total mRNA. Depletion of Hel66 resulted in a significant increase in the levels of mRNA after 48 h of RNAi induction. Levels of the total mRNA increased slightly within the first 24 h and then to about 1.3-fold after 48 h of Hel66 RNAi induction (Figure 29C, D), suggesting an accumulation of total mRNA following Hel66 depletion.

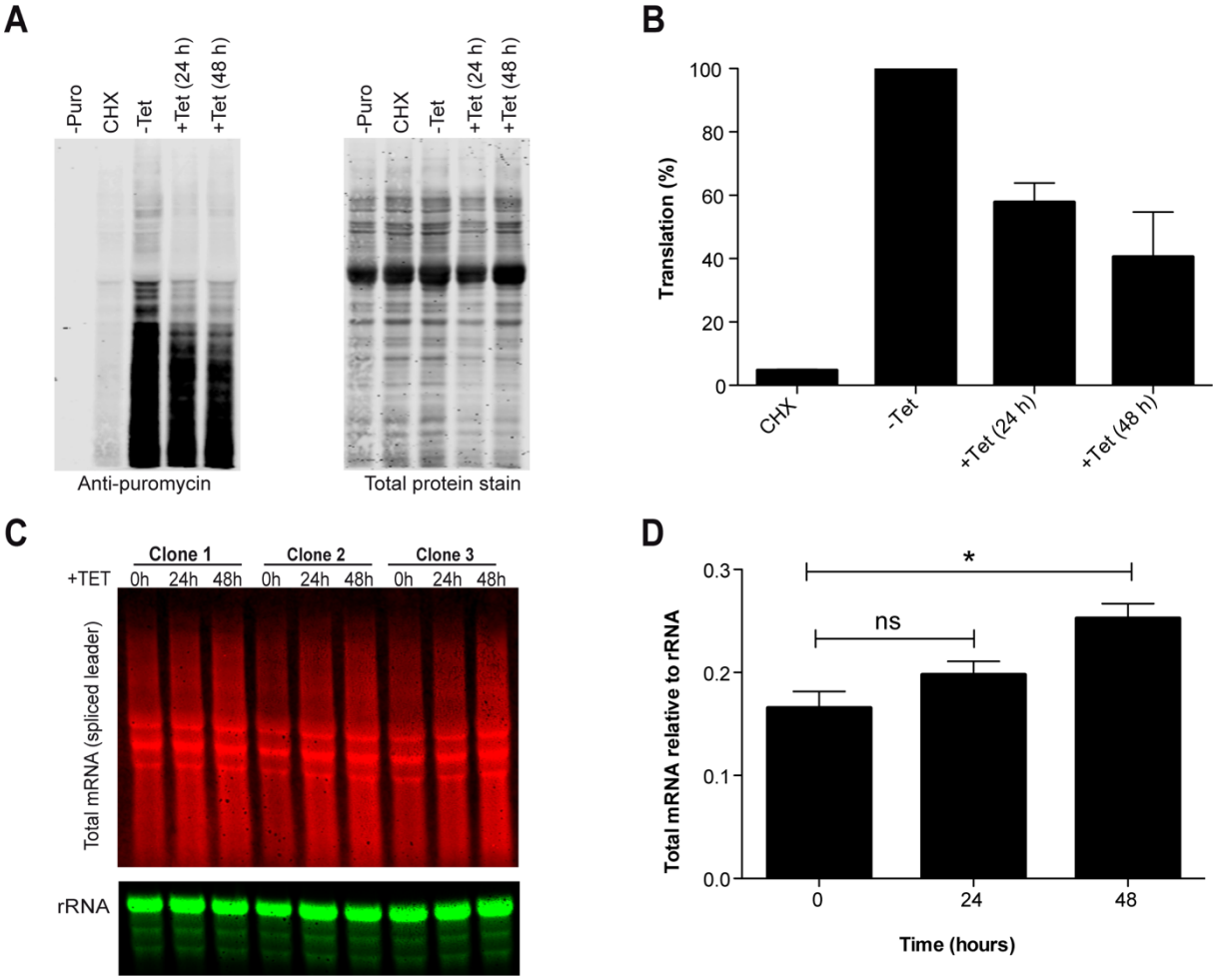

Figure 29. Loss of Hel66 inhibits translation and causes accumulation of total mRNA. (A) Left: Representative Western blot showing puromycin detection after the SUnSET assay. Lane -Puro: uninduced cells with no puromycin treatment; Lane CHX: uninduced cells treated with cycloheximide (inhibitor of translation) 30 min prior to puromycin treatment; Lane -Tet: uninduced cells treated with puromycin; Lane +Tet (24 h): cells treated with puromycin after 24 h of RNAi induction; Lane +Tet (48 h): cells treated with puromycin after 48 h of RNAi induction. Right: Loading control (total protein stain). (B) Quantification of puromycin incorporation in nascent polypeptides (translation) from the Western blots. Data are from three independent clones and are presented as mean ± standard error of the mean (SEM). (C) Northern blot probed for total mRNA using an oligo antisense to the spliced leader mRNA. A probe antisense to mature rRNA served as loading control. (D) Quantification of total mRNA amounts from the Northern blot. The averages of the three independent clones normalised to rRNA with error bars representing standard error of the mean are shown below. Differences in total mRNA amounts in induced cells in comparison to uninduced cells (0h) was considered statistically significant (*) when $P < 0.05$ (t-test). Figure 29 panels A and B are adapted from Bakari-Soale et al. (2021).

3.4.8 Absence of phenotypic effect upon overexpression of ATPase dead Hel66

The characteristic motifs of the DExD/H family of proteins are highly conserved in Hel66. The amino acid residues involved in either RNA binding or catalytic activity are conserved as well. The DExD/H box proteins have an RNA-dependent ATPase activity and the walker B motif/motif II (DExD/H) has been shown to be important for this catalytic activity. Mutation of the glutamic acid (E) to glutamine (Q) in motif II has been shown to terminate ATP hydrolysis activity in many DExD/H box proteins (Carroll et al., 2011; Kramer et al., 2010). Therefore, to investigate whether the putative ATPase activity of Hel66 is essential for its role in rRNA processing, the glutamic acid (E) in the motif II of Hel66 (DEVD) was mutated to glutamine (Q), generating a dead ATPase mutant (E254Q). HA-tagged copies of the wild-type Hel66 (Hel66$_{WT}$::HA) and the dead ATPase mutant (Hel66$_{E254Q}$::HA) were overexpressed in 13-90 BSF *T. brucei* cells. There was no observable growth phenotype on the cells upon overexpression of either Hel66$_{WT}$::HA or Hel66$_{E254Q}$::HA (Figure 30A,B). The tagged proteins were overexpressed in both cases (Figure 30C, D) but the degree of overexpression could not be estimated since the endogenous Hel66 levels could not be measured. With the absence of an observable phenotypic effect, the role of ATPase activity in Hel66 was not further explored in this study.

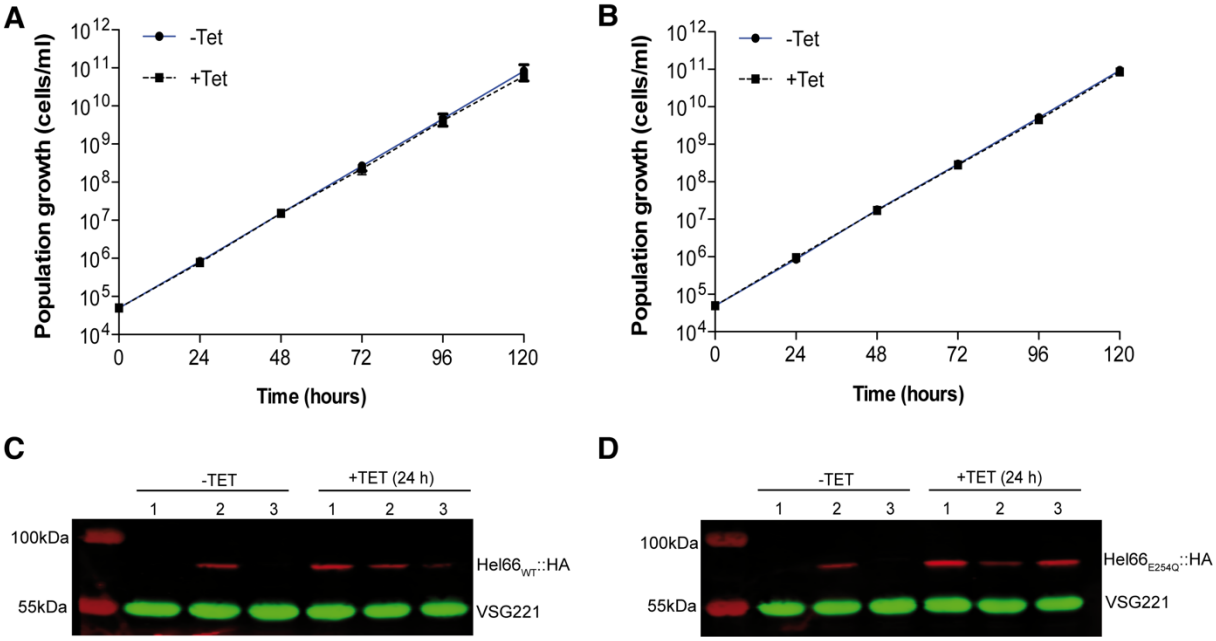

Figure 30. No growth phenotype upon overexpression of wild-type and dead ATPase mutant Hel66. (A) Cumulative growth curve of uninduced (-Tet) and induced (+Tet) cells for overexpression of Hel66$_{WT}$::HA. (B) Cumulative growth curve of uninduced (-Tet) and induced (+Tet) cells for overexpression of Hel66$_{E254Q}$::HA. Data are means ± standard error of the mean (SEM) of three independent clones. (C,D) Western blots showing the overexpression of Hel66$_{WT}$::HA (C) and Hel66$_{E254Q}$::HA (D). VSG221 served as loading control.

4.0 DISCUSSION

The VSG is an abundantly expressed protein that is essential in BSF African trypanosomes. The control of the expression of this protein is a complex multifactorial process which is highly regulated. Generally, gene expression in African trypanosomes is largely post-transcriptional and the 3´ UTRs of most genes are critical for the regulation of transcript levels. In *T. brucei*, a conserved 16mer motif within the VSG 3´ UTR has been shown to be essential for the stability of *VSG* mRNAs. This motif is 100 % conserved in all of the > 1000 VSG genes present in this parasite but absent in VSGs in *T. congolense* and *T. vivax*. To date, no stabilisation or destabilisation motif has been described to have such a high conservation in a group of genes. For instance, the stability-associated 16 nucleotide sequence element in the 3´ UTR of procyclin transcripts is highly conserved but has nucleotide substitutions occurring within the motif (Furger et al., 1997). Also, the motifs M24 and M23 in yeast and the miR-381 and miR-219 sequence motifs in humans all show high conservation but are not 100% conserved (Shalgi et al., 2005). This prompted a further probe into what other roles this motif could be involved in besides stabilisation of the *VSG* transcript. Also, as part of an attempt to identify proteins interacting with the 16mer motif, a novel nucleolar protein involved in ribosome biogenesis was identified and characterised.

4.1 100 % conservation of the 16mer motif is not essential for cell viability and expression of functional levels of VSG

The VSG 3´ UTR has been previously shown to be important for stabilisation of mRNA of a CAT reporter gene (Berberof et al., 1995). Through mutational analysis it was identified that sequence elements within the last 97 bp before the polyadenylation site of the transcript were responsible for the stabilisation effect. This region was found to contain conserved motifs including the 8mer, 16mer and the polypyrimidine tract. A subsequent study demonstrated that indeed the VSG 3´ UTR was essential for maintenance of *VSG* mRNA stability and hence responsible for the production of functional levels of the VSG (Ridewood et al., 2017). The authors also showed that the 100 % conserved 16mer motif is the essential sequence element responsible for maintenance of *VSG* mRNA stability. Scrambling of the 16mer motif resulted in a drastic reduction in the mRNA half-life from over 100 min to about 20 min (Ridewood et al., 2017). Christopher Batram earlier in his Diplom and PhD theses (Batram, Diplom thesis 2009; PhD thesis 2013) also demonstrated that deletion of either part of or the entire 16mer motif caused a drastic decrease in *VSG* mRNA levels, thus impacting high VSG protein production. A tolerated mutation of the 16mer motif involving substitution of the first three

nucleotides of the motif (TGA to ACT) was however shown to produce VSG protein levels sufficient for cell viability. All of these studies used ectopic expression of a second VSG for the functional analysis of the 16mer motif since the endogenous VSG is essential for parasite viability. The endogenous VSG with its intact VSG 3´ UTR was therefore always present in the expression site. The presence of a second VSG with its intact 3´ UTR in the expression site could however influence the outcome of the functional analysis, hence the need to use single-expresser cells to confirm the outcome.

The main hypothesis of this study was that the 16mer is involved in other roles besides *VSG* mRNA stability, hence its 100 % conservation. If this were true, it would be expected that not every mutation disrupting this 100 % conservation should impact the production of functional VSG levels. To investigate this, it was necessary to generate transgenic *T. brucei* cells expressing a single VSG (single-expresser cell) harbouring a mutated 16mer motif. Before generation of the single-expresser cells, double-expresser cell lines were first generated by inserting an ectopic VSG121 downstream of the endogenous VSG221 in 13-90 cells. This was to ensure that the ectopic VSG121 was situated in a telomere proximal position. It was observed that placing the ectopic VSG downstream of the endogenous VSG resulted in reduction of the endogenous *VSG* mRNA and protein at approximately inverse amount to the level of expression of the ectopic VSG. This was in agreement with previous studies where expression of an ectopic VSG regardless of the genomic location resulted in modulation of the endogenous VSG levels (Aroko et al., 2021; Batram et al., 2014; Ridewood et al., 2017; Smith et al., 2009; Zimmermann et al., 2017). In previously-generated constitutive double-expresser cells where the ectopic VSG was placed upstream of the endogenous VSG, modulation of the endogenous VSG resulted in approximately equal amounts of both VSGs (Muñoz-Jordán et al., 1996; Ridewood et al., 2017; Smith et al., 2009). In this study however, it was observed that the ectopic *VSG121* mRNA and the endogenous *VSG221* mRNA where expressed in a ratio of 30:80 when the ectopic VSG was placed downstream of the endogenous VSG. This observation could be due to low transcriptional activity downstream of the endogenous VSG or a position-dependent repression of the ectopic VSG due to its proximity to the telomere (Horn & Cross, 1997). Therefore, although *VSG* mRNA regulation is independent of the genomic location of an ectopic VSG, the position of the ectopic VSG within the expression site impacts the degree of modulation of the endogenous VSG. With only 30 % of the ectopic *VSG121* mRNA expressed, introducing mutations into the 16mer motif resulted in expression of less than 10 % of *VSG121* mRNA and less than 25 % of *VSG121* protein in the double-expresser cell lines. The low ectopic VSG expression levels led to loss of the endogenous *VSG* mRNA regulation in all of the mutant cell

lines except $221^{ES}121_{inver8mer}$. The $221^{ES}121_{inver8mer}$ cell line expressed the most *VSG121* mRNA and VSG121 protein, suggesting that the ectopic VSG amounts in this mutant cell line could still be sufficient for modulation of the endogenous VSG. Using reporter cell lines in which a GFP ORF flanked by a VSG121 3′ UTR with either an intact 16mer motif ($221^{ES}GFP_{WT}$) or a mutated 16mer motif ($221^{ES}GFP_{\Delta 46\text{-}52}$), the endogenous *VSG* mRNA was also not regulated in either of the cell lines although high GFP amounts were expressed. About 2.7-fold less *GFP* mRNA was however expressed in $221^{ES}GFP_{\Delta 46\text{-}52}$ cells compared to $221^{ES}GFP_{WT}$ cells, an outcome supporting the finding that the 16mer motif is essential for mRNA stability. Erick Aroko recently demonstrated in his PhD thesis that adding an endoplasmic reticulum (ER) import signal to the GFP ORF in $221^{ES}GFP_{WT}$ results in the regulation of the endogenous *VSG221* mRNA (Aroko et al, 2021; Aroko, PhD thesis 2021). These data together with my finding that no *VSG* mRNA regulation occurs with low ectopic VSG expression in mutant double-expressers and also when the GFP in the reporter cell was not targeted to the ER suggest that high levels of expression and targeting of the ectopic mRNA to the ER are necessary for the regulation/modulation of the endogenous VSG.

Based on Christopher Batram's findings that substitution of the first three nucleotides of the 16mer motif (TGA to ACT referred to as N46-48 in this thesis) does not significantly impact VSG expression (Batram, Diplom thesis 2009; PhD thesis 2013), transgenic *T. brucei* cell lines expressing a single VSG (single-expresser cell) harbouring either an intact 16mer motif or the tolerated 16mer mutation were generated by deleting the endogenous VSG221 from the double-expresser cell lines $221^{ES}121_{WT}$ and $221^{ES}121_{N46\text{-}48}$, respectively. The 16mer mutant single-expresser cell line $\Delta221^{ES}121_{N46\text{-}48}$ expressed 43 % less *VSG121* mRNA and grew approximately 2 h slower per cycle compared to the control cell line with the intact 16mer motif ($\Delta221^{ES}121_{WT}$). VSG121 protein levels were however similar and comparable to wild-type amounts in both cell lines, suggesting that the intact 16mer motif is necessary for mRNA abundance but not required for expression of functional levels of VSG protein. With regards to the differences in population doubling times between the two cell lines, the slower growth in $\Delta221^{ES}121_{N46\text{-}48}$ cells may be due to lower amount of *VSG121* mRNA expressed in this cell line. As the rate of synthesis of a protein is dependent on the concentration and translation efficiency of its mRNA (Polymenis & Aramayo, 2015), $\Delta221^{ES}121_{N46\text{-}48}$ cells may take a longer time to translate functional levels of the VSG protein due to the lower mRNA levels.

The lower levels of *VSG* mRNA in $\Delta221^{ES}121_{N46\text{-}48}$ single-expresser cells compared to $\Delta221^{ES}121_{WT}$ cells is possibly a result of differences in stability of *VSG121* transcripts produced by the two cell lines. 16mer-dependent N^6-methyladenosine (m^6A) modification of *VSG*

poly(A) tails has been proposed as a novel mechanism contributing to *VSG* transcript stability (Viegas et al., 2020). It was therefore investigated whether the difference in *VSG121* mRNA levels in the two single-expresser cell lines is due to differences in m^6A modification as a result of the 16mer mutation in $\Delta221^{ES}121_{N46\text{-}48}$ cells. Analysis of the m^6A modification in *VSG121* mRNA from the two cell lines revealed that both cell lines had similar m^6A levels, suggesting that 100 % conservation of the 16mer motif is not required for m^6A modification of *VSG* transcripts.

4.2 The intact 16mer motif is not required in the active VSG for silencing

The single active VSG is expressed from the telomeric region of the active ES. Changing of the actively expressed VSG occurs either via a gene conversion event where a new VSG is moved into the active ES or via telomere exchange, where VSGs located at telomeric regions are swapped between a silent ES and the active ES (Li, 2015; Taylor & Rudenko, 2006). VSG switching can also occur by transcriptional switch (*in situ* switch), where a new ES is activated and the previously active ES is transcriptionally silenced resulting in silencing of the active VSG (Taylor & Rudenko, 2006). Apart from the transcriptional repression, post-transcriptional processes such as mRNA stability may also be involved in silencing of the active VSG during an *in situ* switch. As the 16mer motif is involved in stability of the *VSG* transcript, the role of its 100 % conservation in VSG silencing during an *in situ* switch was investigated. It was previously shown that high-level expression of an ectopic VSG initiates attenuation of the active ES and silencing of the VSG (Batram et al., 2014). Using this approach, VSG221 was inducibly expressed at high levels in the cell lines $\Delta221^{ES}121_{WT}221^{Tet}$ and $\Delta221^{ES}121_{N46\text{-}48}221^{Tet}$ which have an expression site resident VSG121 with an intact 16mer motif or a mutated 16mer motif, respectively. The expression site resident *VSG121* mRNA and protein levels decreased with similar kinetics in both cell lines, suggesting that the 100 % conservation of the 16mer motif is not essential in the active VSG 3´UTR for efficient silencing to occur.

Uptake of bloodstream trypanosomes into the tsetse fly midgut results in differentiation of bloodstream form trypanosomes to procyclic forms. A key feature of this differentiation process is the replacement of the VSG coat with a new surface coat, procyclin (Roditi et al., 1989). Repression of the VSG expression occurs almost immediately with procyclin expression simultaneously induced (Gruszynski et al., 2006). This results in destabilisation of the *VSG* mRNA as the half-life is reduced from ~ 4.5 h in bloodstream trypanosomes to 1.2 h in the transforming cells (Ehlers et al., 1987). Since the 16mer motif is essential for maintenance of *VSG* mRNA stability, the role of its 100 % conservation in VSG silencing during differentiation

from BSF to PCF was investigated using an *in vitro* differentiation assay that mimics the process in the tsetse fly midgut. Upon induction of differentiation in the single-expresser cell lines $\Delta221^{ES}121_{WT}$ and $\Delta221^{ES}121_{N46\text{-}48}$, it was found that the expression site resident VSG was silenced with similar kinetics in both cell lines. This suggested that regardless of the presence of an intact 16mer motif in the active VSG 3´ UTR, silencing of the active VSG occurs rapidly and efficiently. 100 % conservation of the 16mer motif is therefore not essential for VSG silencing during differentiation. As sequence elements present in the 3´ UTR are responsible for transcriptional crosstalk between the VSG and procyclin (Berberof et al., 1995; Pays, 2005; Vanhamme et al., 1995), it was directly investigated whether this crosstalk was affected by the mutation in the 16mer motif. It was however found that the crosstalk between the VSG and procyclin was not dependent on 100 % conservation of the 16mer. This could explain the reason why we found no differences in the kinetics of silencing when the motif was mutated. The 8mer motif has recently been implicated in a fail-safe mechanisms that ensures destabilisation of accidently-produced VSG transcripts in procyclics (do Nascimento et al., 2021). The presence of the 8mer motif in both single-expresser cell lines could also account for the similar kinetics of VSG silencing during the differentiation process.

Overall, it was established that 100 % conservation of the 16mer motif within the active VSG 3´ UTR is neither essential for silencing of the active VSG during parasite differentiation nor during a transcriptional VSG switch.

4.3 Efficient silencing and exchange of the active VSG during an *in situ* switch requires 100 % conservation of the 16mer motif in a silent VSG

A transcriptional/*in situ* VSG switch involves activation of a new ES and silencing of the previously active ES. Therefore having established that the 100 % conservation of the 16mer motif within the active VSG is not essential for silencing, it was investigated whether the intact 16mer is required in the new VSG to trigger efficient silencing and exchange of the active VSG during switching. For this, the overexpression system mimicking transcriptional VSG switching, in which the VSG is expressed from the ribosomal spacer region driven by a tetracycline inducible T7 promoter was used (Batram et al., 2014). Previous studies have established that high level expression of an ectopic wild-type VSG triggers efficient silencing and exchange of the active VSG (Batram et al., 2014; Zimmermann et al., 2017; Henning, Bachelor thesis 2012; Specht, Master thesis 2013). The induction of overexpression of the ectopic VSG produces a phenotypic plasticity in pleomorphic trypanosomes (Zimmermann et al., 2017). In monomorphic trypanosomes however, the phenotypic plasticity occurs depending

on the active BES and the specific VSG isoform being overexpressed (Batram et al., 2014; Henning, Bachelor thesis 2012; Specht, Master thesis 2013; Goos, Master thesis 2013). The general observation from all of the wild-type VSG overexpression was that the active VSG was always silenced and after 24 h the ectopic VSG became the abundantly-expressed VSG at both the mRNA and protein levels.

To test whether the 100 % conservation of the 16mer motif is essential for this process, $VSG121_{WT}$ (with an intact 16mer motif) and $VSG121_{N46-48}$ (with a mutated 16mer motif) were overexpressed in 13-90 BSF cells. With $VSG121_{WT}$ overexpression, the endogenous VSG 221 was silenced with similar kinetics as earlier described (Batram et al., 2014). The *VSG221* mRNA decreased to ~ 32 % by 24 h with the VSG221 protein levels decreasing to ~ 40 %. The ectopic *VSG121* mRNA and VSG121 protein were however at 88 % and 74 %, respectively, indicating efficient silencing and exchange of the endogenous VSG. The cells also stalled growth after 24 h for 4 days similar to observations made by Batram et al., (2014). However, overexpression of $VSG121_{N46-48}$ failed to efficiently silence and replace the endogenous/active VSG221. The endogenous *VSG221* mRNA only decreased to 55 – 60 % within the first 24 h and the VSG221 protein remained the abundantly expressed VSG at 59 %. The ectopic VSG121 protein on the other hand was not well expressed, as only ~ 35 % of wild-type levels were detected after 24 h although the *VSG121* mRNA was at ~ 92 %. Mutation of the 16mer motif therefore prevented efficient silencing and exchange of the endogenous VSG221, suggesting that the 100 % conservation of the 16mer motif is necessary to trigger efficient VSG silencing and coat exchange during a transcriptional VSG switch.

Similarly, $VSG118_{WT}$ (with an intact 16mer motif) and $VSG118_{N46-48}$ (with a mutant 16mer motif) were overexpressed in 13-90 BSF cells. Overexpression of $VSG118_{WT}$ resulted in cells with two distinct phenotypes, slow-growers and fast-growers (Henning, Bachelor thesis 2012; Figure 17B). It was observed that regardless of the growth phenotype, the endogenous VSG221 was efficiently silenced with the *VSG221* mRNA and VSG221 protein levels decreasing to 36 % and 32 %, respectively by 24 h. Overexpression of $VSG118_{N46-48}$ however produced contrasting results between the two clonal populations. In the fast-growers, the endogenous VSG221 was efficiently silenced at both the mRNA and protein levels. In the slow growers however, the endogenous VSG221 was not efficiently silenced. The *VSG221* mRNA levels decreased to about 42 % by 24 h and the protein levels remained above 50 %. These data therefore confirm that the 100 % conservation of the 16mer motif is important for VSG silencing and coat exchange. The high conservation of the motif is essential in a silent VSG for transcriptional switching from the active VSG to this silent VSG. This finding further supports

the assumption that the mechanism of antigenic switching operates differently in *T. brucei* compared to *T. congolense* and *T. vivax*. These trypanosome species lack the conserved 16mer motif and may have evolved different sequence elements for regulation of stability and transcriptional VSG switching.

The observed differences in the VSG silencing dynamics between the fast- and slow-growers in VSG118 overexpression could however be a result of the ES activity. The ES activity has been shown to influence phenotypic plasticity with slow-growers having attenuated ES activity (Zimmermann et al., 2017). Since the effect of the 16mer mutation was only observed in clonal cells exhibiting reduced growth upon overexpression of the ectopic VSG, the activity of ES may be influencing exchange of the VSG coat. The mechanism or process by which this occurs could however not be determined in this study.

4.4 Identification of 16mer binding proteins

Trans-acting proteins binding to DNA or RNA are essential for sustenance of cellular activity. Identifying the RNA or DNA binding proteins interacting with specific cis-regulatory elements could therefore provide information on the function of the sequence elements. It was hypothesised that the high conservation of the 16mer motif is as a result of the motif acting as a DNA binding domain facilitating gene expression. For this reason a pull-down assay was carried out with the 16mer DNA. Further analysis of the potential interaction partners however showed that none of the essential proteins significantly impacted VSG expression. The functions of these proteins were therefore not further explored.

Since it had already been established that the 16mer motif is essential for mRNA stability, pull-down assays were also carried out to identify the RNA binding protein(s) interacting with the motif. Hel66 (Tb927.10.1720), a putative RNA-dependent DEAD/H RNA helicase was the only protein enriched when the intact VSG 3´ UTR was used as bait, but not when using three control RNAs (VSG 3´ UTR with a scrambled 16mer and 8mer, the reverse complement of the VSG 3´ UTR with a scrambled 16mer and 8mer and VSG 3´ UTR with a reverse complement of the 16mer). Hel66 was therefore further characterised to determine its function in the parasite.

4.5 Hel66 is a novel nucleolar DExD/H box protein involved in ribosome biogenesis in *T. brucei*

Biogenesis of ribosomes in eukaryotes is a complex process involving small nucleolar RNAs and multiple proteins including, ribosomal proteins, RNA helicases, GTPases and ATPases (Henras et al., 2008; Henras, 2015; Woolford & Baserga, 2013). Of these proteins, most RNA

helicases of the DExD/H protein family have been implicated in eukaryotic ribosome biogenesis. DExD/H proteins play important roles in RNA metabolism and are thought to promote structural rearrangement at various stages during rRNA processing and ribosome assembly. In *Trypanosoma brucei*, there are about 51 DExD/H helicases, many of which remain uncharacterised (Aslett et al., 2010; Berriman et al., 2005; Gargantini et al., 2012). Although Hel66 was identified from a screen to find interaction partners of the 16mer motif within the VSG 3´ UTR, it was found that Hel66 does not directly interact with the 16mer RNA and is not involved in VSG expression. Hel66 was also absent from a recent *VSG*-mRNA proteome (do Nascimento et al., 2021), suggesting that Hel66 interaction observed in the pull-down was unspecific.

Some of the DExD/H box proteins have been shown to localise either in nucleolar (19) or nucleoplasmic (9) regions, indicating that members of this protein family may be involved in ribosome processing (Berriman et al., 2005). Hel66 was found to localise to the nucleolus. As many as 17 DExD/H RNA helicases, mostly those with nucleolar localisation, have been assigned to yeast homologues with known functions in ribosome biogenesis (Michaeli, 2012; Rajan et al., 2019; Umaer et al., 2014). These 17 proteins were also co-purified in *Leishmania tarentolae* with at least one of four proteins (UTP18, SSF1, RPL24 and ARX1) involved in ribosome processing (Barandun et al., 2017; Rajan et al., 2019; Sanghai et al., 2018; Wu et al., 2016). Hel66 is absent from the 17 DExD/H proteins assigned to a yeast homologue (Michaeli, 2012; Rajan et al., 2019; Umaer et al., 2014). A reverse BLAST of the closest hits from yeast do not return Hel66. Hel66 may therefore be a trypanosome-unique protein involved in ribosomal processing steps that may be absent in yeast. Copurification of Hel66 with SSF1 (Shulamit Michaeli, personal information) however suggests its function in early processing of the large ribosomal subunit. Mtr4, which localises to the nucleolus, is the only other DExD/H RNA helicase that has been identified in the context of rRNA maturation in *T. brucei* (Cristodero & Clayton, 2007; Dean et al., 2017). Depletion of Mtr4 causes accumulation of ribosomal precursor RNAs (Cristodero & Clayton, 2007). A recent study in yeast provided evidence of a direct function of Mtr4 in rRNA maturation, as the protein was shown to mediate a direct interaction between the exosome and the 90S pre-ribosomal subunit (Lau et al., 2021).

Knockdown of Hel66 via RNAi resulted in growth arrest with a perturbation in the cell cycle. This effect may be a response to stress due to impairment of nucleolar function. In mammalian cells, disruption of nucleolar function leads to perturbation of the nucleolar structure, cell cycle arrest and cell death (James et al., 2014; Mayer & Grummt, 2005; Rubbi & Milner, 2003). Several studies have demonstrated similar severe growth phenotypes in trypanosomes upon

depletion of proteins involved in ribosome biogenesis (Cristodero & Clayton, 2007; Droll et al., 2010; Faktorová et al., 2018; Rink et al., 2019).

Hel66 depletion also resulted in accumulation of the first cleavage products of the rRNA processing pathway, presenting direct evidence of the involvement of this protein in ribosome biogenesis. The accumulation of the 3.4 kb and 5.8 kb intermediates indicates impaired cleavage of these processing intermediates after the initial cleavage at the B1 site located upstream of the 5.8S. This is apparent as seen in the decreased levels of downstream intermediates (0.61kb and 5.0kb). In yeast, there are reports that a feedback effect from the loss of *trans*-acting factors required for processing of the 60S rRNA inhibits or slows processing of 18S rRNA (Venema & Tollervey, 1999). Although the precise timing of the action of Hel66 cannot be defined due to varying degree of stability of the processing intermediates, the data shows an involvement of Hel66 in rRNA processing pathways of both ribosomal subunits. A slightly larger processing intermediate of the 5.8S rRNA processing pathway was detected 48 h post-induction of Hel66 depletion. This indicated that interruption of the processing pathway resulted in stabilisation of a yet undescribed processing intermediate of the 5.8S rRNA. Loss of Hel66 therefore prevents normal rRNA processing of both the 40S and 60S subunits thereby potentially altering the normal composition of the ribosomal subunits (Rink et al., 2019).

4.6 Conclusion and future perspectives

The role of the 100 % conserved 16mer motif in the maintenance of *VSG* mRNA stability has already been established although the mechanism for the stabilisation is still unclear. As a stabilisation motif, the 100 % conservation of the 16mer motif is surprising and raises the question as to whether the motif plays additional roles besides *VSG* mRNA stability. In this study, it was established that indeed the intact 16mer motif is not essential for expression of functional levels of VSG. A new role of the motif has also been identified, as the 100 % conservation of the motif was found to be essential for triggering efficient VSG silencing and coat exchange during an *in situ* VSG switch. This additional role of the 16mer motif in antigenic variation, which is an important process in the parasite, may be the driving force for the 100 % conservation of the motif in all VSG isoforms. As experiments mimicking *in situ* switch in the most direct approach were used to investigate the VSG switching, future studies could directly investigate this in the parasite using a suitable *in vivo* switch assay.

Although the direct interaction partners of the 16mer motif could not be identified in this study, Hel66 — an essential protein involved in rRNA processing — was characterised. The process of rRNA biogenesis is significantly different in trypanosomes compared to opisthokonts. The

characterisation of Hel66, a trypanosome specific DExD/H helicase involved in rRNA processing of both ribosomal subunits, might ultimately yield insights into the unique features of the rRNA processing pathways in trypanosomes. Hel66 could be characterised further by exploring its association with other proteins, as this may present information on the direct role of the protein in the rRNA processing pathway.

5.0 REFERENCES

Acosta-Serrano, A., Cole, R. N., Mehlert, A., Lee, M. G. S., Ferguson, M. A. J., & Englund, P. T. (1999). The procyclin repertoire of *Trypanosoma brucei*. Identification and structural characterization of the Glu-Pro-rich polypeptides. *Journal of Biological Chemistry*, *274*, 29763–29771.

Acosta-Serrano, A., Vassella, E., Liniger, M., Renggli, C. K., Brun, R., Roditi, I., & Englund, P. T. (2001). The surface coat of procyclic *Trypanosoma brucei*: Programmed expression and proteolytic cleavage of procyclin in the tsetse fly. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 1513–1518.

Agabian, N. (1990). Trans Splicing of Nuclear Pre-mRNAs. *Cell*, *61*, 1157–1160.

Alsford, S., & Horn, D. (2012). Cell-cycle-regulated control of VSG expression site silencing by histones and histone chaperones ASF1A and CAF-1b in *Trypanosoma brucei*. *Nucleic Acids Research*, *40*, 10150–10160.

Alsford, S., Kawahara, T., Glover, L., & Horn, D. (2005). Tagging a *T. brucei* RRNA locus improves stable transfection efficiency and circumvents inducible expression position effects. *Molecular and Biochemical Parasitology*, *144*, 142–148.

Alsford, S., Kawahara, T., Isamah, C., & Horn, D. (2007). A sirtuin in the African trypanosome is involved in both DNA repair and telomeric gene silencing but is not required for antigenic variation. *Molecular Microbiology*, *63*, 724–736.

Amodeo, S., Jakob, M., & Ochsenreiter, T. (2018). Characterization of the novel mitochondrial genome replication factor MiRF172 in *Trypanosoma brucei*. *Journal of Cell Science*, *131*, jcs211730.

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

Aroko, E. O. (2021). *Trans-regulation of* Trypanosoma brucei *variant surface glycoprotein (VSG) mRNA and structural analysis of a* Trypanosoma vivax *VSG using X-ray crystallography*. Universität Würzburg.

Aroko, E. O., Bakari-Soale, M., Batram, C., Jones, N. G., & Engstler, M. (2021). Endoplasmic reticulum-targeting but not translation is required for mRNA balancing in trypanosomes. *BioRxiv : The Preprint Server for Biology*.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., … Wang, H. (2010). TriTrypDB: A functional genomic resource for the Trypanosomatidae. *Nucleic Acids Research*, *38*, 457–462.

Bakari-Soale, M., Ikenga, N. J., Scheibe, M., Butter, F., Jones, N. G., Kramer, S., & Engstler, M. (2021). The nucleolar DExD/H protein Hel66 is involved in ribosome biogenesis in *Trypanosoma brucei*. *Scientific Reports*, *11*, 18325.

Barandun, J., Chaker-Margot, M., Hunziker, M., Molloy, K. R., Chait, B. T., & Klinge, S. (2017). The complete structure of the small-subunit processome. *Nature Structural and*

Molecular Biology, *24*, 944–953.

Barnwell, E. M., Van Deursen, F. J., Jeacock, L., Smith, K. A., Maizels, R. M., Acosta-Serrano, A., & Matthews, K. (2010). Developmental regulation and extracellular release of a VSG expression-site-associated gene product from *Trypanosoma brucei* bloodstream forms. *Journal of Cell Science*, *123*, 3401–3411.

Batram, C. (2009). *Functional analysis of the variant surface glycoprotein 3' untranslated region in* Trypanosoma brucei. Technische Universität Darmstadt.

Batram, C. (2013). *Die Kontrolle der monoallelen Expression, antigenen Variation und Entwicklung in* Trypanosoma brucei. Universität Würzburg.

Batram, C., Jones, N. G., Janzen, C. J., Markert, S. M., & Engstler, M. (2014). Expression site attenuation mechanistically links antigenic variation and development in *Trypanosoma brucei. ELife*, *3*, e02324.

Benmerzouga, I., Concepción-Acevedo, J., Kim, H. S., Vandoros, A. V., Cross, G. A. M., Klingbeil, M. M., & Li, B. (2013). *Trypanosoma brucei* Orc1 is essential for nuclear DNA replication and affects both VSG silencing and VSG switching. *Molecular Microbiology*, *87*, 196–210.

Berberof, M., Vanhamme, L., Tebabi, P., Pays, A., Jefferies, D., Welburn, S., & Pays, E. (1995). The 3'-terminal region of the mRNAs for VSG and procyclin can confer stage specificity to gene expression in *Trypanosoma brucei*. *The EMBO Journal*, *14*, 2925–2934.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., … El-Sayed, N. M. (2005). The genome of the African trypanosome *Trypanosoma brucei. Science*, *309*, 416–422.

Blum, M. L., Down, J. A., Gurnett, A. M., Carrington, M., Turner, M. J., & Wiley, D. C. (1993). A structural motif in the variant surface glycoproteins of *Trypanosoma brucei. Nature*, *362*, 603–609.

Boisvert, F. M., Ahmad, Y., Gierliński, M., Charrière, F., Lamont, D., Scott, M., … Lamond, A. I. (2012). A quantitative spatial proteomics analysis of proteome turnover in human cells. *Molecular and Cellular Proteomics*, *11*.M111.011.429

Borst, P., & Ulbert, S. (2001). Control of VSG gene expression sites. *Molecular and Biochemical Parasitology*, *114*(1), 17–27. https://doi.org/10.1016/S0166-6851(01)00243-2

Bouteille, B., & Buguet, A. (2012). The detection and treatment of human African trypanosomiasis. *Research and Reports in Tropical Medicine*, *3*, 35-45.

Brandenburg, J., Schimanski, B., Nogoceke, E., Nguyen, T. N., Padovan, J. C., Chait, B. T., … Günzl, A. (2007). Multifunctional class I transcription in *Trypanosoma brucei* depends on a novel protein complex. *EMBO Journal*, *26*, 4856–4866.

Brun, R., & Schönenberger, M. (1981). Stimulating effect of citrate and cis-aconitate on the transformation of *Trypanosoma brucei* bloodstream forms to procyclic forms in vitro. *Zeitschrift Für Parasitenkunde Parasitology Research*, *66*, 17–24.

Burkard, G., Fragoso, C. M., & Roditi, I. (2007). Highly efficient stable transformation of bloodstream forms of *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, *153*, 220–223.

Büscher, P., Cecchi, G., Jamonneau, V., & Priotto, G. (2017). Human African trypanosomiasis. *The Lancet*, *390*, 2397–2409.

Campbell, D. A., Kubo, K., Clark, C. G., & Boothroyd, J. C. (1987). Precise identification of cleavage sites involved in the unusual processing of trypanosome ribosomal RNA. *Journal of Molecular Biology*, *196*, 113–124.

Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Shimer, S., & Cerami, A. (1986). Identification of a common nucleotide sequence in the 3'-untranslated region of mRNA molecules specifying inflammatory mediators. *Proceedings of the National Academy of Sciences of the United States of America*, *83*, 1670–1674.

Carrington, M., Miller, N., Blumn, M., Roditi, I., Wiley, D., & Turner, M. (1991). Variant specific glycoprotein of *Trypanosoma brucei* consists of two domains each having an independently conserved pattern of cysteine residues. *Journal of Molecular Biology*, *221*, 823-835.

Carroll, J. S., Munchel, S. E., & Weis, K. (2011). The DExD/H box ATpase Dhh1 functions in translational repression, mRNA decay, and processing body dynamics. *Journal of Cell Biology*, *194*, 527–537.

Cesaro, G., Carneiro, F. R. G., Ávila, A. R., Zanchin, N. I. T., & Guimarães, B. G. (2019). *Trypanosoma brucei* RRP44 is involved in an early stage of large ribosomal subunit RNA maturation. *RNA Biology*, *16*, 133–143.

Cestari, I., & Stuart, K. (2015). Inositol phosphate pathway controls transcription of telomeric expression sites in trypanosomes. *Proceedings of the National Academy of Sciences*, *112*, e2803–2812.

Cestari, I., & Stuart, K. (2018). Transcriptional Regulation of Telomeric Expression Sites and Antigenic Variation in Trypanosomes. *Current Genomics*, *19*, 119–132.

Chaves, I., Rudenko, G., Dirks-Mulder, A., Cross, M., & Borst, P. (1999). Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *EMBO Journal*, *18*, 4846–4855.

Chen, C. Y. A., & Shyu, A. Bin. (1995). AU-rich elements: characterization and importance in mRNA degradation. *Trends in Biochemical Sciences*, *20*, 465–470.

Clayton, C. (2019). Regulation of gene expression in trypanosomatids: Living with polycistronic transcription. *Open Biology*, *9*, 190072

Clayton, C. E. (2014). Networks of gene expression regulation in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, *195*, 96–106.

Clayton, C., & Shapira, M. (2007). Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Molecular and Biochemical Parasitology*, *156*, 93–101.

Clayton, Christine E. (2002). Life without transcriptional control? From fly to man and back

again. *EMBO Journal*, *21*, 1881–1888.

Cordingley, J. S., & Turner, M. J. (1980). 6.5 S RNA; Preliminary characterisation of unusual small RNAs in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, *1*, 91–96.

Cristodero, M., & Clayton, C. E. (2007). Trypanosome MTR4 is involved in rRNA processing. *Nucleic Acids Research*, *35*, 7023–7030.

Cross, G. A. (1975). Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology*, *71*(3), 393–417.

Cross, G. A. M., Kim, H. S., & Wickstead, B. (2014). Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Molecular and Biochemical Parasitology*, *195*, 59–73.

Czichos, J., Nonnengaesser, C., & Overath, P. (1986). *Trypanosoma brucei*: cis-Aconitate and temperature reduction as triggers of synchronous transformation of bloodstream to procyclic trypomastigotes in vitro. *Experimental Parasitology*, *62*, 283–291.

da Silva, M. S., Hovel-Miner, G. A., Briggs, E. M., Elias, M. C., & McCulloch, R. (2018). Evaluation of mechanisms that may generate DNA lesions triggering antigenic variation in African trypanosomes. *PLoS Pathogens*, *14*, e1007321.

Dagnachew, S., Terefe, G., Abebe, G., Barry, D., McCulloch, R., & Goddeeris, B. (2015). In vivo experimental drug resistance study in *Trypanosoma vivax* isolates from tsetse infested and non-tsetse infested areas of Northwest Ethiopia. *Acta Tropica*, *146*, 95–100.

Daniels, J., Gull, K., & Wickstead, B. (2010). Cell Biology of the Trypanosome Genome. *Microbiology and Molecular Biology Reviews*, *74*, 552–569.

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

Dean, S., Marchetti, R., Kirk, K., & Matthews, K. R. (2009). A surface transporter family conveys the trypanosome differentiation signal. *Nature*, *459*, 213–217.

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

Deitsch, K. W., Lukehart, S. A., & Stringer, J. R. (2009). Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nature Reviews Microbiology*, *7*, 493–503.

Denninger, V., Fullbrook, A., Bessat, M., Ersfeld, K., & Rudenko, G. (2010). The FACT subunit TbSpt16 is involved in cell cycle specific control of VSG expression sites in *Trypanosoma brucei*. *Molecular Microbiology*, *78*, 459–474.

Desquesnes, M., Holzmuller, P., Lai, D. H., Dargantes, A., Lun, Z. R., & Jittaplapong, S. (2013). *Trypanosoma evansi* and surra: A review and perspectives on origin, history, distribution, taxonomy, morphology, hosts, and pathogenic effects. *BioMed Research International*, *2013*, 194176.

do Nascimento, L. M., Egler, F., Arnold, K., Papavasiliou, N., Clayton, C., & Erben, E. (2021). Functional insights from a surface antigen mRNA-bound proteome. *ELife*, *10*, e68136.

Donelson, J. E., & Rice-Ficht, A. C. (1985). Molecular biology of trypanosome antigenic variation. *Microbiological Reviews*, *49*, 107–125.

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

Droll, D., Minia, I., Fadda, A., Singh, A., Stewart, M., Queiroz, R., & Clayton, C. (2013). Post-Transcriptional Regulation of the Trypanosome Heat Shock Response by a Zinc Finger Protein. *PLoS Pathogens*, *9*. e1003286.

DuBois, K. N., Alsford, S., Holden, J. M., Buisson, J., Swiderski, M., Bart, J. M., … Field, M. C. (2012). NUP-1 is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. *PLoS Biology*, *10*. e1001287

Duraisingh, M. T., & Horn, D. (2016). Epigenetic Regulation of Virulence Gene Expression in Parasitic Protozoa. *Cell Host and Microbe*, *19*, 629–640.

Ehlers, B., Czichos, J., & Overath, P. (1987). RNA turnover in *Trypanosoma brucei. Molecular and Cellular Biology*, *7*, 1242–1249.

El-Sayed, N. M., Hegde, P., Quackenbush, J., Melville, S. E., & Donelson, J. E. (2000). The African trypanosome genome. *International Journal for Parasitology*, *30*, 329–345.

Engstler, M., & Boshart, M. (2004). Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in *Trypanosoma brucei*. *Genes and Development*, *18*, 2798–2811.

Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N., & Overath, P. (2007). Hydrodynamic Flow-Mediated Protein Sorting on the Cell Surface of Trypanosomes. *Cell*, *131*, 505–515.

Erben, E., Chakraborty, C., & Clayton, C. (2013). The CAF1-NOT complex of trypanosomes. *Frontiers in Genetics*, *4*, 299.

Erben, E. D., Fadda, A., Lueong, S., Hoheisel, J. D., & Clayton, C. (2014). A Genome-Wide Tethering Screen Reveals Novel Potential Post-Transcriptional Regulators in *Trypanosoma brucei*. *PLoS Pathogens*, *10*(6). e1004178.

Ersfeld, K. (2011). Nuclear architecture, genome and chromatin organisation in Trypanosoma brucei. *Research in Microbiology*, *162*, 626–636.

Estévez, A. M. (2008). The RNA-binding protein TbDRBD3 regulates the stability of a specific subset of mRNAs in trypanosomes. *Nucleic Acids Research*, *36*, 4573–4586.

Estévez, A. M., Kempf, T., & Clayton, C. (2001). The exosome of *Trypanosoma brucei. EMBO Journal*, *20*, 3831–3839.

Fadda, A., Ryten, M., Droll, D., Rojas, F., Färber, V., Haanstra, J. R., … Clayton, C. (2014). Transcriptome-wide analysis of trypanosome mRNA decay reveals complex degradation

kinetics and suggests a role for co-transcriptional degradation in determining mRNA levels. *Molecular Microbiology*, *94*, 307–326.

Fairman-Williams, M. E., Guenther, U. P., & Jankowsky, E. (2010). SF1 and SF2 helicases: Family matters. *Current Opinion in Structural Biology*, *20*, 313–324.

Faktorová, D., Bär, A., Hashimi, H., McKenney, K., Horák, A., Schnaufer, A., … Lukeš, J. (2018). TbUTP10, a protein involved in early stages of pre-18S rRNA processing in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, *225*, 84–93.

Faria, J., Glover, L., Hutchinson, S., Boehm, C., Field, M. C., & Horn, D. (2019). Monoallelic expression and epigenetic inheritance sustained by a *Trypanosoma brucei* variant surface glycoprotein exclusion complex. *Nature Communications*, *10*, 25–28.

Ferguson, M. A. J., Homans, S. W., Dwek, R. A., & Rademacher, T. W. (1988). Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. *Science*, *239*, 753–759.

Fernández-Moya, S. M., Carrington, M., & Estévez, A. M. (2014). A short RNA stem-loop is necessary and sufficient for repression of gene expression during early logarithmic phase in trypanosomes. *Nucleic Acids Research*, *42*, 7201–7209.

Figueiredo, L. M., Janzen, C. J., & Cross, G. A. M. (2008). A histone methyltransferase modulates antigenic variation in African trypanosomes. *PLoS Biology*, *6*, 1539–1548.

Finelle, P. (1973). African trypanosomiasis. Part I. Disease and Chemotherapy. *World Animal Review*, *8*, 24–29.

Fuller-Pace, F. V. (2006). DExD/H box RNA helicases: Multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Research*, *34*, 4206–4215.

Furger, A., Schürch, N., Kurath, U., & Roditi, I. (1997). Elements in the 3' untranslated region of procyclin mRNA regulate expression in insect forms of *Trypanosoma brucei* by modulating RNA stability and translation. *Molecular and Cellular Biology*, *17*, 4372–4380.

Gargantini, Pablo R., Lujan, H. D., & Pereira, C. A. (2012). In silico analysis of trypanosomatids' helicases. *FEMS Microbiology Letters*, *335*(2), 123–129.

Gargantini, Pablo Rubén, Serradell, M. del C., Ríos, D. N., Tenaglia, A. H., & Luján, H. D. (2016). Antigenic variation in the intestinal parasite *Giardia lamblia*. *Current Opinion in Microbiology*, *32*, 52–58.

Gerhardy, S., Menet, A. M., Peña, C., Petkowski, J. J., & Panse, V. G. (2014). Assembly and nuclear export of pre-ribosomal particles in budding yeast. *Chromosoma*, *123*, 327–344.

Gillian, A. L., & Svaren, J. (2004). The Ddx20/DP103 Dead Box Protein Represses Transcriptional Activation by Egr2/Krox-20. *Journal of Biological Chemistry*, *279*, 9056–9063.

Giordani, F., Morrison, L. J., Rowan, T. G., De Koning, H. P., & Barrett, M. P. (2016). The animal trypanosomiases and their chemotherapy: A review. *Parasitology*, *143*, 1862–1889.

Glover, L., Hutchinson, S., Alsford, S., & Horn, D. (2016). VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes. *Proceedings of the National Academy of Sciences*, *113*, 7225–7230.

Glover, L., Hutchinson, S., Alsford, S., Mcculloch, R., Field, M. C., & Horn, D. (2013). Antigenic variation in African trypanosomes: The importance of chromosomal and nuclear context in VSG expression control. *Cellular Microbiology*, *15*, 1984–1993.

Goodman, C. A., & Hornberger, T. A. (2013). Measuring protein synthesis with SUnSET: A valid alternative to traditional techniques? *Exercise and Sport Sciences Reviews*, *41*, 107–115.

Goos, C. (2013). *Funktionsanalyse des variablen Oberflächenproteins von* Trypanosoma brucei *– Die Rolle von N-verknüpften Glykanen*. Universität Würzburg.

Goos, C., Dejung, M., Janzen, C. J., Butter, F., & Kramer, S. (2017). The nuclear proteome of *Trypanosoma brucei*. *PLoS ONE*, *12*, e0181884.

Goos, C., Dejung, M., Wehman, A. M., M-Natus, E., Schmidt, J., Sunter, J., … Kramer, S. (2019). Trypanosomes can initiate nuclear export co-transcriptionally. *Nucleic Acids Research*, *47*, 266–282.

Grace, D., Randolph, T., Affognon, H., Dramane, D., Diall, O., & Clausen, P. H. (2009). Characterisation and validation of farmers' knowledge and practice of cattle trypanosomosis management in the cotton zone of West Africa. *Acta Tropica*, *111*, 137–143.

Graham, S. V., Wymer, B., & Barry, J. D. (1998). Activity of a Trypanosome Metacyclic Variant Surface Glycoprotein Gene Promoter Is Dependent upon Life Cycle Stage and Chromosomal Context. *Molecular and Cellular Biology*, *18*, 1137–1146.

Grünfelder, C. G., Engstler, M., Weise, F., Schwarz, H., Stierhof, Y. D., Boshart, M., & Overath, P. (2002). Accumulation of a GPI-anchored protein at the cell surface requires sorting at multiple intracellular levels. *Traffic*, *3*, 547–559.

Gruszynski, A. E., van Deursen, F. J., Albareda, M. C., Best, A., Chaudhary, K., Cliffe, L. J., … Bangs, J. D. (2006). Regulation of surface coat exchange by differentiating African trypanosomes. *Molecular and Biochemical Parasitology*, *147*, 211–223.

Guizetti, J., & Scherf, A. (2013). Silence, activate, poise and switch! Mechanisms of antigenic variation in Plasmodium falciparum. *Cellular Microbiology*, *15*, 718–726.

Günzl, A. (2010). The pre-mRNA splicing machinery of trypanosomes: Complex or simplified? *Eukaryotic Cell*, *9*, 1159–1170.

Günzl, A., Bruderer, T., Laufer, G., Schimanski, B., Tu, L. C., Chung, H. M., … Lee, M. G. S. (2003). RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. *Eukaryotic Cell*, *2*, 542–551.

Hajduk, S. L., Moore, D. R., Vasudevacharya, J., Siqueira, H., Torri, A. F., Tytler, E. M., & Esko, J. D. (1989). Lysis of *Trypanosoma brucei* by a toxic subspecies of human high density lipoprotein. *Journal of Biological Chemistry*, *264*, 5210–5217.

Heckman, K. L., & Pease, L. R. (2007). Gene splicing and mutagenesis by PCR-driven overlap extension. *Nature Protocols*, *2*, 924–932.

Hehl, A., Vassella, E., Braun, R., & Roditi, I. (1994). A conserved stem-loop structure in the 3' untranslated region of procyclin mRNAs regulates expression in *Trypanosoma brucei. Proceedings of the National Academy of Sciences of the United States of America*, *91*, 370–374.

Henning, J. (2012). *Analyse der Funktion von N-Glykanen in variablen Oberflächenglykoproteinen von* Trypanosoma brucei *am Beispiel von MITat1.5*. Universität Würzburg.

Henras, A. K., Soudet, J., Gérus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A., & Henry, Y. (2008). The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cellular and Molecular Life Sciences*, *65*, 2334–2359.

Henras, Anthony K., Plisson-Chastang, C., O'Donohue, M. F., Chakraborty, A., & Gleizes, P. E. (2015). An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdisciplinary Reviews: RNA*, *6*, 225–242.

Hertz-Fowler, C., Figueiredo, L. M., Quail, M. A., Becker, M., Jackson, A., Bason, N., … Berriman, M. (2008). Telomeric expression sites are highly conserved in *Trypanosoma brucei. PLoS ONE*, *3*. e0003527.

Hirumi, H., & Hirumi, K. (1989). Continuous Cultivation of *Trypanosoma brucei* Blood Stream Forms in a Medium Containing a Low Concentration of Serum Protein without Feeder Cell Layers. *The Journal of Parasitology*, *75*, 985–989.

Horn, D. (2014). Antigenic variation in African trypanosomes. *Molecular and Biochemical Parasitology*, *195*, 123–129.

Horn, D., & Cross, G. A. M. (1997). Position-dependent and promoter-specific regulation of gene expression in *Trypanosoma brucei*. *EMBO Journal*, *16*, 7422–7431.

Hotz, H. R., Hartmann, C., Huober, K., Hug, M., & Clayton, C. (1997). Mechanisms of developmental regulation in *Trypanosoma brucei:* A polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation. *Nucleic Acids Research*, *25*, 3017–3025.

Hovel-Miner, G., Mugnier, M. R., Goldwater, B., Cross, G. A. M., & Papavasiliou, F. N. (2016). A Conserved DNA Repeat Promotes Selection of a Diverse Repertoire of *Trypanosoma brucei* Surface Antigens from the Genomic Archive. *PLoS Genetics*, *12*, e1005994.

Hutchinson, S., Glover, L., & Horn, D. (2016). High-resolution analysis of multi-copy variant surface glycoprotein gene expression sites in African trypanosomes. *BMC Genomics*, *17*, 806.

Ikenga, N. J. (2020). *Analysis of a Putative RNA-Binding Protein in* Trypanosoma brucei. Universität Würzburg.

Jackson, A. P., Berry, A., Aslett, M., Allison, H. C., Burton, P., Vavrova-Anderson, J., … Berriman, M. (2012). Antigenic diversity is generated by distinct evolutionary

mechanisms in African trypanosome species. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 3416–3421.

Jaé, N., Wang, P., Gu, T., Hühn, M., Palfi, Z., Urlaub, H., & Bindereif, A. (2010). Essential role of a trypanosome U4-specific sm core protein in small nuclear ribonucleoprotein assembly and splicing. *Eukaryotic Cell*, *9*, 379–386.

James, A., Wang, Y., Raje, H., Rosby, R., & DiMario, P. (2014). Nucleolar stress with and without p53. *Nucleus*, *5*, 402–426.

Jankowsky, E. (2011). RNA helicases at work: Binding and rearranging. *Trends in Biochemical Sciences*, *36*, 19–29.

Janzen, C. J., Lander, F., Dreesen, O., & Cross, G. A. M. (2004). Telomere length regulation and transcriptional silencing in KU80-deficient *Trypanosoma brucei*. *Nucleic Acids Research*, *32*, 6575–6584.

Jeacock, L., Faria, J., & Horn, D. (2018). Codon usage bias controls mRNA and protein abundance in trypanosomatids. *ELife*, *7*, 32496.

Jehi, S. E., Li, X., Sandhu, R., Ye, F., Benmerzouga, I., Zhang, M., … Li, B. (2014). Suppression of subtelomeric VSG switching by *Trypanosoma brucei* TRF requires its TTAGGG repeat-binding activity. *Nucleic Acids Research*, *42*, 12899–12911.

Jehi, S. E., Wu, F., & Li, B. (2014). *Trypanosoma brucei* TIF2 suppresses VSG switching by maintaining subtelomere integrity. *Cell Research*, *24*, 870–885.

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

Jensen, B. C., Wang, Q., Kifer, C. T., & Parsons, M. (2003). The NOG1 GTP-binding protein is required for biogenesis of the 60 S ribosomal subunit. *Journal of Biological Chemistry*, *278*, 32204–32211.

Jojic, B., Amodeo, S., Bregy, I., & Ochsenreiter, T. (2018). Distinct 3′ UTRs regulate the life-cycle-specific expression of two *TCTP* paralogs in *Trypanosoma brucei*. *Journal of Cell Science*, *131*, jcs206417.

Jones, N. G., Thomas, E. B., Brown, E., Dickens, N. J., Hammarton, T. C., & Mottram, J. C. (2014). Regulators of *Trypanosoma brucei* Cell Cycle Progression and Differentiation Identified Using a Kinome-Wide RNAi Screen. *PLoS Pathogens*, *10*, e1003886.

Kala, S., Mehta, V., Yip, C. W., Moshiri, H., Najafabadi, H. S., Ma, R., … Salavati, R. (2017). The interaction of a *Trypanosoma brucei* KH-domain protein with a ribonuclease is implicated in ribosome processing. *Molecular and Biochemical Parasitology*, *211*, 94–103.

Kassem, A., Pays, E., & Vanhamme, L. (2014). Transcription is initiated on silent variant surface glycoprotein expression sites despite monoallelic expression in *Trypanosoma brucei*. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 8943–8948.

Keene, J. D. (2007). RNA regulons: Coordination of post-transcriptional events. *Nature Reviews Genetics*, *8*, 533–543.

Kennedy, P. G. E. (2004). Human African trypanosomiasis of the CNS: Current issues and challenges. *Journal of Clinical Investigation*, *113*, 496–504.

Kennedy, P. G. E. (2008). The continuing problem of human African trypanosomiasis (sleeping sickness). *Annals of Neurology*, *64*, 116–126.

Kim, H. S., Park, S. H., Günzl, A., & Cross, G. A. M. (2013). MCM-BP Is Required for Repression of Life-Cycle Specific Genes Transcribed by RNA Polymerase I in the Mammalian Infectious Form of *Trypanosoma brucei*. *PLoS ONE*, *8*, e0057001.

Knoppe, T. N., Bauer, B., McDermott, J. J., Peregrine, A. S., Mehlitz, D., & Clausen, P. H. (2006). Isometamidium sensitivity of *Trypanosoma congolense* stocks from cattle in West Africa tested in mice and the drug incubation infectivity test. *Acta Tropica*, *97*, 108–116.

Kolev, N. G., Franklin, J. B., Carmi, S., Shi, H., Michaeli, S., & Tschudi, C. (2010). The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution. *PLoS Pathogens*, *6*, e1001090.

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

Kramer, S., Queiroz, R., Ellis, L., Hoheisel, J. D., Clayton, C., & Carrington, M. (2010). The RNA helicase DHH1 is central to the correct expression of many developmentally regulated mRNAs in trypanosomes. *Journal of Cell Science*, *123*, 699–711.

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

Kramer, Susanne. (2021). Nuclear mRNA maturation and mRNA export control: from trypanosomes to opisthokonts. *Parasitology, 148*, 1196–1218.

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

Kramer, Susanne, Marnef, A., Standart, N., & Carrington, M. (2012). Inhibition of mRNA maturation in trypanosomes causes the formation of novel foci at the nuclear periphery containing cytoplasmic regulators of mRNA fate. *Journal of Cell Science*, *125*, 2896–2909.

Kraus, A. J., Brink, B. G., & Siegel, T. N. (2019). Efficient and specific oligo-based depletion of rRNA. *Scientific Reports*, *9,* 12281.

Krause, M., & Hirsh, D. (1987). A trans-spliced leader sequence on actin mRNA in *C. elegans. Cell*, *49*(6), 753–761.

Kressler, D., Linder, P., & de la Cruz, J. (1999). Protein trans-Acting Factors Involved in Ribosome Biogenesis in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, *19*, 7897–7912.

Landeira, D., Bart, J. M., Van Tyne, D., & Navarro, M. (2009). Cohesin regulates VSG monoallelic expression in trypanosomes. *Journal of Cell Biology*, *186*, 243–254.

Lau, B., Cheng, J., Flemming, D., La Venuta, G., Berninghausen, O., Beckmann, R., & Hurt, E. (2021). Structure of the Maturing 90S Pre-ribosome in Association with the RNA Exosome. *Molecular Cell*, *81*, 293-303.

Lebowitz, J. H., Smith, H. Q., Rusche, L., & Beverley, S. M. (1993). Coupling of poly(A) site selection and trans-splicing in Leishmania. *Genes and Development*, *7*, 996–1007.

Li, B. (2015). DNA double-strand breaks and telomeres play important roles in *Trypanosoma brucei* antigenic variation. *Eukaryotic Cell*, *14*, 196–205.

Liang, X. H., Haritan, A., Uliel, S., & Michaeli, S. (2003). trans and cis splicing in trypanosomatids: Mechanism, factors, and regulation. *Eukaryotic Cell*, *2*, 830–840.

Linder, P., & Jankowsky, E. (2011). From unwinding to clamping — the DEAD box RNA helicase family. *Nature Publishing Group*, *12*, 505–516.

Lindner, A. K., Lejon, V., Chappuis, F., Seixas, J., Kazumba, L., Barrett, M. P., … Franco, J. R. (2019). New WHO guidelines for treatment of gambiense human African trypanosomiasis including fexinidazole: substantial changes for clinical practice. *The Lancet Infectious Diseases*, *20*, e38–46.

López-Farfán, D., Bart, J. M., Rojas-Barros, D. I., & Navarro, M. (2014). SUMOylation by the E3 Ligase TbSIZ1/PIAS1 Positively Regulates VSG Expression in *Trypanosoma brucei. PLoS Pathogens*, *10*, e1004545.

Lueong, S., Merce, C., Fischer, B., Hoheisel, J. D., & Erben, E. D. (2016). Gene expression regulatory networks in *Trypanosoma brucei:* Insights into the role of the mRNA-binding proteome. *Molecular Microbiology*, *100*, 457–471.

MacGregor, P., & Matthews, K. R. (2012). Identification of the regulatory elements controlling the transmission stage-specific gene expression of PAD1 in *Trypanosoma brucei. Nucleic Acids Research*, *40*, 7705–7717.

Mair, G., Shi, H., Li, H., Djikeng, A., Aviles, H. O., Bishop, J. R., … Tschudi, C. (2000). A new twist in trypanosome RNA metabolism: Cis-splicing of pre-mRNA. *Rna*, *6*, 163–169.

Maishman, L., Obado, S. O., Alsford, S., Bart, J. M., Chen, W. M., Ratushny, A. V., … Field, M. C. (2016). Co-dependence between trypanosome nuclear lamina components in nuclear stability and control of gene expression. *Nucleic Acids Research*, *44*, 10554–10570.

Marcello, L., & Barry, J. D. (2007). Analysis of the VSG gene silent archive in *Trypanosoma brucei* reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure. *Genome Research*, *17*, 1344–1352.

Martin, R., Straub, A. U., Doebele, C., & Bohnsack, M. T. (2013). DExD/H-box RNA helicases in ribosome biogenesis. *RNA Biology*, *10*, 4–18.

Martínez-Calvillo, S., Yan, S., Nguyen, D., Fox, M., Stuart, K., & Myler, P. J. (2003). Transcription of Leishmania major Friedlin chromosome 1 initiates in both directions

within a single region. *Molecular Cell*, *11*, 1291–1299.

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

Maudlin, I. E., Kelly, S., Schwede, A., & Carrington, M. (2020). VSG mRNA levels are regulated by the production of functional VSG protein. *Molecular and Biochemical Parasitology*, *241,* 111348

Mayer, C., & Grummt, I. (2005). Cellular Stress and Nucleolar Function. *Cell Cycle*, *4*, 1036–1038.

Mayho, M., Fenn, K., Craddy, P., Crosthwaite, S., & Matthews, K. (2006). Post-transcriptional control of nuclear-encoded cytochrome oxidase subunits in *Trypanosoma brucei:* Evidence for genome-wide conservation of life-cycle stage-specific regulatory elements. *Nucleic Acids Research*, *34*, 5312–5324.

McCulloch, R. (2004). Antigenic variation in African trypanosomes: Monitoring progress. *Trends in Parasitology*, *20*(3), 117–121.

Michaeli, S. (2012). *rRNA Biogenesis in Trypanosomes. In Binderief A. (eds) RNA Metabolism in Trypanosomes*. *Nucleic Acids and Molecular Biology, vol 28. Springer, Berlin, Heidelberg*.

Molina-Portela, M. P., Samanovic, M., & Raper, J. (2008). Distinct roles of apolipoprotein components within the trypanosome lytic factor complex revealed in a novel transgenic mouse model. *Journal of Experimental Medicine*, *205*, 1721–1728.

Monahan, K., & Lomvardas, S. (2015). Monoallelic Expression of Olfactory Receptors. *Annual Review of Cell and Developmental Biology*, *31*, 721–740.

Monk, S. L., Simmonds, P., & Matthews, K. R. (2013). A short bifunctional element operates to positively or negatively regulate ESAG9 expression in different developmental forms of *Trypanosoma brucei*. *Journal of Cell Science*, *126*, 2294–2304.

Mony, B. M., MacGregor, P., Ivens, A., Rojas, F., Cowton, A., Young, J., … Matthews, K. (2014). Genome-wide dissection of the quorum sensing signalling pathway in *Trypanosoma brucei. Nature*, *505*, 681–685.

Mony, B. M., & Matthews, K. R. (2015). Assembling the components of the quorum sensing pathway in African trypanosomes. *Molecular Microbiology*, *96*, 220–232.

Morrison, L. J., Marcello, L., & McCulloch, R. (2009). Antigenic variation in the African trypanosome: Molecular mechanisms and phenotypic complexity. *Cellular Microbiology*, *11*, 1724–1734.

Mowatt, M. R., & Clayton, C. E. (1987). Developmental regulation of a novel repetitive protein of *Trypanosoma brucei*. *Molecular and Cellular Biology*, *7*, 2838–2844.

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

Muñoz-Jordán, J. L., Davies, K. P., & Cross, G. a. (1996). Stable expression of mosaic coats of variant surface glycoproteins in *Trypanosoma brucei*. *Science*, *272*, 1795–1797.

Murphy, W. J., Watkins, K. P., & Agabian, N. (1986). Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: Evidence for Trans splicing. *Cell*, *47*, 517–525.

Nakajima, T., Uchida, C., Anderson, S. F., Chee-Gun, L., Hurwitz, J., Parvin, J. D., & Montminy, M. (1997). RNA helicase A mediates association of CBP with RNA polymerase II. *Cell*, *90*, 1107–1112.

Nantulya, V. M. (1990). Trypanosomiasis in domestic animals: the problems of diagnosis. *Revue Scientifique et Technique (International Office of Epizootics)*, *9*(2), 357–367. https://doi.org/10.20506/rst.9.2.507

Narayanan, M. S., Kushwaha, M., Ersfeld, K., Fullbrook, A., Stanne, T. M., & Rudenko, G. (2011). NLP is a novel transcription regulator involved in VSG expression site control in *Trypanosoma brucei. Nucleic Acids Research*, *39*, 2018–2031.

Narayanan, M. S., & Rudenko, G. (2013). TDP1 is an HMG chromatin protein facilitating RNA polymerase i transcription in African trypanosomes. *Nucleic Acids Research*, *41*, 2981–2992.

Nascimento, J. de F., Kelly, S., Sunter, J., & Carrington, M. (2018). Codon choice directs constitutive mRNA levels in trypanosomes. *ELife*, *7*, e32467

Navarro, M., & Gull, K. (2001). A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. *Nature*, *414*, 759–763.

Nguyen, T. N., Müller, L. S. M., Park, S. H., Siegel, T. N., & Günzl, A. (2014). Promoter occupancy of the basal class I transcription factor A differs strongly between active and silent VSG expression sites in *Trypanosoma brucei*. *Nucleic Acids Research*, *42*, 3164–3176.

Nguyen, T. N., Nguyen, B. N., Lee, J. H., Panigrahi, A. K., & Günzl, A. (2012). Characterization of a novel class I transcription factor A (CITFA) subunit that is indispensable for transcription by the multifunctional RNA polymerase I of *Trypanosoma brucei*. *Eukaryotic Cell*, *11*, 1573–1581.

Nilsson, D., Gunasekera, K., Mani, J., Osteras, M., Farinelli, L., Baerlocher, L., … Ochsenreiter, T. (2010). Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of *Trypanosoma brucei*. *PLoS Pathogens*, *6*, e1001037.

Oberholzer, M., Morand, S., Kunz, S., & Seebeck, T. (2006). A vector series for rapid PCR-mediated C-terminal in situ tagging of *Trypanosoma brucei* genes. *Molecular and Biochemical Parasitology*, *14*, 117–120.

Ouellette, M., & Papadopoulou, B. (2009). Coordinated gene expression by post-transcriptional regulons in African trypanosomes. *Journal of Biology*, *8*, 1001-1004.

Overath, P., Czichos, J., & Haas, C. (1986). The effect of citrate/cis-aconitate on oxidative metabolism during transformation of *Trypanosoma brucei*. *European Journal of*

Biochemistry, *160*, 175–182.

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

Paterou, A., Walrad, P., Craddy, P., Fenn, K., & Matthews, K. (2006). Identification and stage-specific association with the translational apparatus of TbZFP3, a CCCH protein that promotes trypanosome life-cycle development. *Journal of Biological Chemistry*, *281*, 39002–39013.

Pays, E. (2005). Regulation of antigen gene expression in *Trypanosoma brucei. Trends in Parasitology*, *21*, 517–520.

Pays, E. (2006). The variant surface glycoprotein as a tool for adaptation in African trypanosomes. *Microbes and Infection*, *8*, 930–937.

Pays, E., & Vanhollebeke, B. (2009). Human innate immunity against African trypanosomes. *Current Opinion in Immunology*, *21*, 493–498.

Pedram, M., & Donelson, J. E. (1999). The anatomy and transcription of a monocistronic expression site for a metacyclic variant surface glycoprotein gene in *Trypanosoma brucei. Journal of Biological Chemistry*, *274*, 16876–16883.

Pena, A. C., Pimentel, M. R., Manso, H., Vaz-Drago, R., Pinto-Neves, D., Aresta-Branco, F., … Figueiredo, L. M. (2014). *Trypanosoma brucei* histone H1 inhibits RNA polymerase I transcription and is important for parasite fitness in vivo. *Molecular Microbiology*, *93*, 645–663.

Pérez-Morga, D., Vanhollebeke, B., Paturiaux-Hanocq, F., Nolan, D. P., Lins, L., Homblé, F., … Pays, E. (2005). Microbiology: Apolipoprotein L-I promotes trypanosome lysis by forming pores in lysosomal membranes. *Science*, *309*, 469–472.

Pitula, J., Ruyechan, W. T., & Williams, N. (2002). Two novel RNA binding proteins from *Trypanosoma brucei* are associated with 5S rRNA. *Biochemical and Biophysical Research Communications*, *290*, 569–576.

Polymenis, M., & Aramayo, R. (2015). Translate to divide: control of the cell cycle by protein synthesis. *Microbial Cell*, *2*, 94–104.

Povelones, M. L., Gluenz, E., Dembek, M., Gull, K., & Rudenko, G. (2012). Histone H1 Plays a Role in Heterochromatin Formation and VSG Expression Site Silencing in *Trypanosoma brucei*. *PLoS Pathogens*, *8*, e1003010.

Preußer, C., Jaé, N., & Bindereif, A. (2012). MRNA splicing in trypanosomes. *International Journal of Medical Microbiology*, *302*, 221–224.

Priotto, G., Kasparian, S., Mutombo, W., Ngouama, D., Ghorashian, S., Arnold, U., … Kande, V. (2009). Nifurtimox-eflornithine combination therapy for second-stage African Trypanosoma brucei gambiense trypanosomiasis: a multicentre, randomised, phase III, non-inferiority trial. *The Lancet*, *374*, 56–64.

Radwanska, M., Vereecke, N., Deleeuw, V., Pinto, J., & Magez, S. (2018). Salivarian

Trypanosomosis: A Review of Parasites Involved, Their Global Distribution and Their Interaction With the Innate and Adaptive Mammalian Host Immune System. *Frontiers in Immunology*, *9*, 2253.

Rajan, K. S., Chikne, V., Decker, K., Waldman Ben-Asher, H., & Michaeli, S. (2019). Unique Aspects of rRNA Biogenesis in Trypanosomatids. *Trends in Parasitology*, *35*(10), 778–794.

Rajkovic, A., Davis, R. E., Simonsen, J. N., & Rottman, F. M. (1990). A spliced leader is present on a subset of mRNAs from the human parasite *Schistosoma mansoni. Proceedings of the National Academy of Sciences of the United States of America*, *87*(22), 8879–8883.

Ramey-Butler, K., Ullu, E., Kolev, N. G., & Tschudi, C. (2015). Synchronous expression of individual metacyclic variant surface glycoprotein genes in *Trypanosoma brucei. Molecular and Biochemical Parasitology*, *200*, 1–4.

Raper, J., Fung, R., Ghiso, J., Nussenzweig, V., & Tomlinson, S. (1999). Characterization of a novel trypanosome lytic factor from human serum. *Infection and Immunity*, *67*, 1910–1916.

Redmond, S., Vadivelu, J., & Field, M. C. (2003). RNAit: An automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, *128*, 115–118.

Reis, H., Schwebs, M., Dietz, S., Janzen, C. J., & Butter, F. (2018). TelAP1 links telomere complexes with developmental expression site silencing in African trypanosomes. *Nucleic Acids Research*, *46*, 2820–2833.

Reuner, B., Vassella, E., Yutzy, B., & Boshart, M. (1997). Cell density triggers slender to stumpy differentiation of *Trypanosoma brucei* bloodstream forms in culture. *Molecular and Biochemical Parasitology*, *90*(1), 269–280.

Reynolds, D., Hofmeister, B. T., Cliffe, L., Alabady, M., Siegel, T. N., Schmitz, R. J., & Sabatini, R. (2016). Histone H3 Variant Regulates RNA Polymerase II Transcription Termination and Dual Strand Transcription of siRNA Loci in *Trypanosoma brucei*. *PLoS Genetics*, *12*, e1005758

Richardson, J. P., Beecroft, R. P., Tolson, D. L., Liu, M. K., & Pearson, T. W. (1988). Procyclin: an unusual immunodominant glycoprotein surface antigen from the procyclic stage of African trypanosomes. *Molecular and Biochemical Parasitology*, *31*, 203–216.

Rico, E., Ivens, A., Glover, L., Horn, D., & Matthews, K. R. (2017). Genome-wide RNAi selection identifies a regulator of transmission stage-enriched gene families and cell-type differentiation in *Trypanosoma brucei*. *PLoS Pathogens*, *13*, e1006279.

Rico, E., Rojas, F., Mony, B. M., Szoor, B., MacGregor, P., & Matthews, K. R. (2013). Bloodstream form pre-adaptation to the tsetse fly in *Trypanosoma brucei*. *Frontiers in Cellular and Infection Microbiology*, *3*, 78.

Ridewood, S., Ooi, C. P., Hall, B., Trenaman, A., Wand, N. V., Sioutas, G., … Rudenko, G. (2017). The role of genomic location and flanking 3′UTR in the generation of functional levels of variant surface glycoprotein in *Trypanosoma brucei. Molecular Microbiology*,

106(4), 614–634.

Rink, C., Ciganda, M., & Williams, N. (2019). The Nuclear Export Receptors TbMex67 and TbMtr2 Are Required for Ribosome Biogenesis in *Trypanosoma brucei* . *MSphere*, *4*, e00343-19.

Robinson, N. P., Burman, N., Melville, S. E., & Barry, J. D. (1999). Predominance of Duplicative VSG Gene Conversion in Antigenic Variation in African Trypanosomes. *Molecular and Cellular Biology*, *19*, 5839–5846.

Roditi, I., Carrington, M., & Turner, M. (1987). Expression of a polypeptide containing a dipeptide repeat is confined to the insect stage of *Trypanosoma brucei. Nature, 325, 272–274.*

Roditi, I., Schwarz, H., Pearson, T. W., Beecroft, R. P., Liu, M. K., Richardson, J. P., … Overath, P. (1989). Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. *Journal of Cell Biology*, *108*, 737–746.

Rojas, F., Silvester, E., Young, J., Milne, R., Tettey, M., Houston, D. R., … Matthews, K. R. (2019). Oligopeptide Signaling through TbGPR89 Drives Trypanosome Quorum Sensing. *Cell*, *176*, 306–317.

Rose, C., Casas-Sánchez, A., Dyer, N. A., Solórzano, C., Beckett, A. J., Middlehurst, B., … Acosta-Serrano, Á. (2020). *Trypanosoma brucei* colonizes the tsetse gut via an immature peritrophic matrix in the proventriculus. *Nature Microbiology*, *5*, 909–916.

Rotureau, B., & Van Den Abbeele, J. (2013). Through the dark continent: African trypanosome development in the tsetse fly. *Frontiers in Cellular and Infection Microbiology*, *3*, 53.

Rubbi, C. P., & Milner, J. (2003). Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO Journal*, *22*, 6068–6077.

Sakyiama, J., Zimmer, S. L., Ciganda, M., Williams, N., & Read, L. K. (2013). Ribosome biogenesis requires a highly diverged XRN family 5′→3′ exoribonuclease for rRNA processing in *Trypanosoma brucei. Rna*, *19*, 1419–1431.

Salmon, D., Vanwalleghem, G., Morias, Y., Denoeud, J., Krumbholz, C., Lhommé, F., … Pays, E. (2012). Adenylate cyclases of *Trypanosoma brucei* inhibit the innate immune response of the host. *Science*, *337*, 463–466.

Sanghai, Z. A., Miller, L., Molloy, K. R., Barandun, J., Hunziker, M., Chaker-Margot, M., … Klinge, S. (2018). Modular assembly of the nucleolar pre-60S ribosomal subunit. *Nature*, *556*, 126–129.

Schell, D., Evers, R., Preis, D., Ziegelbauer, K., Kiefer, H., Lottspeich, F., … Overath, P. (1991). A transferrin-binding protein of *Trypanosoma brucei* is encoded by one of the genes in the variant surface glycoprotein gene expression site. *EMBO Journal*, *10*, 1061–1066.

Schindelin, J., Arganda-carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., … Cardona, A. (2012). Fiji : an open-source platform for biological-image analysis. *Nature Methods*, *9*, 676–682.

Schnare, M. N., Spencer, D. F., & Gray, M. W. (1983). Primary structures of four novel small ribosomal RNAs from *Crithidia fasciculata. Canadian Journal of Biochemistry and Cell Biology*, *61*, 38–45.

Schulz, D., Mugnier, M. R., Paulsen, E. M., Kim, H. S., Chung, C. wa W., Tough, D. F., … Debler, E. W. (2015). Bromodomain Proteins Contribute to Maintenance of Bloodstream Form Stage Identity in the African Trypanosome. *PLoS Biology*, *13,* e1002316.

Schulz, D., Zaringhalam, M., Papavasiliou, F. N., & Kim, H. S. (2016). Base J and H3.V Regulate Transcriptional Termination in *Trypanosoma brucei. PLoS Genetics*, *12,* e1005762.

Schuster, S., Krüger, T., Subota, I., Thusek, S., Rotureau, B., Beilhack, A., & Engstler, M. (2017). Developmental adaptations of trypanosome motility to the tsetse fly host environments unravel a multifaceted in vivo microswimmer system. *ELife*, *6*, e27656.

Schuster, S., Lisack, J., Subota, I., Zimmermann, H., Reuter, C., Müller, T., … Engstler, M. (2021). Unexpected plasticity in the life cycle of *Trypanosoma brucei*. *ELife*, *10*, e66028.

Schwede, A., & Carrington, M. (2010). Bloodstream form trypanosome plasma membrane proteins: Antigenic variation and invariant antigens. *Parasitology*, *137*, 2029–2039.

Schwede, A., Jones, N., Engstler, M., & Carrington, M. (2011). The VSG C-terminal domain is inaccessible to antibodies on live trypanosomes. *Molecular and Biochemical Parasitology*, *175*, 201–204.

Schwede, A., Macleod, O. J. S., MacGregor, P., & Carrington, M. (2015). How Does the VSG Coat of Bloodstream Form African Trypanosomes Interact with External Proteins? *PLoS Pathogens*, *11,* e1005259.

Seed, J. R., & Wenck, M. A. (2003). Role of the long slender to short stumpy transition in the life cycle of the African trypanosomes. *Kinetoplastid Biology and Disease*, *2*, 3.

Shajani, Z., Sykes, M. T., & Williamson, J. R. (2011). Assembly of bacterial ribosomes. *Annual Review of Biochemistry*, *80*, 501–526.

Shalgi, R., Lapidot, M., Shamir, R., & Pilpel, Y. (2005). A catalog of stability-associated sequence elements in 3' UTRs of yeast mRNAs. *Genome Biology*, *6*, 1–15.

Shaw, A. P. M., Cecchi, G., Wint, G. R. W., Mattioli, R. C., & Robinson, T. P. (2014). Mapping the economic benefits to livestock keepers from intervening against bovine trypanosomosis in Eastern Africa. *Preventive Veterinary Medicine*, *113*, 197–210.

Shaw, G., & Kamen, R. (1986). A conserved adenine-uridine sequence from the 3' untranslated region of granulocyte-monocyte colony stimulating factor messenger RNA mediates selective messenger RNA degradation. *Cell*, *46*, 659–668.

Siegel, T. Nicolai, Gunasekera, K., Cross, G. A. M., & Ochsenreiter, T. (2011). Gene expression in *Trypanosoma brucei*: Lessons from high-throughput RNA sequencing. *Trends in Parasitology*, *27*, 434–441.

Siegel, Tim Nicolai, Hekstra, D. R., Wang, X., Dewell, S., & Cross, G. A. M. (2010). Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and

identification of splicing and polyadenylation sites. *Nucleic Acids Research*, *38*, 4946–4957.

Silva Pereira, S., Jackson, A. P., & Figueiredo, L. M. (2021). Evolution of the variant surface glycoprotein family in African trypanosomes. *Trends in Parasitology*, 1–14.

Simarro, P. P., Cecchi, G., Franco, J. R., Paone, M., Diarra, A., Ruiz-Postigo, J. A., … Jannin, J. G. (2012). Estimating and Mapping the Population at Risk of Sleeping Sickness. *PLoS Neglected Tropical Diseases*, *6*, e0001859.

Simarro, P. P., Cecchi, G., Paone, M., Franco, J. R., Diarra, A., Ruiz, J. A., … Jannin, J. G. (2010). The Atlas of human African trypanosomiasis: A contribution to global mapping of neglected tropical diseases. *International Journal of Health Geographics*, *9*, 1–18.

Sloan, K. E., & Bohnsack, M. T. (2018). Unravelling the Mechanisms of RNA Helicase Regulation. *Trends in Biochemical Sciences*, *43*, 237–250.

Smith, T. K., Bringaud, F., Nolan, D. P., & Figueiredo, L. M. (2017). Metabolic reprogramming during the *Trypanosoma brucei* life cycle. *F1000Research*, *6*, 683.

Smith, T. K., Vasileva, N., Gluenz, E., Terry, S., Portman, N., Kramer, S., … Rudenko, G. (2009). Blocking variant surface glycoprotein synthesis in Trypanosoma brucei triggers a general arrest in translation initiation. *PLoS ONE*, *4*, e0007532.

Specht, A. (2013). *Charakterisierung der Funktion der N-Glykane des variablen Oberflächenglykoproteins MITat1.5 von* Trypanosoma brucei. universität Würzburg.

Stanne, T., Narayanan, M. S., Ridewood, S., Ling, A., Witmer, K., Kushwaha, M., … Rudenko, G. (2015). Identification of the ISWI chromatin remodeling complex of the early branching eukaryote *Trypanosoma brucei*. *Journal of Biological Chemistry*, *290*, 26954–26967.

Stockdale, C., Swiderski, M. R., Barry, J. D., & McCulloch, R. (2008). Antigenic variation in *Trypanosoma brucei*: Joining the DOTs. *PLoS Biology*, *6*, 1386–1391.

Strunk, B. S., & Karbstein, K. (2009). Powering through ribosome assembly. *Rna*, *15*, 2083–2104.

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

Sutton, R. E., & Boothroyd, J. C. (1986). Evidence for Trans splicing in trypanosomes. *Cell*, *47*, 527–535.

Tanner, N. K., & Linder, P. (2001). DExD/H box RNA helicases: From generic motors to specific dissociation functions. *Molecular Cell*, *8*, 251–262.

Taylor, J. E., & Rudenko, G. (2006). Switching trypanosome coats: what's in the wardrobe? *Trends in Genetics*, *22*, 614–620.

Tessier, L. H., Keller, M., Chan, R. L., Fournier, R., Weil, J. H., & Imbault, P. (1991). Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by trans-

splicing in Euglena. *EMBO Journal*, *10*, 2621–2625.

Tetley, L., & Vickerman, K. (1985). Differentiation in *Trypanosoma brucei*: Host-parasite cell junctions and their persistence during acquisition of the variable antigen coat. *Journal of Cell Science*, *74*, 1–19.

Turner, C. M. R., Aslam, N., & Dye, C. (1995). Replication, differentiation, growth and the virulence of *Trypanosoma brucei* infections. *Parasitology*, *111*, 289–300.

Umaer, K., Ciganda, M., & Williams, N. (2014). Ribosome biogenesis in African trypanosomes requires conserved and trypanosome-specific factors. *Eukaryotic Cell*, *13*, 727–737.

Urwyler, S., Studer, E., Renggli, C. K., & Roditi, I. (2007). A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of *Trypanosoma brucei. Molecular Microbiology*, *63*, 218–228.

Uzureau, P., Uzureau, S., Lecordier, L., Fontaine, F., Tebabi, P., Homblé, F., … Pays, E. (2013). Mechanism of *Trypanosoma brucei* gambiense resistance to human serum. *Nature*, *501*, 430–434.

Vandenberghe, A. E., Meedel, T. H., & Hastings, K. E. M. (2001). mRNA 5'-leader trans-splicing in the chordates Amanda. *Genes and Development*, *15*, 294–303.

Vanhamme, L., Berberof, M., Le Ray, D., & Pays, E. (1995). Stimuli of differentiation regulate RNA elongation in the transcription units for the major stage-specific antigens of *Trypanosoma brucei*. *Nucleic Acids Research*, *23*, 1862–1869.

Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D. P., Lins, L., Van Den Abbeele, J., … Pays, E. (2003). Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature*, *422*, 83–87.

Vanhollebeke, B., Nielsen, M. J., Watanabe, Y., Truc, P., Vanhamme, L., Nakajima, K., … Pays, E. (2007). Distinct roles of haptoglobin-related protein and apolipoprotein L-I in trypanolysis by human serum. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 4118–4123.

Vanhollebeke, B., Uzureau, P., Monteyne, D., Pérez-Morga, D., & Pays, E. (2010). Cellular and molecular remodeling of the endocytic pathway during differentiation of *Trypanosoma brucei* bloodstream forms. *Eukaryotic Cell*, *9*, 1272–1282.

Vassella, E., Probst, M., Schneider, A., Studer, E., Renggli, C. K., & Roditi, I. (2004). Expression of a Major Surface Protein of *Trypanosoma brucei* Insect Forms Is Controlled by the Activity of Mitochondrial Enzymes. *Mol Biol Cell*, *15*, 5318–5328.

Vassella, E., Van Den Abbeele, J., Bütikofer, P., Renggli, C. K., Furger, A., Brun, R., & Roditi, I. (2000). A major surface glycoprotein of *Trypanosoma brucei* is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia. *Genes and Development*, *14*, 615–626.

Venema, J., & Tollervey, D. (1999). Ribosome synthesis in *Saccharomyces cerevisiae. Annual Review of Genetics*, *33*, 261–311.

Vettermann, C., & Schlissel, M. S. (2010). Allelic exclusion of immunoglobulin genes: Models

and mechanisms. *Immunological Reviews*, *237*, 22–42.

Viegas, I. J., Macedo, J. P. De, Niz, M. De, Rodrigues, J. A., Aresta-branco, F., Jaffrey, S. R., & Figueiredo, L. M. (2020). N 6 -methyladenosine in poly (A) tails stabilize VSG transcripts. *BioRxiv Preprint*, 1–25.

Walrad, P. B., Capewell, P., Fenn, K., & Matthews, K. R. (2012). The post-transcriptional trans-acting regulator, TbZFP3, co-ordinates transmission-stage enriched mRNAs in Trypanosoma brucei. *Nucleic Acids Research*, *40*(7), 2869–2883. https://doi.org/10.1093/nar/gkr1106

Wang, Q. P., Kawahara, T., & Horn, D. (2010). Histone deacetylases play distinct roles in telomeric VSG expression site silencing in African trypanosomes. *Molecular Microbiology*, *77*, 1237–1245.

Wang, Y. N., Wang, M., & Field, M. C. (2010). *Trypanosoma brucei*: Trypanosome-specific endoplasmic reticulum proteins involved in variant surface glycoprotein expression. *Experimental Parasitology*, *125*, 208–221.

Wedel, C., Förstner, K. U., Derr, R., & Siegel, T. N. (2017). GT -rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes . *The EMBO Journal*, *36*, 2581–2594.

White, T. C., Rudenko, G., & Borst, P. (1986). Three small RNAs within the 10 kb trypanosome rRNA transcription unit are analogous to domain VII of other eukaryotic 28S rRNAs. *Nucleic Acids Research*, *14*, 9471–9489.

Wild, T., Horvath, P., Wyler, E., Widmann, B., Badertscher, L., Zemp, I., … Kutay, U. (2010). A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biology*, *8,* e1000522.

Wilson, B. J., Bates, G. J., Nicol, S. M., Gregory, D. J., Perkins, N. D., & Fuller-Pace, F. V. (2004). The p68 and p72 DEAD box RNA helicases interact with HDAC1 and repress transcription in a promoter-specific manner. *BMC Molecular Biology*, *5*, 1–15.

Wirtz, E., Leal, S., Ochatt, C., & Cross, G. A. M. (1999). A tightly regulated inducible expression system for dominant negative approaches in *Trypanosoma brucei*. *Molecular & Biochemical Parasitology*, *99*, 89–101.

Woolford, J. L., & Baserga, S. J. (2013). Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics*, *195*, 643–681.

World Health Organization, (WHO). (2020). Trypanosomiasis, human African (sleeping sickness). Retrieved from https://www.who.int/en/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)

Wu, S., Tutuncuoglu, B., Yan, K., Brown, H., Zhang, Y., Tan, D., … Gao, N. (2016). Diverse roles of assembly factors revealed by structures of late nuclear pre-60S ribosomes. *Nature*, *534*, 133–137.

Xong, H. Van, Vanhamme, L., Chamekh, M., Chimfwembe, C. E., Van Den Abbeele, J., Pays, A., … Pays, E. (1998). A VSG expression site-associated gene confers resistance to human serum in *Trypanosoma rhodesiense*. *Cell*, *95*, 839–846.

Yang, X., Figueiredo, L. M., Espinal, A., Okubo, E., & Li, B. (2009). RAP1 Is Essential for Silencing Telomeric Variant Surface Glycoprotein Genes in *Trypanosoma brucei*. *Cell*, *137*, 99–109.

Zhang, Y., Forys, J. T., Miceli, A. P., Gwinn, A. S., & Weber, J. D. (2011). Identification of DHX33 as a Mediator of rRNA Synthesis and Cell Growth. *Molecular and Cellular Biology*, *31*, 4676–4691.

Ziegelbauer, K., Quinten, M., Schwarz, H., Pearson, T. W., & Overath, P. (1990). Synchronous differentiation of *Trypanosoma brucei* from bloodstream to procyclic forms in vitro. *European Journal of Biochemistry*, *192*, 373–378.

Zimmermann, H., Subota, I., Batram, C., Kramer, S., Janzen, C. J., Jones, N. G., & Engstler, M. (2017). A quorum sensing-independent path to stumpy development in *Trypanosoma brucei. PLoS Pathogens*, *13*, e1006324.

6.0 APPENDIX

6.1 Supplementary figures

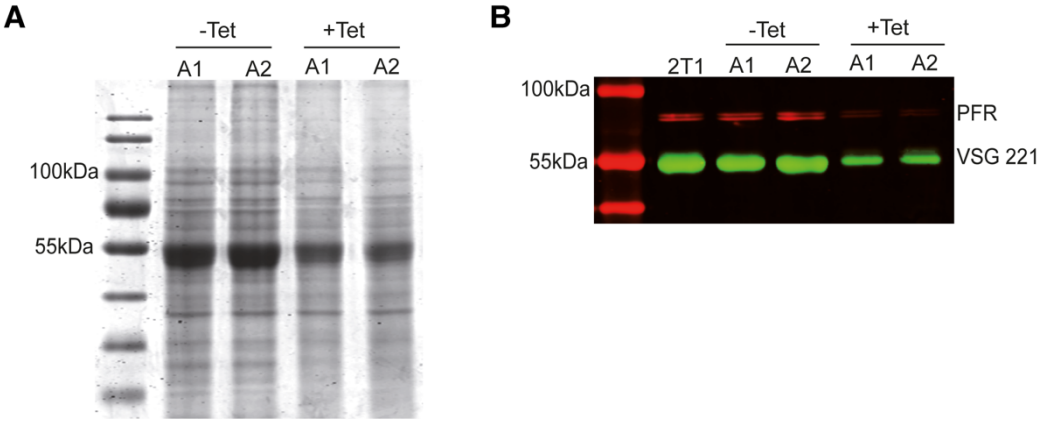

Figure 31. Depletion of RPC40 results in global reduction in protein amounts. Coomassie stained SDS-Page gel (A) and Western blot (B) showing reduced protein levels after RNAi depletion of RPC40 for 48 h.

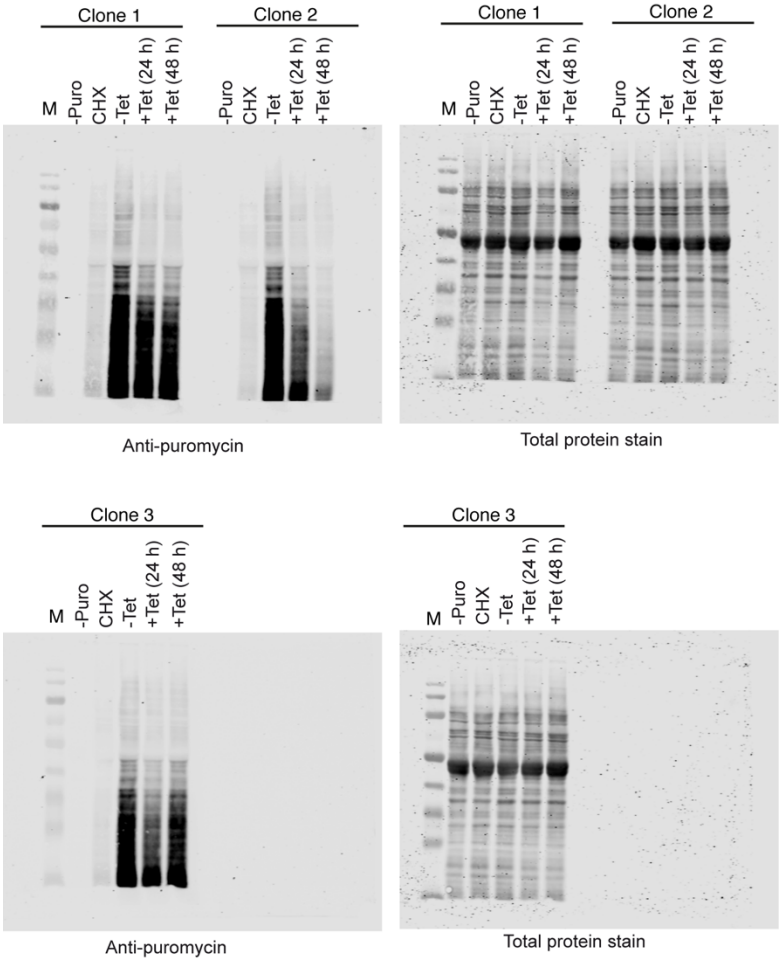

Figure 32. Loss of Hel66 results in global reduction in translation. Images showing data from all three clonal cell lines after SUnSET assay. The figure is adapted from Bakari-Soale et al., 2021.

6.2 Publications

Parts of this thesis are included in the following publications or manuscripts in preparation;

Bakari-Soale M., Ikenga N. J., Scheibe M., Butter F., Jones N. G., Kramer S., Engstler M. (2021). The nucleolar DExD/H protein Hel66 is involved in ribosome biogenesis in *Trypanosoma brucei. Sci Rep* 11, 18325. https://doi.org/10.1038/s41598-021-97020-0

Aroko E. O., **Bakari- Soale M**., Batram C., Jones N. G., Engstler M. (2021). Endoplasmic reticulum-targeting but not translation is required for mRNA balancing in trypanosomes. **Preprint**: bioRxiv. doi: 10.1101/2021.05.05.442555.

Bakari-Soale M., Zimmermann H., Batram C., Kramer S., Jones N. G., Engstler M. (2021). Dual role of the 16mer motif within the Variant Surface Glycoprotein of *Trypanosoma brucei*. **Manuscript in preparation**

6.3 Acknowledgement

I am grateful to God for his goodness and protection throughout my doctoral studies. I thank my primary supervisor Prof. Dr. Markus Engstler for his mentorship, support and encouragement during my studies. I also thank my other thesis advisory committee members, Prof. Dr. Klaus Brehm and Prof. Dr. Mark Carrington for the fruitful discussions and encouragement during our annual meetings.

I am grateful to Dr. Nicola Jones for the help and support I received during this work. Thank you for all the discussions on new ideas and experiments, and for proof reading of the thesis. I am also grateful to Prof. Dr. Susanne Kramer for the help with my experiments and manuscript writing. Thank you for also translating the summary of my thesis into German. A huge thank you to Dr. Brooke Morriswood for taking time out of his busy schedule to proof-read my thesis in such a short notice. Thank you to Dr. Falk Butter and Marion Scheibe for the collaborative work on identification of candidate 16mer binding proteins.

I would also like to thank past and present members of the department for creating a lovely working environment and always ready to help when needed. Special thanks to Reinhild, Kathrin, Lidia, Elizabeth Meyer-Natus, Elina and the entire technical staff for ensuring that things run smoothly in the lab. I also thank Uli and Manu for the immense help in taking care of all admirative documentation on time.

Thank you to all my friends who have been part of this challenging yet exciting journey. Special mention to Kingsley, Isaac, Erick and Paula for really being there for me this past few months.

Special thank you to my parents and my siblings for the love, support, patience and prayers. I appreciate the encouraging words I received whenever I was down.

I also thank the Graduate School of Life Sciences for awarding me the GSLS fellowship and providing the initial funds for my studies.

6.4 Curriculum vitae (CV)