Spring 5-1-2022

Disrupting Monoallelic Expression of Variant Surface Glycoprotein in Trypanosoma brucei by a Non-lethal Mutation in Class I Transcription Factor A

Sarah Platt
sarah.platt@uconn.edu

Disrupting Monoallelic Expression of Variant Surface Glycoprotein in *Trypanosoma brucei* by a Non-lethal Mutation in Class I Transcription Factor A

Sarah Platt

University of Connecticut

Department of Molecular and Cell Biology

Honors Thesis Advisor: Arthur Günzl, PhD

Honors Advisor: Aoife Heaslip, PhD

May 12th, 2022

ABSTRACT

Human African trypanosomiasis (HAT) is a lethal disease caused by protozoan hemoflagellates of the genus *Trypanosoma*. Humans are vulnerable to two subspecies, *Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense*. At the crux of HAT lethality lie two uncommon genetic expression phenomena: monoallelic expression and antigenic variation. Combined, these mechanisms effectively shield trypanosomes from host immune systems, prolonging infections. Variant Surface Glycoproteins (VSGs) are the key outer membrane proteins involved in antigenic variation. By continuously changing the composition of cell surface antigens, trypanosomes can survive bouts of immunological detection and eventually traverse the blood-brain barrier. There are over two thousand *VSG* variants within the trypanosome genome, but only one is expressed at a time. Transcription of the active *VSG* is initiated by the Class I Transcription Factor A (CITFA) complex. This basal promoter-binding complex is composed of eight subunits: CITFA1-7 and dynein light chain LC8. Preliminary data suggested that CITFA7 dimerizes via LC8 and is important for expression of only a single VSG at a time. Results of a co-immunoprecipitation assay provide evidence against the possibility that CITFA7 dimerizes *in vivo*, while semi-quantitative and quantitative reverse-transcription PCR data support the hypothesis that CITFA7 contributes to monoallelic expression. Visualization of silent VSG derepression with indirect immunofluorescence was attempted to see whether disruption of monoallelic expression by the expression of a non-lethal mutation in *CITFA7* alters the normally punctuated location of the CITFA complex within the nucleus. Results from this effort lay the groundwork for future quantitative analysis with direct fluorescence.

INTRODUCTION

Trypanosomatids, early diverged parasites of the order *Kinetoplastida,* give rise to a number of diseases including sleeping sickness, Chagas disease, and leishmaniasis. *Trypanosoma brucei* is a species within this order that is responsible for sleeping sickness, and nagana in cattle (or human African trypanosomiasis (HAT) and animal trypanosomiasis, respectively). Humans are vulnerable to two out of the three *T. brucei* subspecies. *Trypanosoma brucei gambiense* leads to a more chronic condition that prevails throughout central and western Africa, whereas *Trypanosoma brucei rhodesiense* leads to a more acute condition that is currently restricted to Uganda (WHO, 2022). *Trypanosoma brucei brucei*, the third subspecies, causes animal trypanosomiasis and poses minimal risk to humans due to its susceptibility to the trypanolytic factor of the human innate immune system (Radwanska et al., 2018). Thus, *T. b. brucei* can be utilized in the laboratory as a relatively safe proxy for studying gene expression in the human-infectious subspecies.

Research efforts have not yet amounted to an antiparasitic vaccine against HAT. However, cases of HAT have plummeted to as low as 663 cases in 2020 due to large-scale efforts to control vector spread and to diagnose and treat active cases in the early stages of the disease (CDC, 2019; Tikhonenkov et al., 2021). Despite these strides in containment, the disease still disproportionately affects people living in remote and impoverished areas with limited access to healthcare (WHO, 2022). Thus, the search for potential drug targets and vaccine design remain important efforts that could potentially also elucidate effective treatment strategies for related diseases.

Trypanosomes are transmitted by the tsetse fly and undergo eight morphological transitions that prepare them for the different environments they encounter over the course of their life cycle. When a tsetse fly bites an infected mammal, trypanosomes collect in the fly's midgut where they morph into their mitotically dividing procyclic form, and subsequently migrate to the fly's salivary glands as epimastigotes where they differentiate into non-proliferative metacyclic trypomastigotes. Following a blood meal, when metacyclics are transferred from the fly to the host, they often remain subcutaneous for several weeks, causing a skin chancre to form. Once they break into the bloodstream, metacyclics transition into their slender form (CDC, 2019; Horn, 2014; Ginger, 2002). At each of these life cycle stages, trypanosomes undergo differentiation which entails changes in morphology, nutrient-specific metabolism, organelle structure and organization, and epigenetic regulation of gene transcription to keep up with the specific challenges posed by each new environment. Consequently, the composition of cell surface proteins changes distinctly at each stage (Fenn et al., 2007; Quintana et al., 2021; Briggs et al., 2021).

As procyclics, trypanosomes express procyclins (or procyclic acidic repetitive protein) on their cell surface, while as metacyclics, trypanosomes monoallelically express metacyclic variant surface glycoproteins (mVSGs) as a preadaptation to mammalian blood. Up to seven days after entering the bloodstream of a host, *mVSG* transcription is silenced and replaced by monoallelic *VSG* expression. Both *mVSG*s and *VSG*s are transcribed and processed into functional mRNA through Spliced Leader (SL) *trans*-splicing (Ginger et al., 2002; Günzl, 2010). Most *VSG* genes are located in subtelomeric regions of chromosomes or harbored within minichromosomes as functional or pseudo-

*VSG*s. However, there are only fifteen *VSG*s in 427 strain trypanosomes contained within telomeric domains called bloodstream expression sites (BES), and it is only one BES *VSG* that is actively transcribed at a time. Monoallelic *VSG* expression results in a uniform cell surface coat, and VSG makes up an impressive 10% of all cell proteins expressed by slender forms at any given time (Van der Ploeg et al., 1982). Restriction of expression to a single gene variant is not unique to trypanosomes. In fact, genes encoding olfactory receptors and immunoglobulin genes have been found to exhibit monoallelic expression (Schultz and Papavasiliou, 2016; Monahan et al., 2015; Vettermann and Schlissel, 2010). However, for *T. brucei* and other infectious organisms such as the malaria-causing *Plasmodium falciparum*, monoallelic expression confers a particular adaptive advantage. By systematically altering the identity of their antigens through the process of antigenic variation, these parasites can evade detection by host immune systems, making them particularly pervasive and persistent in their hosts (Turner, 1984). Teasing apart the different molecular players that regulate monoallelic expression in these organisms is critical to understanding their pathogenesis and developing effective treatments against their associated diseases.

Monoallelic expression and antigenic variation work hand in hand. In *P. falciparum*, the *var* gene family—which encodes the adhesion surface molecule Erythrocyte Membrane Protein 1—is silenced by default, and one *var* gene is selectively activated early during the parasite's development in its host's blood. Although switching events of *var* gene variants are well characterized and essential to the parasite's survival, the molecular mechanisms of these events are still under investigation (Deitsch et al., 2017). Similarly, in *T. brucei*, monoallelic activation of *VSG* transcription occurs early in

development to prepare for the host's bloodstream, but unlike for *P. falciparum*, the mechanism underlying switching events have been investigated at length (Horn, 2014; Mugnier et al., 2015; Morrison et al., 2009). The three most common types of VSG switching events in *T. brucei* occur through gene conversion, telomere exchange, and transcriptional switching. In the former two cases, the VSG coat is changed through genetic recombination during meiotic division in the metacyclic and bloodstream stages whereby a *VSG* contained within a silent BES, a subtelomeric region, or a minichromosome replaces the *VSG* in an active BES. In the latter case, active *VSG*s undergo rare epigenetic switching events that lead to the monoallelic expression of a different *VSG* (Dreesen et al., 2006). Transcriptional switching is particularly useful in chronic infections; while trypanosomes expressing a particular VSG may trigger their host's adaptive immune response such that they are recognized by antibodies, other trypanosomes having randomly switched *VSG* expression may live on and proliferate because the host has not yet produced antibodies specific against the new VSG (Horn, 2014; Hertz-Fowler et al., 2008).

T. brucei has many unique features, including an unusual mode of transcription. Whereas in *P. falciparum*, *var* genes are transcribed by RNA polymerase II (RNA pol II) and *rRNA* is transcribed by RNA pol I, as in all other eukaryotes, *T. brucei* possesses a unique RNA pol I that is recruited to both *rRNA* promoters in the nucleolus as well as to promoters upstream of *procyclin* and *VSG* genes (Günzl et al., 2003). The active *VSG* gene is always located within a telomeric BES, which contains several expression site–associated genes (*ESAG*s) and a single *VSG* gene just upstream of the telomeric repeats (Figure 1). *T. brucei* Lister 427 cells, the strain used in this study, possesses exactly

fifteen BESs (Hertz-Fowler et al., 2008). Transcription of a single *VSG* from the active BES always occurs within a sub-nucleolar compartment called the expression site body (ESB), the formation of which is attributed to the trypanosome-specific exclusion protein VEX1 (Navarro and Gull, 2001; Faria et al., 2019).

BES repression occurs through telomeric silencing, accomplished, in part, by the telomere-binding protein RAP1 and the histone methyltransferase DOT1B (Figueiredo et al., 2008; Yang et al., 2009). Silencing almost completely prevents transcription of the downstream *VSG* gene but still allows for some transcriptional activity at the repressed BES promoter lying 50-70 Kb upstream of the telomeric repeats. (Hertz-Fowler et al., 2008). Interestingly, transcriptional activity occurring at BESs gradually decreases along its length such that promoter proximal regions are transcribed at levels far higher than telomere proximal regions. In one model of this phenomenon, the recruitment of histone modifying enzymes by telomeric proteins spreads repressive chromatin inwards to diminish transcription elongation along silent BESs (Yang et al., 2009; Wang et al., 2010). In one study, when the *RAP1* gene was knocked down, silent BESs became de-repressed but the activity levels of these BESs remained magnitudes below that of the active BES, suggesting that there is a separate, dedicated mechanism for transcription

initiation at BESs (Yang et al., 2009).

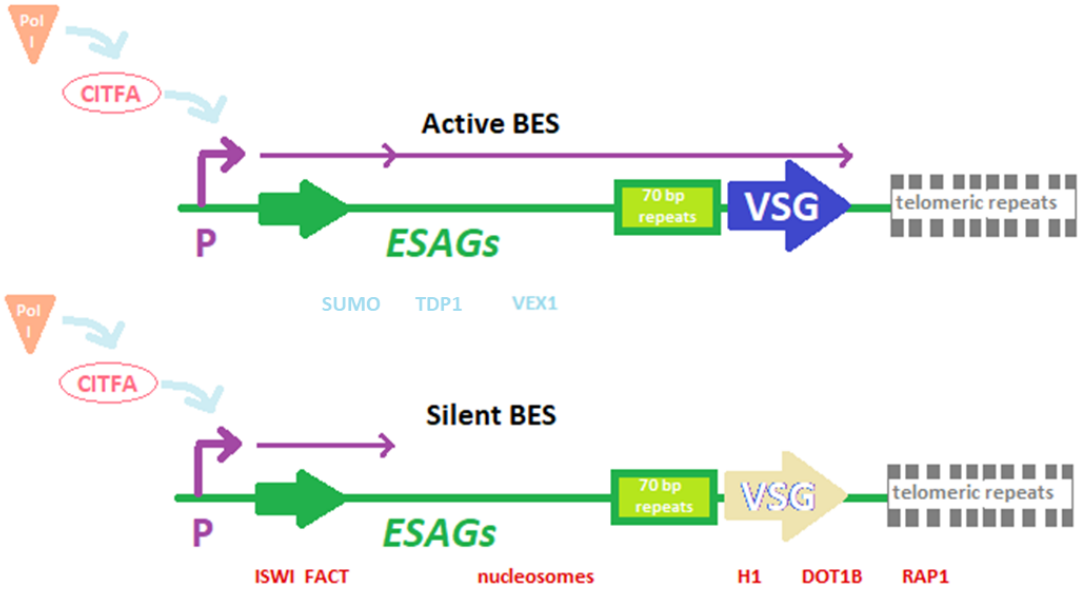


Figure 1 Schematic of active and silent bloodstream expression sites in *T. brucei*. Image (not to scale) shows RNA pol I, CITFA complex, promoter (P), *ESAG*s (Expression Site Associated Genes), 70 base pair repeats, *VSG* (variant surface glycoprotein), and telomeric repeats. Proteins associated with the active BES include SUMO (small ubiquitin-like modifier that positively regulates monoallelic expression by post-translational modification (Ye et al., 2015; Saura et al., 2019)), TDP1 (essential high-mobilitygroup box protein that facilitates RNA pol I transcription and is abundant in the nucleolus and ESB (Narayanan et al., 2013)) and VEX1 (RNA pol I-dependent-VSG exclusion protein which assembles with VEX2 to sustain the ESB (Glover et al., 2016)). Proteins associated with BES repression include ISWI (members of the ISWI chromatin-remodeling family are involved in BES downregulation in promoter-proximal regions (Hughes et al., 2007)), FACT (may reassemble nucleosomes behind RNA pol I as it elongates, repressing VSG transcription (Denninger, & Rudenko, 2014)), H1 (inhibits RNA pol I transcription (Pena et al., 2014)), DOT1B (chromatin-modifying enzyme required for complete BES silencing (Figueiredo et al., 2008)), and RAP1 (an intrinsic component of the *T. brucei* telomere complex that is essential for complete silencing of BESs (Yang et al., 2009).

Despite what is known about *VSG* expression, regulation of BES transcription activation remains particularly elusive. Characterization of the CITFA complex provided an important line of evidence that there is a distinct activation mechanism of monoallelic transcription. CITFA is composed of subunits CITFA1-7 as well as the highly conserved dimerizing protein, dynein light chain LC8 (Nguyen et al., 2012). CITFA directly binds to BES and rRNA gene promoters and is essential for RNA pol I to initiate transcription

(Brandenburg et al., 2007). When CITFA subunit genes were completely ablated, transcription of rRNA genes and *VSG* plummeted and cells died within two days, demonstrating the importance of CITFA for cell viability (Nguyen et al., 2012). Surprisingly, chromatin immunoprecipitation (ChIP) of either subunits CITFA7 or CITFA2 revealed that CITFA occupied the promoter of the active BES promoter several fold more than the promoter of a repressed BES, a phenotype that prevailed even after trypanosomes had switched consecutively between the two BESs. Since a genome-wide ChIP-seq analysis ruled out the possibility of CITFA being enriched at the active BES through stable interaction with transcribing RNA pol I, this finding strongly indicated that BES activation involves a distinct promoter-centered mechanism dependent on CITFA (Günzl et al., 2015; Nguyen et al., 2014).

CITFA subunits have been found to be highly enriched in the ESB and the nucleolus, and not detectable in other parts of the nucleus where silent BESs are located (Nguyen et al., 2014; Nguyen et al., 2012). This localization pattern was found to remain intact even after *CITFA1*, a DNA-binding subunit of CITFA, was silenced. A subsequent immunoblot assessment of the relative abundances of other CITFA subunits, including CITFA2, CITFA6, and CITFA7 following *CITFA1* silencing showed negligible reductions in these other subunits, indicating that CITFA1 is not essential for CITFA complex assembly. Furthermore, indirect immunofluorescence of CITFA3 after *CITFA1* silencing showed CITFA3 within the presumed ESB. Given that CITFA1 binds directly to the BES promoter, yet CITFA3 remained within the ESB when CITFA1 was depleted, this suggests that CITFA subunits are sequestered within the ESB independent of transcription initiation at the promoter (Park et al., 2014). In light of this later finding, an

earlier study found that when trypanosomes were forced to simultaneously express two mutant BESs containing different antibiotic resistance genes, RNA fluorescence in situ hybridization resulted in an image of two fluorescent spots very close together and with diminished fluorescence (Chaves et al., 1999). This finding now seems to suggest that a key factor(s) involved in BES activation is localized within the ESB, and when two BES are forcibly co-expressed, competition arises for this limiting factor(s), perhaps the factor being CITFA.

Published data from our laboratory has shown that CITFA7 is not directly required for RNA pol I transcription *in vitro*, nor assembly and integrity of the CITFA complex (Nguyen et al., 2014; Brandenburg et al., 2007; Nguyen et al., 2012). Yet, unpublished results from a knockdown of *CITFA7* showed an increase in the expression of repressed BESs severalfold. Further preliminary data suggested that this function of CITFA7 may require LC8-mediated dimerization. Employing yeast-two-hybrid technology, our laboratory found evidence that CITFA2 and CITFA7 interact with the dimerizing protein LC8 and identified in both putative binding sites for LC8 similar to sites found in human and yeast genes (A. Günzl lab, unpublished). While an anti-CITFA7 antibody did not interfere with *VSG* transcription *in vitro*, the same assay with an antibody against CITFA2 abolished transcription, consistent with the understanding that dimerized CITFA2 enables the CITFA complex to recognize and bind to BES promoters, while CITFA7 does not affect promoter binding (Nguyen et al., 2012). Furthermore, LC8-mediated dimerization of CITFA2 proved essential for CITFA to bind to the BES promoter, indicating that only dimerized CITFA2 could stably interact with the promoter (Kirkham, et al., 2017). Although LC8-mediated dimerization of CITFA7 was not

rigorously verified, the hypothesis emerged that it could also be required for BES activation, however, not through direct binding to DNA but perhaps instead through an interaction with a scaffold protein, leading to sequestration of the complex within the ESB. Based on initial semi-quantitative PCR results, trypanosomes with an LC8 binding site mutation in CITFA7 exhibited elevated expression of *VSG* mRNA from a silent BES (T. Nguyen & A. Günzl, unpublished). Given this preliminary evidence, the goal of this study was to confirm CITFA7 dimerization *in vivo* and to determine whether a non-lethal mutation in *CITFA7* could disrupt the localization of the CITFA complex and therefore monoallelic expression. At present, researchers in this subject are primarily focused on telomeric silencing as the main mechanism underlying monoallelic *VSG* expression, overlooking the essential role of a dedicated, promoter-centered activation mechanism. This study provides further evidence that BES activation by CITFA7-mediated sequestration restricts *VSG* expression to a single allele.

METHODS

Cell Lines and Gene Silencing

For detection of CITFA7 by indirect immunofluorescence and for efficient pulldown in co-immunoprecipitation, a modified *T. brucei brucei* strain 427 containing a tagged copy of CITFA7 was used (Cross, 1975). More specifically, the composite ~20 kDa PTP tag had been C-terminally fused to CITFA7 and contained a Protein C epitope, a TEV protease cleavage site, and tandem Protein A domains (ProtA) by methods described previously (Schimanski, 2005). For *CITFA7* silencing, so-called single marker bloodstream trypanosomes, which express both T7 RNA pol and the tetracycline repressor, had been stably transfected with a linearized plasmid that contained 500 b.p. of

CITFA7 3’ untranslated region (UTR) in a sense-loop-antisense arrangement under the control of the TET-regulated T7 promoter (Wirtz et al., 1999). In a subsequent transfection, a linearized plasmid was targeted for integration into one of the two *CITFA7* alleles to add the PTP tag sequence to the 3’end of the coding region. The plasmid either left the coding region unchanged (WT cell line) or replaced 269’-GTGCAAGTTGAGTGG-283’ (encoding VQVEW) with GCAGCTGCAGCTGCC (encoding AAAAA). Wildtype (WT) and mutant cell lines expressing PTP-tagged CITFA7 were independently derived clonal cell lines from the same transfection (WT, Mut1, MutX). The mutant *CITFA7-PTP* is therefore identical in Mut1 and MutX clones. Specific knockdown of the endogenous copy of *CITFA7* was achieved by treatment with doxycycline (dox) (2 µg/mL) for 3 days (3d) or 10 days (10d) which induced the expression of a hairpin RNA that targeted the 3’ UTR of the endogenous *CITFA7* mRNA via an RNAi pathway, while not affecting *CITFA7-PTP* mRNA containing a different 3/ UTR (Nguyen et al., 2012).

Cell Culturing

Bloodstream form (BSF) trypanosome cultures were grown at 37°C, 5% P_{CO2} in HMI-9 medium (as described by Hirumi and Hirumi, 1989) supplemented with 10% fetal bovine serum. To maintain genetic modifications, cells were grown in the presence of 2 µg/mL blasticidin, 2.5 µg/mL G418, and 1 µg/mL phleomycin. 10 mL cultures were grown in a 50 mL flat-bottom, filter-cap flask to a density of $2x10^6$ cells/mL before splitting or harvesting.

Protein analysis

Co-IP assays were carried out to assess CITFA7 dimerization *in vivo* by an established method (Park et al., 2012). For protein extraction, WT, Mut1, and MutX cells were grown to a density of 2 x10^6 cells/mL in 350 mL HMI-9 medium per cell line. Cells were harvested by centrifugation at 2,700 g, 4°C, for 10 minutes, and resuspended and washed three times in ice-cold trypanosome wash solution (100 mM NaCl, 3 mM $MgCl_2$, and 20 mM of Tris-HCl at pH 7.5). Cells were then resuspended in extraction (E) buffer (150 mM sucrose, 20 mM L-glutamic acid, 20 mM HEPES-KOH at pH 7.7, and 3mM $MgCl_2$), vortexed and centrifuged at 2,700 g, 4°C, for 10 minutes, and resuspended in one packed-cell-volume of E buffer. 1 mM DTT, 10 µg/mL aprotinin, 10 µg/mL leupeptin, 20 µL phosphatase inhibitor cocktail 2 (Sigma), and ~200 µL of 800 µm pre-equilibrated Low Binding Silica Beads (OPS Diagnostics, LLC) were added to the suspended cells, which were then subjected to five rounds of a vortex-ice incubation-vortex treatment (shock frozen in liquid nitrogen, immediately hand-thawed). 20 µL of broken cell suspension was combined with SDS sample buffer (dH_2O, SDS, 200mM βME) up to 2X to be the input (Inp) sample, and the remaining extract was spun down at 25,000 g 4°C, and shock frozen in liquid nitrogen.

For precipitation of CITFA7-PTP, 25 µL of settled human IgG beads (GE Healthcare) were equilibrated and washed 3 times with 800 µL of TET150 buffer (containing 150 NaCl, 20 mM Tris-HCl at pH 8, 3 mM $MgCl_2$ and 0.1% Tween 20). Extract was mixed with 50 µL of TET150-suspended beads and chilled in a cool chamber at 3°C for 1 hour, inverting frequently to facilitate antibody-antigen binding between the ProtA domain of the PTP tag and the Fc domain of the IgG antibody. Extract was subsequently centrifuged at 3,000 g for 30 seconds, after which the supernatant (SN)

sample was collected, and SDS sample buffer was added up to 2X. Bound beads were washed six times with 800 µL of TET150 buffer (150 mM NaCl, 20 mM Tris–HCl, pH 8.0, 3 mM $MgCl_2$, 0.1% Tween 20), and CITFA7-PTP was eluted with 80 µL SDS sample buffer, into which CITFA7-PTP was directly released. 20 µL of samples (2X Inp, 2X SN, 3.5X P) were boiled for 5 minutes and loaded into a 6% stacking-15% separating SDS-polyacrylamide gel along with a 1X SDS buffer control sample and run at 80 minutes through the stacking gel and then at 120 V through the separating gel over 1 hour 20 minutes. Separated proteins were electroblotted onto a polyvinylidene difluoride membrane for 1 hour at 100 V at 20°C, which was stored in 1X TBS at 4°C overnight.

To visualize proteins by immunoblotting, the membrane was blocked in 5% milk blocking solution (2.5 g non-fat dry milk, 50 mL 1X TBST), and stripped with 0.5 M NaOH for 6 minutes before reprobing. Endogenous CITFA7 was detected with purified polyclonal α-rat anti-CITFA7 antibody (1:1000), the ProtA domain of CITFA7-PTP was detected directly with peroxidase anti-peroxidase (PAP) complex (1:5000), the negative control, CRK9, was detected with purified polyclonal α-rat anti-CRK9 antibody (1:1000), and the positive control, CITFA6, was detected with purified polyclonal α-rat anti-CITFA6 antibody (1:2500). Primary antibodies were detected with anti-rat IgG POD secondary antibody (1:5000). BM Chemiluminescence Western Blotting Kit (Sigma Aldrich) was used to detect the peroxidase-labeled secondary antibody and the chemiluminescent substrate luminol for imaging.

RNA analysis

To qualitatively assess relative *VSG* mRNA abundance produced from active and silent BESs, RNA was isolated from trypanosomes growing in exponential phase before

and after 3d and 10d dox inductions. 10 mL cultures were centrifuged at 2,000 g, 4°C, for 7 minutes, and cells were resuspended in 700 µL of TRIzol™ Reagent (Invitrogen), vortexed vigorously for 10 seconds and stored overnight at -80°C. Samples were thawed for 5 minutes at room temperature (RT). 200 µL of chloroform was added, the samples were again vortexed and incubated for 2 minutes at RT, and centrifuged at 12,000 g, 4°C, for 15 minutes. The aqueous phase was transferred to a new tube and 500 µL of isopropanol was added and centrifuged at 12,000 g, 4°C, for 10 minutes. The pellet was washed with 900 µL of 70% ethanol, and centrifuged at 12,000 g, 4°C, for 5 minutes. After drying for approximately 10 minutes, the pellet was resuspended in 30 µL of dH_2O.

To avoid PCR amplification of genomic DNA contaminations, a DNA digest was performed on isolated RNA prior to cDNA synthesis following the standard protocol (Sigma-Aldrich). RNA was reversed transcribed by SuperScript IV Reverse Transcriptase with Oligo-dT or random hexamer primers to make cDNA. To assess the quality of the RNA and to see if there were discrepancies in the relative concentrations between RNA samples, RNA gel electrophoresis was performed using a Reliant™ Gel System (Lonza). 2 µg of RNA was added to 8 µL of Ambion Glyoxal Sample Loading Dye and incubated at 50°C for 30 minutes. The samples were spun briefly and placed on ice before loading. The gel was run in 1X MOPS at 70V (10X MOPS: 200 mM MOPS, 50 mM Sodium acetate, 10 mM EDTA at pH 7.0) through a precast 1.25% SeaKem Gold agarose RNA gel, incubated on a shaker in a mixture of 100 mL DMPC with 1 µg/mL ethidium bromide in a plastic container, and covered in polyvinyl-chloride wrap for 30 minutes. The gel then was washed 3 times in milli-Q water and imaged in a transilluminator.

For standard RT-PCR of *CITFA7*, a 4% agarose gel was successful at resolving the small, amplified cDNA product. A PstI restriction enzyme digest was performed on amplified cDNA using NEBuffer™ r3.1 according to standard protocol (NEB).

For semi-RT-qPCR, T_m was calculated for each primer to optimize the annealing temperatures, and primer length taken into account when selecting extension times. The linear range of product increase was found to be 29 cycles for cDNA of silent *VSG*s.

For RT-qPCR, melting curves and standard curves were analyzed before calculating relative mRNA abundance from cycle threshold (C_t) values using the Delta-Delta-C_t Method (Livak and Schmittgen, 2001). SYBR Green I dye was used in RT-qPCR reactions to obtain C_t values, and these values for *VSG* sequences were normalized against those C_t values of *a-tubulin* which is expressed at high levels like the active *VSG* and is transcribed by RNA pol II.

DNAs

DNA oligonucleotides specific to the coding sequence of *CITFA7* used for RT-PCR have been previously described (Park et. al., 2014). Oligonucleotides used for semi-RT-qPCR and RT-qPCR included the following: 5'-AGTTTCTTAGCAACTGACTGC-3'and 5'-CACCTACGTTTTCGAGCTGT-3' for the single silent *VSG* mRNA (dubbed VSGbR2), 5'-AAATGGGAGGGTRAAACTTGCAAAG-3' and 5'-CAAGGCCACAAATGCAGCAG-3' for consensus oligonucleotides for four BES-localized silent *VSG*s mRNA, 5'-AGCTTTTTGGCAACCTCTTTGCCAGG-3' and 5'-CCTTATCGACCGTTTTGTCCGCAATGGT-3' for the active *VSG* mRNA (dubbed VSG221), 5'-GTGCATTGAACGTGGATCTG-3' and 5'-CGGATGGTGCTCGTTACGTG-3' for *a-tubulin*, and 5'-

TCATCAAACTGTGCCGATTAC-3'and 5'-CTATTGAAGCAATATCGG-3' for the 18S rRNA coding region.

Indirect immunofluorescence

Localization of CITFA7 within the nucleus was examined through an established indirect immunofluorescence method (Nguyen et al., 2012). Coverslips were washed in acetone and silanized with 2% aminopropyltriethoxysilane for 1 minute on a shaker. BSF cells were grown to exponential phase and harvested at a density of 2 x 10^6 cells/mL. 10 mL of each cell culture were spun at 800 g for five minutes at RT, washed in 1X PBS, and fixed with a final concentration of 2% formaldehyde for 10 minutes at RT. Fixation was terminated by adding glycine to a final concentration of 0.125 M and incubating for 5 minutes at RT. Cells were then washed three times and resuspended in 1X PBS to a density of 3.5 x 10^6 cells/mL, and 250 µL were spotted on silanized coverslips and left to settle for 30 minutes. Unadhered cells were gently washed away twice with 1X PBS for 5 minutes each and fixed cells were stored at 4°C overnight. Fixe cells were then placed in a circular 6-well plate and spun at 800 g for 2 minutes at RT, and rehydrated in 1X PBS for 10 minutes, permeabilized with 0.1% NP40 in 1X PBS for exactly 5 minutes, washed twice in 1 mL 1X PBS, and washed once with 500 µL PBG (1X phosphate-buffered saline (PBS), 0.2% fish gelatin and 0.5% BSA) for 5 minutes each. The cells were blocked twice in 250 µL PBG for 10 minutes and incubated with primary α-rabbit anti-ProtA antibody (1:40,000 dilution in 1X PBG) for 1.5 hours at RT. After washing six times on shaker in 1X PBS containing 0.05% Tween 20, slips were incubated with secondary Alexa Fluor™ 594 antibody (Thermo Fisher; 1:400 dilution in 1X PBG) along with DAPI (1:200) in 1X PBG for 45 minutes with a foil covering to prevent fluorophore

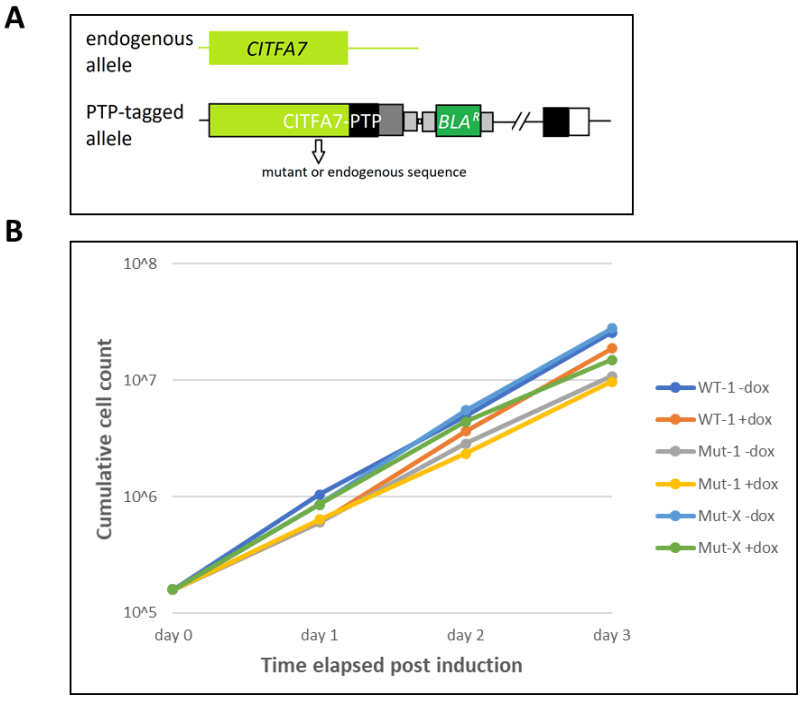
bleaching. The slips were washed six times in 1X PBS on a shaker to remove unbound antibodies and mounted onto a slide with 7 μL Vectashield per slip. Cells were visualized and imaged with a Zeiss 780 Upright Confocal microscope equipped with a 63X objective and a 561 nm Diode laser. Exposure time was adjusted to optimize the detection of nuclear fluorescent signals. Images were analyzed in Fiji-ImageJ and brightness and contrast were adjusted for better visualization.

RESULTS

Mutation in the putative LC8 binding site of CITFA7 is nonlethal. Previously, a sedimentation analysis strongly indicated that subunit CITFA2 dimerizes through an interaction with LC8 (Kirkham et al., 2016). Results of a subsequent yeast-two-hybrid analysis suggested that CITFA7, in addition to CITFA2, dimerizes *in vitro*. Given CITFA2's interaction with LC8, it was thought that CITFA7 might also dimerize via LC8. Using an LC8 binding partner prediction system (Rapali et al., 2011), the 89'-GVQVEW-94' motif within CITFA7 was identified as the potential LC8 binding site (J. Kirkham, unpublished).

To elucidate whether CITFA7 does, in fact, dimerize *in vivo*, 427 strain BSF trypanosomes were modified to express a PTP-tagged version of CITFA7 from one allele while leaving the remaining *CITFA7* allele unmodified. A mutation was then introduced into clones of the CITFA7-PTP-expressing cell lines such that five amino acids in the putative LC8 binding site were replaced with five alanines. Because all cells derived from a parent line expressing a T7 RNA pol and bacterial tetracycline (TET) repressor, silencing of the endogenous copy of *CITFA7* could be reversibly induced in the presence of doxycycline, a more stable derivative of tetracycline (Figure 2A).

CITFA7 has been found to be essential for cell viability—when *CITFA7* was silenced, transcription of RNA pol I-transcribed genes plummeted and cells died (Nguyen et al., 2012). However, silencing of the endogenous *CITFA7* in mutant clones expressing *CITFA7-PTP* did not lead to a decline in cell growth over 3d and 10d inductions, confirming that the CITFA7 mutation is nonlethal (Figure 2B; growth curve for 10d not shown). To check that both wild type and mutant cells were, in fact, expressing *CITFA7-PTP*, standard PCR was performed. Based on the fact that only the mutant *CITFA7-PTP* sequence contains a PstI restriction site, this sequence difference was leveraged to confirm the expression of the mutant gene in mutant cells after 10d induction. Gel electrophoresis results show fragmented PCR-amplified DNA from cDNA of mutant *CITFA7-PTP* in the mutant lane following PstI digest, whereas only unfragmented cDNA appear in the WT lane, as expected (Figure 2C).



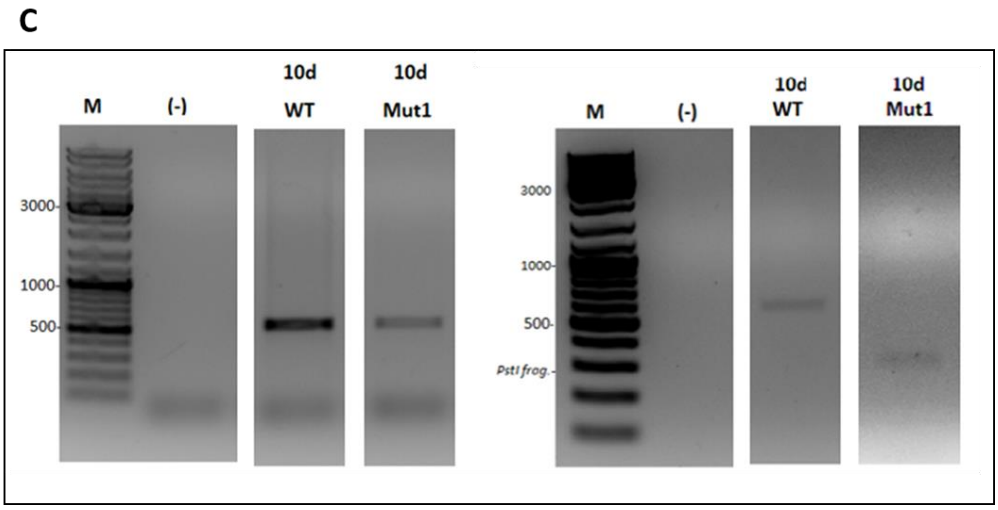


Figure 2 Verification of non-lethality and expression of mutant *CITFA7*. (A) Schematic of the CITFA7 locus in smC7PTP-WT or -Mut1 or -MutX cells expressing *CITFA7-PTP* from one allele (not to scale). **(B)** Growth curve of representative WT, Mut1 and MutX cells over 3d induction, silencing endogenous *CITFA7* (counts performed by A. Günzl). **(C)** Before (left) and after (right) restriction enzyme digest with PstI performed on *CITFA7-PTP* PCR products.

CITFA7 likely does not dimerize *in vivo*. To test for dimerization of CITFA7 *in vivo*, co-IP was carried out against the PTP tag in CITFA7-PTP with CITFA7-PTP-WT protein extracts, and results were assessed with an immunoblot. If dimerization occurred, it would be expected that pulldown of CITFA7-PTP would co-precipitate untagged CITFA7, leaving little CITFA7 in the supernatant and the majority in the pellet sample. As expected, the negative control, cyclin-dependent kinase CRK9, was detected in the input and supernatant sample lanes, but not in the pellet sample lane. Conversely, the positive control, CITFA6, was detected in the input and pellet lanes, but only faintly detectable in the supernatant, indicating successful co-precipitation with CITFA7-PTP (Figure 3, top panel). Likewise, the PTP tag of CITFA7-PTP was visualized clearly in the input and pellet lanes and was depleted from the supernatant sample. When the immunoblot was reprobed with anti-CITFA7 antibody, the protein was successfully

detected in the input lane, as expected. However, counter to the hypothesis, CITFA7 did not appear to co-precipitate with CITFA7-PTP, as it was both clearly detected in the supernatant lane and virtually absent in pellet lane, opposite to expectations (Figure 3). These results suggest that CITFA7 may not dimerize *in vivo*. However, there is also the possibility that the CITFA7-LC8 interaction did not withstand co-IP, given that different LC8 binding sites have different affinities for LC8 (Rapali et al., 2011; Jespersen et al., 2019).

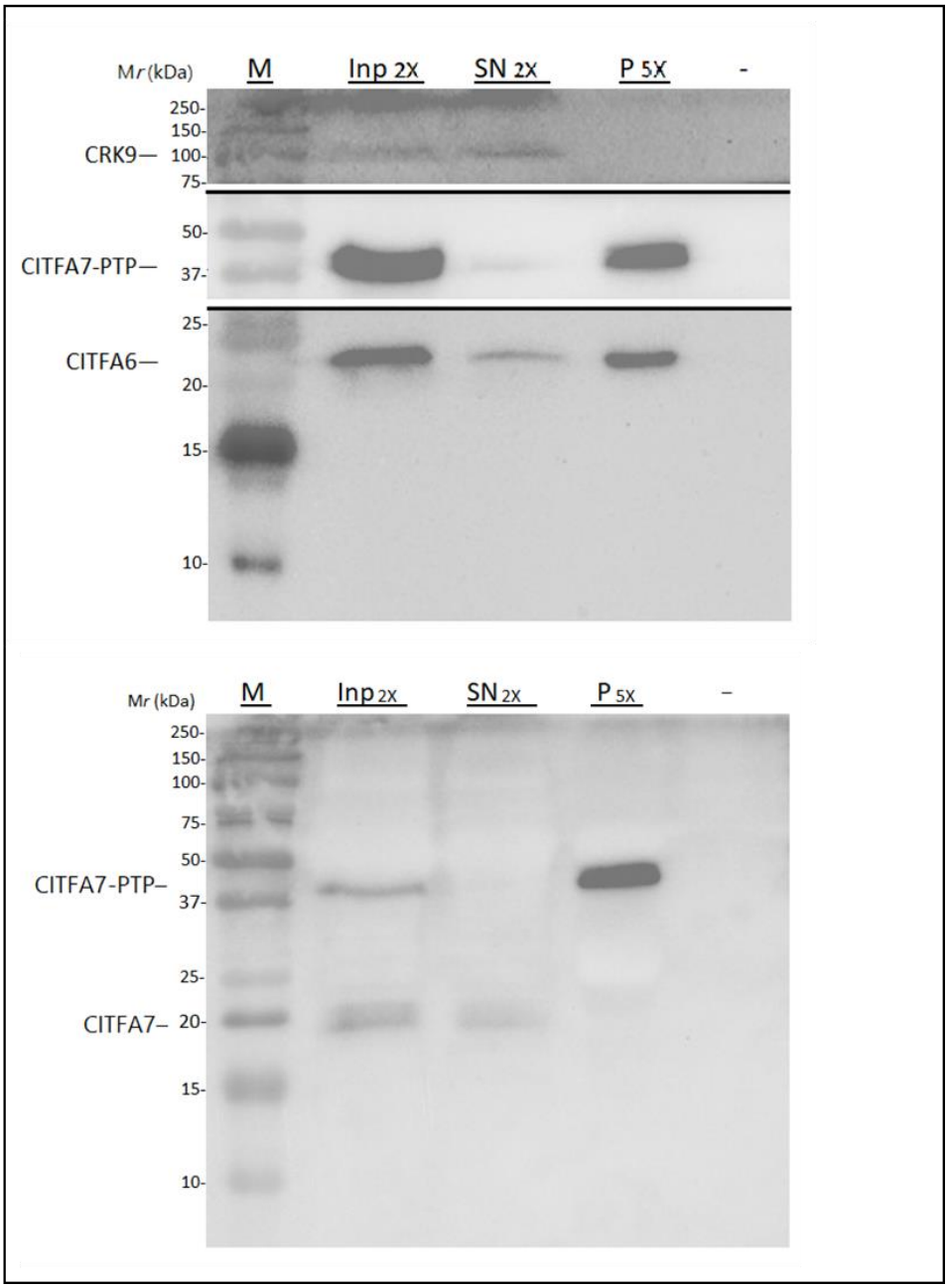
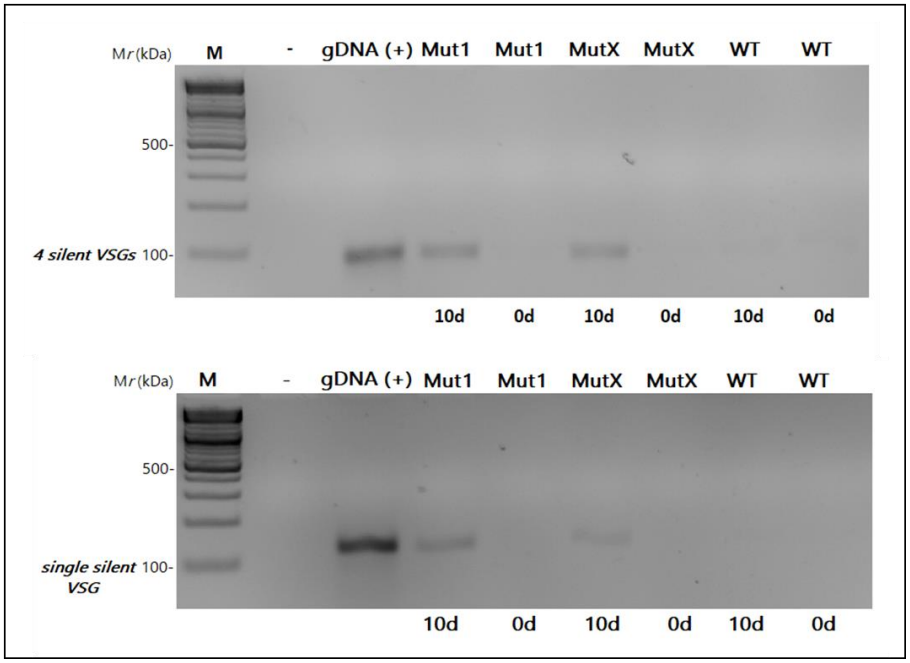


Figure 3 Immunoblot testing for CITFA7 dimerization. CITFA7-PTP was precipitated from extract with IgG beads, and proteins in the input (Inp), supernatant (SN), and pellet (P) samples were separated by SDS-PAGE. PVDF membrane blotted with protein from WT CITFA7-PTP-expressing cells was probed with the following primary antibodies: purified polyclonal rat anti-CRK9 antibody to detect CRK9 (negative control), and purified, polyclonal rat anti-CITFA6 antibody to detect CITFA6 (positive control) [controls in top panel]; purified, polyclonal rat anti-CITFA7 antibody to detect endogenous CITFA7, and PAP reagent to detect CITFA7-PTP [bottom panel]. Anti-rat IgG POD secondary antibody was used to detect primary antibodies.

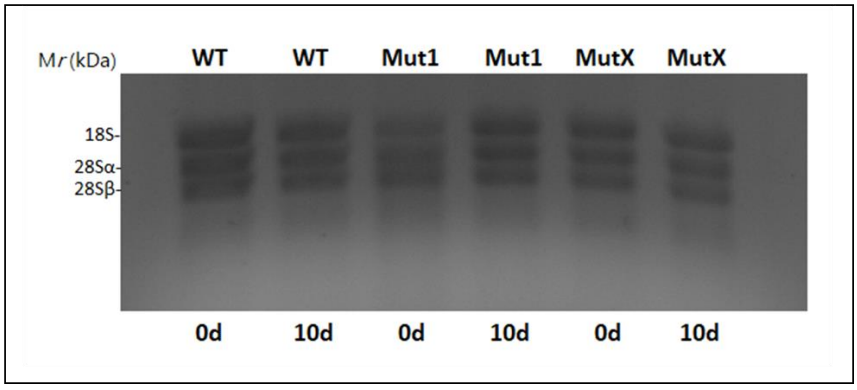
Knockdown of endogenous CITFA7 leads to increased *VSG* expression from silent BESs. Although dimerization may not be the mechanism by which CITFA7 contributes to monoallelic *VSG* expression from a single active BES, results from a preliminary semi-RT-qPCR on silent *VSG*s suggested that the putative LC8 binding site mutation introduced into *CITFA7-PTP* disrupts monoallelic expression. In particular, after silencing endogenous *CITFA7* with doxycycline (leaving only PTP-tagged *CITFA7*, either WT or mutant, expressed), gel electrophoresis images showed elevated expression of silent *VSG*s indicated by strong bands in mutant cell lanes but not WT lanes (T. Nguyen, data not shown). To confirm this initial result both semi-quantitatively and quantitatively, Mut1, MutX, and WT clones in exponential growth were subjected to a 10d (and in a subsequent trial, 3d) dox treatment. Total RNA was isolated from cells at the 0d and 10d (or 3d) mark and reverse transcribed. For semi-RT-qPCR, both a single silent *VSG* and four silent *VSG*s (primers were designed taking advantage of regions of sequence consensus) were amplified. Bands were not visible in both WT and mutant cell lanes before *CITFA7* silencing (0d), despite the fact that half of the CITFA7 proteins available to mutant cells before silencing are abnormal. However, strong bands in both Mut1 and MutX clone lanes, indicative of silent *VSG* expression, were noted after 10d silencing, while no bands were noted in the WT lanes (Figure 4A). Before proceeding with RT-qPCR analysis of *VSG* expression, the quality of the isolated RNA was assessed by direct detection of rRNA (Lonza Reliant^{TM} Gel System), and the relative concentrations of cDNA was assessed by RT-qPCR amplifying 18S ribosomal cDNA. RNA appeared intact, with no streaking or smearing noted in any lanes (Figure 4B). Meanwhile, abundance of 18S rRNA was consistent between WT Mut1 samples (fold

range: 1.00-1.19) before and after 10d induction (Figure 4C). RT-qPCR results show that mRNA abundance of the four normally silent *VSG*s increased between 10.8-fold in Mut1 cells, and the expression of a single silent *VSG* increased 13.8-fold, while *VSG* expression did not change in WT cells after 10d. Negligible increases in expression level were found for the active *VSG* in WT (1.00 to 1.05) and Mut1 (1.00 to 1.322) cells, within the range of expected experimental variability (Figure 5A-C).

A



B



C

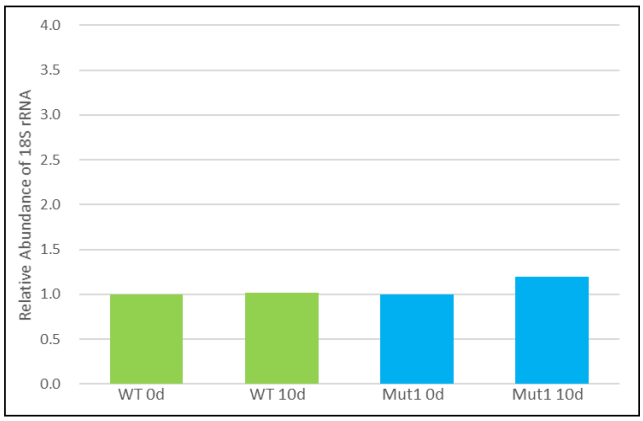


Figure 4 Verification of RNA quality and relative concentrations of cDNA samples. (**A**) Total RNA prepared from 0d and 10d-induced cells was reverse transcribed with Oligo-dT and analyzed by semi-RT-qPCR with oligonucleotides for four Silent *VSG*s, a single silent *VSG*, and the active *VSG*. (**B**) Total RNA separated by Reliant™ precast 1.25% agarose RNA gel and rRNA stained with ethidium bromide. **(C)** RT-qPCR of cDNA (rRNA reverse transcribed with random hexamer) from 0d an 10d-induced cells.

To determine whether the semi-RT and RT-qPCR results could be reproduced with cells induced over a shorter duration, a 3d induction was carried out. Semi-RT-qPCR with 3d cDNA showed qualitatively indistinguishable results from that of the 10d experiment; Mut1 and MutX lanes showed strong bands after 3d of endogenous *CITFA7* silencing (data not shown). Consistent with the 3d semi-RT-qPCR, RT-qPCR results showed that mRNA abundance of four silent *VSG*s increased in Mut1 and MutX cells by 6.9-fold and 5.3-fold, respectively, and decreased slightly (1.00 to 0.59) in WT cells (Figure 4B). At the same time, expression of the active *VSG* decreased negligibly (WT: 1.00 to 0.76; Mut1: 1.00 to 0.99; MutX: 1.00 to 0.86) in all three cell lines (Figure 4A), and expression of the single silent *VSG* increased by 6.2-fold in Mut1 and 6.3-fold in MutX, respectively, while decreasing minimally (1.00 to 0.71) in WT cells (Figure 4C). Thus, the quantitative results align with initial semi-quantitative findings that suggested

that expression mutant *CITFA7* in the absence of the WT form disrupts monoallelic *VSG* expression.

Figure 5 RT-qPCR analysis of mRNA abundances in 0d and 10d-induced WT and Mut1 trypanosomes, and in 0d and 3d-induced WT, Mut1, and MutX cells (*n=2*). **(A)** relative mRNA abundance of active *VSG* **(B)** four silent *VSG*s **(C)** and a single silent *VSG*. mRNA was reverse transcribed with Oligo-dT.*VSG* mRNA relative abundance was normalized against that of the housekeeping gene, *α-tubulin*.

Visualization of CITFA7-PTP within a subnuclear compartment. Given that RT-qPCR results showed an increase in expression of *VSG* from silent BESs following endogenous *CITFA7* silencing, it was expected that mutant CITFA7 might be found in multiple subnuclear locations within cells in which unmodified *CITFA7* upon dox induction, representing newly activated BESs. RNA pol I has previously been found to appear brightly around the nucleolar periphery. The associated protein, RPB7-PTP has also been detected around the perimeter but not within the nucleolus (Park et al., 2011). Thus, it was expected that CITFA7, a subunit of the RNA pol I transcription-initiation complex CITFA, might also be detected as a ring, outlining the edge of the nucleolus.

Indirect immunofluorescence was carried out to compare CITFA7 signals 0d and 3d post-induction. A primary polyclonal rabbit anti-ProtA antibody was used to detect the Protein A domain of the PTP tag on CITFA7, and the constant region of this rabbit antibody was then detected with Alexa Fluor™ 594 secondary antibody. Contrary to expectations, preliminary review of images comparing the number of distinct extranucleolar spots between 0d and 3d induced Mut1 cells did not indicate an increase in CITFA7-PTP foci after 3d, however, CITFA7-PTP signal appeared noticeably larger and/or more diffuse in some 3d dox-induced cells, as shown in representative images (Figure 6A). In multiple cases, CITFA7-PTP in 0d Mut1 cells was detected both encircling the characteristically dark nucleolus and in a single spot outside of this region (Figure 6A). While it is possible this extranucleolar signal could represent the ESB, this pattern was not observed in all Mut1 0d cells, preventing definitive identification of the ESB. As found in the representative z-stack sequences of a 0d Mut1 cell, CITFA7-PTP clearly encircled the nucleolus and appeared to be localized in an extranucleolar spot in

the top left of the cell (Figure 6C). CITFA7-PTP in 3d induced Mut1 cells was also detected around the perimeter of the nucleolus, though a distinct extranucleolar spot was not always discernible (Figure 6D). As can be seen in the 3d Mut1 z-stack sequence, CITFA7-PTP was sometimes detected as multiple distinct spots, as predicted, however it was unclear whether these spots corresponded to the peripheral nucleolar signal or to an extranucleolar compartment (Figure 6E). Indirect fluorescence was performed on MutX and WT cells in addition to Mut1 cells, however, there were unforeseen difficulties in staining and imaging these two clones, negating analysis (data not shown).

Due to the very small size of the trypanosome nucleolus and the ESB, more reliable detection of these subnuclear bodies has been achieved by direct fluorescence microscopy (Nguyen et al., 2014). Thus, to visualize CITFA in our WT and Mut1 cell lines during endogenous CITFA7 silencing by direct fluorescence microscopy, a pre-established CITFA2-mCherry plasmid was transfected into these cells. Since CITFA2 is an essential subunit of the CITFA complex and binds directly to BES and rRNA promoters and given that dox-induced endogenous CITFA7 knockdown seems to increase expression of silent *VSG*s, it is expected that CITFA2 will show as multiple spots within the nucleus after *CITFA7* silencing in mutant CITFA7-PTP cells. In a future study, these cells will be subjected to direct fluorescence microscopy for quantification of fluorescent signals from the ESB and any other extranucleolar regions to which CITFA2 may be directed.

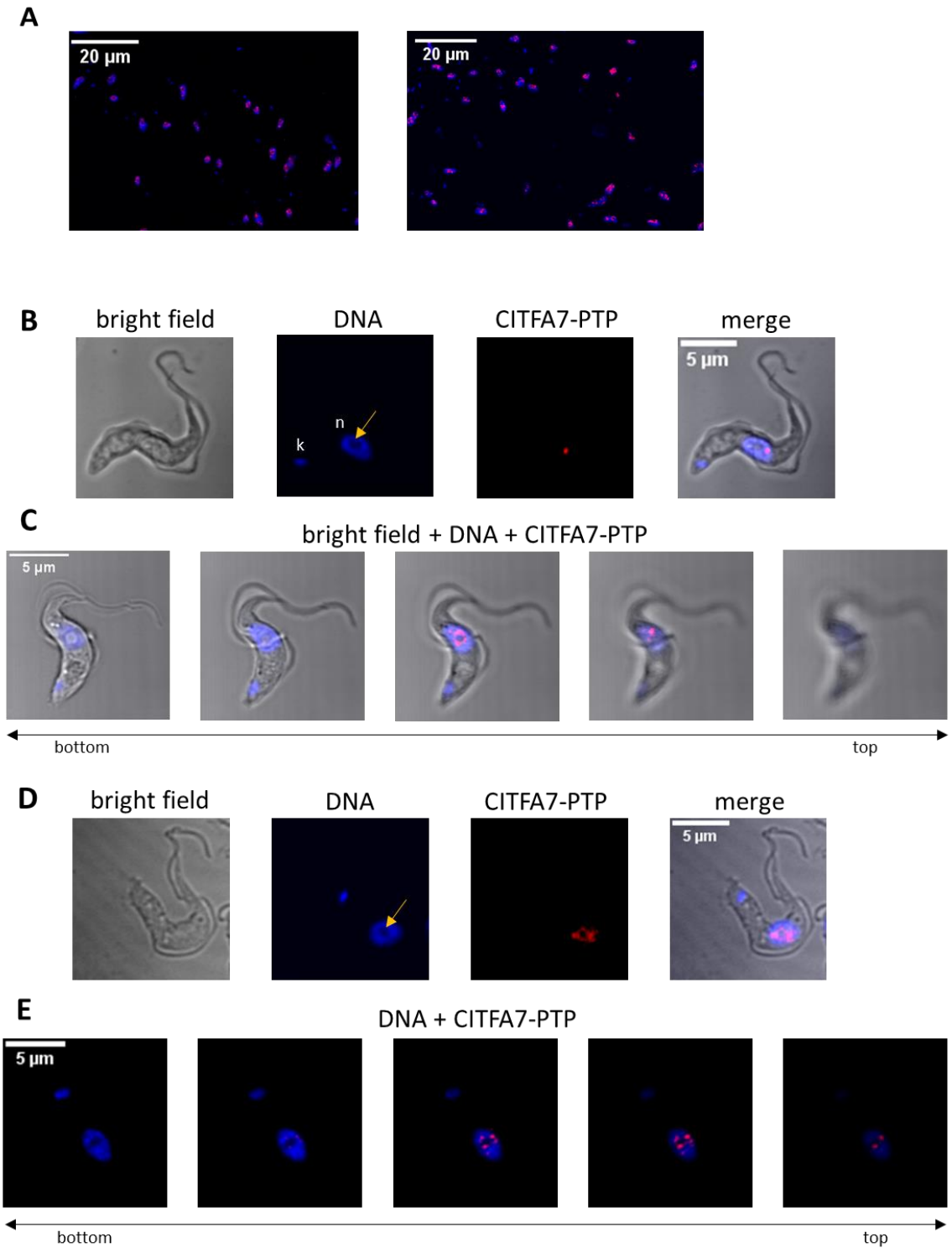


Figure 6 Indirect immunofluorescence microscopy of CITFA7-PTP before (A-B) and after (C-D) 3-day endogenous CITFA7 silencing. Mut1 CITFA7-PTP strain 427 cells were fixed, and DNA was stained with DAPI (n, nucleus; k, kinetoplast). The nucleolus is identified with a yellow arrow. Examples were chosen in which a single or multiple subnuclear spots were visible, and brightness and contrast were enhanced. CITFA7-PTP was detected with a polyclonal anti-ProtA primary antibody and Alexa Fluor™

594 secondary antibody (red). Images were captured with a Zeiss Confocal LSM 780 microscope and processed with Fiji-ImageJ. **(A)** Representative 0d and 3d-induced Mut1 cells. **(B)** Single 0d Mut1 cell. **(C)** Z-stack of 0d Mut1 cell; voxel depth: 0.587 μm; slices (z)=8. **(D)** Single 3d Mut1 cell. **(E)** Z-stack of 3d Mut 1 cell; voxel depth: 0.322 μm; z=10.

DISCUSSION

Monoallelic *VSG* expression seems to require functional CITFA7. Transcription activation at BES promoters has long been proposed as a key regulatory step of monoallelic expression in trypanosomes. Though studies have identified RAP1 and other telomere-associated proteins as critical for maintaining low transcription rates and early attenuation of transcription from silent BES genes, few studies have elucidated the mechanism by which RNA pol I transcription activation at the active BES occurs (Yang et al., 2009; Dreesen et al., 2007). A key distinction should be drawn between BES activation and sustained BES transcription, as all current evidence points towards distinct mechanisms regulating these steps. Recently, with the characterization of the VSG exclusion protein, VEX1, whose sequestration was found to monoallelically restrict *VSG* expression by preventing transcription of silent *VSG* genes, this distinction had become even clearer. VEX1 assembles with a nonsense-mediated-decay helicase ortholog, VEX2, in an RNA pol I-dependent fashion (Glover et al., 2016). While VEX1 has been found to localize within a transcription compartment around the SL RNA gene array with RNA pol II, VEX2 has been found to localize within a transcription compartment with RNA pol I and the active *VSG* gene. With super-resolution microscopy and proximity ligation, these two compartments were found to exist in close spatial proximity, and this closeness has been hypothesized to supply a high enough concentration of *trans*-splicing machinery such that *VSG* transcripts can be efficiently matured (Faria, et al., 2021). However, the role of VEX proteins in BES transcription activation is predicated on the presence of the

CITFA complex within the ESB; given that VEX1 is RNA pol-dependent and that the CITFA complex is RNA pol-I independent, ESB formation depends on RNA pol I activity, and CITFA is required for this activity to be initiated. Thus, it seems that monoallelic BES transcription is primarily determined by CITFA's activity during transcription initiation.

In this study, we built upon a preliminary LC8 binding partner prediction and yeast two-hybrid finding that identified a potential LC8 binding motif in CITFA7, part of the CITFA complex within the nucleus, as well as semi-RT-qPCR findings that showed that depletion of endogenous *CITFA7* (with only a mutant copy of *CITFA7* expressed) leads to an increase in *VSG* mRNA abundance from silent BESs. Our findings provide further quantitative evidence that CITFA7 plays an important role in the proper functioning of the CITFA complex, since when mutant CITFA7 cells were subject to ten-day and later to three-day *CITFA7*-silencing inductions, *VSG* mRNA abundance from silent BES increased 5.3-13.8-fold while active *VSG* mRNA expression remained stable in mutant CITFA7-PTP cells. One question this finding raises is how silencing of unmodified *CITFA7* compares to knockdowns of other genes associated with monoallelic *VSG* expression. From an RNAi screen for loss-of-exclusion of *VSG* transcription, VEX1 was identified, and a subsequent knockdown of VEX1 led to an overall expression increase of 18 previously silent *VSG*s by > 26-fold (Glover et al., 2016), a result which could perhaps, in part, be attributed to the spreading out of CITFA. In contrast, when the high-mobility group box protein, TDP1, which is considered to facilitate RNA pol I transcription, was depleted, growth arrest occurred rapidly, and chromatin repression was observed at the actively-transcribing RNA pol I locus in the ESB. Conversely,

overexpression of *TDPI* was shown to decondense chromatin of silent BES, likely disrupting monoallelic *VSG* expression by facilitating readthrough transcription (Aresto-Branco et al., 2019). Compared to VEX1 and TDP1, which influence RNA pol I activity at BES promoters, knockdown of the telomere-associated protein RAP1 has also been shown to derepress silent BESs. However, in this case, the expression of formerly silent *VSG*s seems to be the result of impairing telomeric silencing rather than directly altering transcription rates. Given VEX1's role in establishing the ESB and CITFA's role in enabling RNA pol I transcription, the increase in silent *VSG* mRNA abundance that we observed after silencing endogenous *CITFA7* could be the result of disrupting its sequestration, causing CITFA to spread out from the ESB and alter the rate of transcription initiation at silent BESs. Thus, while telomeric silencing clearly plays a key role in maintaining monoallelic transcription of the active *VSG*, the establishment of the active BES appears to be distinctly regulated by RNA pol I transcription-associated proteins, including CITFA7.

One important consideration in light of our results is that we only examined abundance of *VSG* mRNA and not pre-mRNA of expression-site-associated genes. BES genes are transcribed polycistronically and then *trans*-spliced, however, the fact that only the active BES has been found to be linked to the SL RNA gene locus would suggest that *trans*-splicing of silent BES genes, including *VSG*, is not as efficient as it is at the active BES, causing less efficient production of mature *VSG* mRNA from these repressed loci (Günzl, 2010; Faria et al., 2021). However, in contrast to VEX1 silencing, when *CITFA7* is mutated, it is possible that *ESAG*s of silent BESs are being transcribed at even higher levels than downstream *VSG* genes in at nonlinear rates, particularly if it is the strict

localization of CITFA within the ESB that is disrupted by the unmodified *CITFA7* knockdown. In this model, RNA pol I has greater access to silent BES promoters with CITFA now available at these locations, and thus more transcription can occur at the beginning of the BES. However, the persistent heterochromatin structure of previously silent BESs may continue to reduce transcription of downstream regions including *VSG* genes, which would explain the still-lower mRNA abundances of silent *VSG*s compared to the active *VSG*, as shown by our data (Pandy et al., 2013). Given that *ESAG*s of BES are highly conserved (over 90%), it is difficult to design primers that can distinguish between active and silent BES *ESAGs* without genetically modifying the *ESAGs* to include unique tags (Hertz-Fowler et al., 2008). In future studies, we would like to investigate abundance of BES promoter proximal *ESAG* transcripts from silent and active BES, in addition to *VSG* mRNA abundance to help clarify the mechanism by which the *CITFA7* mutation alters transcription of *VSG*s.

CITFA7 does not likely contribute to monoallelic expression through dimerization. Our hypothesis that CITFA7 dimerizes *in vivo* derived from clear evidence that CITFA2 dimerizes with LC8 and that such dimerization allows CITFA2 to bind to BES promoters (Kirkham et al., 2016). Since a preliminary finding predicted a putative LC8 binding site In *CITFA7*, we postulated that this motif was important for CITFA7's role in RNA pol I-mediated transcription *in vivo*. A previous study found that CITFA7 consistently copurifies with CITFA2 and CITFA6 in sub-stoichiometric amounts, indicating that it is not essential for the formation or the stability of the complex. This was especially true when co-purification was achieved using PTP-tagged CITFA subunits, which would suggest that the PTP tag destabilized CITFA7's ability to bind to these subunits.

(Brandenburg et al., 2007). In a separate assay, depletion of CITFA7 *in vitro* led to the co-loss of all other subunits except LC8, supporting the alternative hypothesis that CITFA7 serves a stabilizing function (A. Günzl lab, unpublished). However, fluorescence microscopy has shown that when *CITFA1* is silenced, CITFA3 remains localized within the ESB. Thus, even without CITFA binding to BES promoter DNA, CITFA3 remained sequestered, perhaps through an interaction mediated by CITFA7 (Park et al., 2014). Considering our immunoblot results, if CITFA7 interacted with the CITFA complex as a dimer, it would be expected that both the tagged and untagged versions of the protein in the precipitate would be seen in a 1:3 ratio. Yet, this ratio was not observed from co-IP, and instead, our results strongly suggest that CITFA7, in contrast to preliminary results, does not dimerize *in vivo*. In light of this, it is still possible that CITFA7 dimerizes *in vivo* and the method of co-IP that was used in this study was simply too harsh for the lower-affinity of CITFA7 to CITFA7-PTP and to other subunits of the CITFA complex. In fact, a bioinformatics analysis suggested that the LC8 binding motif in *CITFA2* has a higher affinity than the one predicted in *CITFA7* (A. Günzl lab, unpublished). Thus, it appears that while CITFA7 is essential for RNA pol I transcription of rRNA genes and the active *VSG in vivo*, a nonlethal putative LC8 binding site specifically interferes with CITFA's specificity for the active BES.

The question for which we do not yet have a hypothesis is why the transcription of silent BESs only increases when the endogenous copy of *CITFA7* is completed ablated through an RNAi knockdown, even though before silencing, half of all the *CITFA7* genes being expressed are the mutant form. In light of this observation, the unmodified CITFA7 seems to have a rescue effect in CITFA7-PTP mutant cells. Because it is unlikely that

mutant *CITFA7* and endogenous *CITFA7* are expressed at different levels in non-induced cells, and furthermore, because the dox-induced knockdown negligibly affected the mRNA abundance of silent *VSG*s in WT cells, a non-specific effect due to antibiotic treatment can be excluded. There are multiple possibilities that could explain the rescue effect that unmodified CITFA7 seems to have on monoallelic *VSG* expression. For one, there could be low affinity dimerization required for ESB localization, and the mutation of CITFA7 could weaken it such that dimerization is unstable between two mutants, but interaction still occurs between a mutant and WT CITFA7, and between WT and WT CITFA7. Moreover, the unmodified CITFA7 could outcompete the mutant CITFA7 in its association with the CITFA complex, making functional CITFA complexes still available in sufficient quantities for the characteristically high active *VSG* transcription. It is also possible that unmodified CITFA7 might stabilize an unidentified scaffold molecule, whether protein or RNA, which could lead to CITFA's enrichment in the ESB. If this function is impaired in mutant CITFA7, the scaffold molecule might be degraded when mutant CITFA7 is the only form available, causing the loss of CITFA accumulation in the ESB and the spreading out of CITFA complexes to silent BESs. Thus, although our data indicate that CITFA7 does not dimerize, the substitution of five amino acids within a conserved *CITFA7* domain appears to alter CITFA7 such that the CITFA7 complex loses its specificity for the active *VSG* gene. This finding strongly indicates that CITFA7 has a role in restricting *VSG* expression to a single allele at the level of transcription initiation.

At a spatial level, fluorescent microscopy has shown that the site of active *VSG* transcription—the bloodstream expression site—is strictly localized within the expression site body (Navarro & Gull, 2001). CITFA subunits were previously found to be highly

enriched in the ESB and the nucleolus, and not detectable in other parts of the nucleus where silent BESs are located (Nguyen et al., 2014; Nguyen et al., 2012). This localization pattern remained intact even after *CITFA1* (an essential direct DNA-binding subunit of CITFA) was transiently knocked down, which provided strong evidence that CITFA is sequestered in these compartments independent from transcription initiation (Park et al., 2014). In dox-induced cells expressing multiple *VSG*s at high levels, we expected to see mutant CITFA7 fluorescence signals in multiple spots throughout the nucleus and not exclusively within the ESB of the active *VSG*. This prediction was based on the hypothesis that the non-lethal mutations in *CITFA7* would impair CITFA7's function in sequestering or otherwise restricting the location of the CITFA complex. However, our preliminary fluorescence microscopy images are not, at this point, clear enough for us to identify the ESB and possible extranucleolar spots of silent BES compartments. While a single extranucleolar spot was detected in some Mut1 0d dox-treated cells, nonspecific antibody binding in other regions of some cells might explain cases in which spots were seen throughout and sometimes outside of the nucleus. Our methods of indirect fluorescence will therefore be refined and optimized before quantitative analysis can be undertaken.

CONCLUSION

In summary, our co-IP results provide evidence that CITFA7 does not dimerize *in vivo*, while our semi-RT-qPCR and RT-qPCR results confirm preliminary findings that a non-lethal mutation in CITFA7 yields upregulated production of *VSG* mRNA from silent BESs when only mutant CITFA7 is available. To our knowledge, this is the first time expression levels of silent *VSG*s have been changed through disruption of a subunit in a

promoter-binding complex, and thus our finding further supports the key role that activation of BES genes mediated by CITFA plays in establishing monoallelic expression.

ACKNOWLEDGMENTS

I would first like to give my sincerest thanks to Dr. Günzl for his mentorship, patience, and honesty with me, and for investing in my education over these past three years. He has been such a great inspiration to me, and I likely would never have realized my interest in pursuing a MD/PhD if I hadn't had this enriching experience in his laboratory. I would also like to thank Dr. Heaslip for her ongoing support, and for all her advice, especially in her area of expertise–fluorescence microscopy. I also owe gratitude to my fellow WiMSE (Women in Math Science and Engineering) House Learning Community members, whose passion for research helped me realize my own. I of course could not have accomplished this without my family, especially my brother, who has always believed in me and challenged me to reach higher. Finally, I would like to thank the University of Connecticut's Office of Undergraduate Research for supporting this project with a Summer Undergraduate Research Fund and Travel Award.

REFERENCES

Aresta-Branco, F., Sanches-Vaz, M., Bento, F., Rodrigues, J. A., & Figueiredo, L. M. (2019). African trypanosomes expressing multiple VSGs are rapidly eliminated by the host immune system. Proceedings of the National Academy of Sciences of the United States of America, 116(41), 20725–20735. https://doi.org/10.1073/pnas.1905120116.

Brandenburg, J., Schimanski, B., Nogoceke, E., Nguyen, T. N., Padovan, J. C., Chait, B. T., Cross, G. A., & Günzl, A. (2007). Multifunctional class I transcription in Trypanosoma brucei depends on a novel protein complex. The EMBO journal, 26(23), 4856–4866. https://doi.org/10.1038/sj.emboj.7601905.

Briggs, E. M., Rojas, F., McCulloch, R., Matthews, K. R., & Otto, T. D. (2021). Single-cell transcriptomic analysis of bloodstream Trypanosoma brucei reconstructs cell

cycle progression and developmental quorum sensing. Nature communications, 12(1), 5268. https://doi.org/10.1038/s41467-021-25607-2.

Centers for Disease Control and Prevention. (2019). CDC - DPDx - Trypanosomiasis, African. Centers for Disease Control and Prevention. https://www.cdc.gov/dpdx/trypanosomiasisafrican/index.html?msclkid=0bc7709aaea311ecbb5589b7d886162d.

Chaves, I., Rudenko, G., Dirks-Mulder, A., Cross, M., & Borst, P. (1999). Control of variant surface glycoprotein gene-expression sites in Trypanosoma brucei. The EMBO journal, 18(17), 4846–4855. https://doi.org/10.1093/emboj/18.17.4846.

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. https://doi.org/10.1017/s003118200004717x.

Deitsch, K. W., & Dzikowski, R. (2017). Variant Gene Expression and Antigenic Variation by Malaria Parasites. Annual review of microbiology, 71, 625–641. https://doi.org/10.1146/annurev-micro-090816-093841.

Denninger, V., & Rudenko, G. (2014). FACT plays a major role in histone dynamics affecting VSG expression site control in Trypanosoma brucei. Molecular microbiology, 94(4), 945–962. https://doi.org/10.1111/mmi.12812.

Dreesen, O., Li, B., & Cross, G. A. (2007). Telomere structure and function in trypanosomes: a proposal. Nature reviews. Microbiology, 5(1), 70–75. https://doi.org/10.1038/nrmicro1577.

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(1), 3023. https://doi.org/10.1038/s41467-019-10823-8.

Faria, J., Luzak, V., Müller, L., Brink, B. G., Hutchinson, S., Glover, L., Horn, D., & Siegel, T. N. (2021). Spatial integration of transcription and splicing in a dedicated compartment sustains monogenic antigen expression in African trypanosomes. Nature microbiology, 6(3), 289–300. https://doi.org/10.1038/s41564-020-00833-4.

Fenn, K., & Matthews, K. R. (2007). The cell biology of Trypanosoma brucei differentiation. Current opinion in microbiology, 10(6), 539–546. https://doi.org/10.1016/j.mib.2007.09.014

Figueiredo, L. M., Janzen, C. J., & Cross, G. A. (2008). A histone methyltransferase modulates antigenic variation in African trypanosomes. PLoS biology, 6(7), e161. https://doi.org/10.1371/journal.pbio.0060161.

Ginger, M. L., Blundell, P. A., Lewis, A. M., Browitt, A., Günzl, A., & Barry, J. D. (2002). Ex vivo and in vitro identification of a consensus promoter for VSG genes expressed by metacyclic-stage trypanosomes in the tsetse fly. Eukaryotic cell, 1(6), 1000–1009. https://doi.org/10.1128/EC.1.6.1000-1009.2002.

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 of the United States of America, 113(26), 7225–7230. https://doi.org/10.1073/pnas.1600344113.

Günzl A. (2010). The pre-mRNA splicing machinery of trypanosomes: complex or simplified?. Eukaryotic cell, 9(8), 1159–1170. https://doi.org/10.1128/EC.00113-10.

Günzl, A., Bruderer, T., Laufer, G., Schimanski, B., Tu, L. C., Chung, H. M., Lee, P. T., & Lee, M. G. (2003). RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in Trypanosoma brucei. Eukaryotic cell, 2(3), 542–551. https://doi.org/10.1128/EC.2.3.542-551.2003.

Günzl, A., Kirkham, J. K., Nguyen, T. N., Badjatia, N., & Park, S. H. (2015). Mono-allelic VSG expression by RNA polymerase I in Trypanosoma brucei: expression site control from both ends?. Gene, 556(1), 68–73. https://doi.org/10.1016/j.gene.2014.09.047.

Hertz-Fowler C, Figueiredo LM, Quail MA, Becker M, Jackson A, Bason N, et al. (2008). Telomeric Expression Sites Are Highly Conserved in Trypanosoma brucei. PLoS ONE 3(10): e3527. https://doi.org/10.1371/journal.pone.0003527.

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*(6), 985–989. https://pubmed.ncbi.nlm.nih.gov/2614608/.

Horn D. (2014). Antigenic variation in African trypanosomes. Molecular and biochemical parasitology, 195(2), 123–129. https://doi.org/10.1016/j.molbiopara.2014.05.001.

Hughes, K., Wand, M., Foulston, L., Young, R., Harley, K., Terry, S., Ersfeld, K., & Rudenko, G. (2007). A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. The EMBO journal, 26(9), 2400–2410. https://doi.org/10.1038/sj.emboj.7601678.

Jespersen, N., Estelle, A., Waugh, N., Davey, N. E., Blikstad, C., Ammon, Y. C., Akhmanova, A., Ivarsson, Y., Hendrix, D. A., & Barbar, E. (2019). Systematic identification of recognition motifs for the hub protein LC8. Life science alliance, 2(4), e201900366. https://doi.org/10.26508/lsa.201900366.

Kirkham, Justin, "Investigating the Functional Roles of LC8, CITFA, and RPB7 in the Multifunctional RNA Polymerase System of Trypanosoma brucei" (2016). Doctoral Dissertations. 1438. https://opencommons.uconn.edu/dissertations/1438.

Kirkham, J. K., Park, S. H., Nguyen, T. N., Lee, J. H., & Günzl, A. (2017). Dynein Light Chain LC8 Is Required for RNA Polymerase I-Mediated Transcription in Trypanosoma brucei, Facilitating Assembly and Promoter Binding of Class I Transcription Factor A. Molecular and cellular biology, 36(1), 95–107. https://doi.org/10.1128/MCB.00705-15.

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods (San Diego, Calif.)*, *25*(4), 402–408. https://doi.org/10.1006/meth.2001.1262.

Monahan, K., & Lomvardas, S. (2015). Monoallelic expression of olfactory receptors. Annual review of cell and developmental biology, 31, 721–740. https://doi.org/10.1146/annurev-cellbio-100814-125308.

Morrison, L. J., Marcello, L., & McCulloch, R. (2009). Antigenic variation in the African trypanosome: molecular mechanisms and phenotypic complexity. Cellular microbiology, 11(12), 1724–1734. https://doi.org/10.1111/j.1462-5822.2009.01383.x.

Mugnier, M. R., Cross, G. A., & Papavasiliou, F. N. (2015). The in vivo dynamics of antigenic variation in Trypanosoma brucei. Science (New York, N.Y.), 347(6229), 1470–1473. https://doi.org/10.1126/science.aaa4502.

Narayanan, M. S., & Rudenko, G. (2013). TDP1 is an HMG chromatin protein facilitating RNA polymerase I transcription in African trypanosomes. Nucleic acids research, 41(5), 2981–2992. https://doi.org/10.1093/nar/gks1469.

Navarro, M., & Gull, K. (2001). A pol I transcriptional body associated with VSG mono-allelic expression in Trypanosoma brucei. Nature, 414(6865), 759–763. https://doi.org/10.1038/414759a.

Nguyen, T. N., Müller, L. S., 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(5), 3164–3176. https://doi.org/10.1093/nar/gkt1301

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(12), 1573–1581. https://doi.org/10.1128/EC.00250-12.

Pandya, U. M., Sandhu, R., & Li, B. (2013). Silencing subtelomeric VSGs by Trypanosoma brucei RAP1 at the insect stage involves chromatin structure

changes. Nucleic acids research, 41(16), 7673–7682. https://doi.org/10.1093/nar/gkt562.

Park, S. H., Nguyen, T. N., Kirkham, J. K., Lee, J. H., & Günzl, A. (2011). Transcription by the multifunctional RNA polymerase I in Trypanosoma brucei functions independently of RPB7. *Molecular and biochemical parasitology*, *180*(1), 35–42. https://doi.org/10.1016/j.molbiopara.2011.06.008.

Park, S. H., Nguyen, B. N., Kirkham, J. K., Nguyen, T. N., & Günzl, A. (2014). A new strategy of RNA interference that targets heterologous sequences reveals CITFA1 as an essential component of class I transcription factor A in Trypanosoma brucei. Eukaryotic cell, 13(6), 785–795. https://doi.org/10.1128/EC.00014-14.

Pena, A. C., Pimentel, M. R., Manso, H., Vaz-Drago, R., Pinto-Neves, D., Aresta-Branco, F., Rijo-Ferreira, F., Guegan, F., Pedro Coelho, L., Carmo-Fonseca, M., Barbosa-Morais, N. L., & Figueiredo, L. M. (2014). Trypanosoma brucei histone H1 inhibits RNA polymerase I transcription and is important for parasite fitness in vivo. Molecular microbiology, 93(4), 645–663. https://doi.org/10.1111/mmi.12677.

Quintana, J. F., Zoltner, M., & Field, M. C. (2021). Evolving Differentiation in African Trypanosomes. Trends in parasitology, 37(4), 296–303. https://doi.org/10.1016/j.pt.2020.11.003.

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. vol 9. https://www.frontiersin.org/article/10.3389/fimmu.2018.02253.

Rapali, P., Radnai, L., Süveges, D., Harmat, V., Tölgyesi, F., Wahlgren, W. Y., Katona, G., Nyitray, L., & Pál, G. (2011). Directed evolution reveals the binding motif preference of the LC8/DYNLL hub protein and predicts large numbers of novel binders in the human proteome. PloS one, 6(4), e18818. https://doi.org/10.1371/journal.pone.0018818.

Saura, A., Iribarren, P. A., Rojas-Barros, D., Bart, J. M., López-Farfán, D., Andrés-León, E., Vidal-Cobo, I., Boehm, C., Alvarez, V. E., Field, M. C., & Navarro, M. (2019). SUMOylated SNF2PH promotes variant surface glycoprotein expression in bloodstream trypanosomes. EMBO reports, 20(12), e48029. https://doi.org/10.15252/embr.201948029.

Schimanski, B., Nguyen, T. N., & Günzl, A. (2005). Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. Eukaryotic cell, 4(11), 1942–1950. https://doi.org/10.1128/EC.4.11.1942-1950.2005.

Schulz, D., & Papavasiliou, F. N. (2016). The VEXing problem of monoallelic expression in the African trypanosome. Proceedings of the National Academy of Sciences of the United States of America, 113(26), 7017–7019. https://doi.org/10.1073/pnas.1608546113.

Tikhonenkov, D. V., Gawryluk, R., Mylnikov, A. P., & Keeling, P. J. (2021). First finding of free-living representatives of Prokinetoplastina and their nuclear and mitochondrial genomes. Scientific reports, 11(1), 2946. https://doi.org/10.1038/s41598-021-82369-z.

Turner M. J. (1984). Antigenic variation in parasites. Parasitology, 88 (Pt 4), 613–621. https://doi.org/10.1017/s0031182000085528.

Van der Ploeg, L. H., Liu, A. Y., Michels, P. A., De Lange, T., Borst, P., Majumder, H. K., Weber, H., Veeneman, G. H., & Van Boom, J. (1982). RNA splicing is required to make the messenger RNA for a variant surface antigen in trypanosomes. Nucleic acids research, 10(12), 3591–3604. https://doi.org/10.1093/nar/10.12.3591.

Vettermann, C., & Schlissel, M. S. (2010). Allelic exclusion of immunoglobulin genes: models and mechanisms. Immunological reviews, 237(1), 22–42. https://doi.org/10.1111/j.1600-065X.2010.00935.x.

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(5), 1237–1245. https://doi.org/10.1111/j.1365-2958.2010.07284.x.

Wirtz, E., Leal, S., Ochatt, C., & Cross, G. A. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei. Molecular and biochemical parasitology, 99(1), 89–101. https://doi.org/10.1016/s0166-6851(99)00002-x.

World Health Organization. (2022). Trypanosomiasis, human African (sleeping sickness). World Health Organization. https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)?msclkid=be857542aea011ecb663f1fd71575040.

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(1), 99–109. https://doi.org/10.1016/j.cell.2009.01.037.

Ye, K., Zhang, X., Ni, J., Liao, S., & Tu, X. (2015). Identification of enzymes involved in SUMOylation in Trypanosoma brucei. Scientific reports, 5, 10097. https://doi.org/10.1038/srep10097.