5-3-2017

Investigating the Functional Roles of LC8, CITFA, and RPB7 in the Multifunctional RNA Polymerase System of Trypanosoma brucei

Justin Kirkham
UConn Health, jkirkham@uchc.edu

Investigating the Functional Roles of LC8, CITFA, and RPB7 in the Multifunctional RNA Polymerase System of *Trypanosoma brucei*

Justin K. Kirkham, MD, PhD

University of Connecticut, 2017

Abstract of the Thesis

Trypanosoma brucei, a member of the early diverged phylogenetic order Kinetoplastida, is a vector-borne parasite that causes lethal disease in both humans and livestock. Unfortunately, progress has been slow on developing new treatments, and there is a need for new therapeutics, as current therapies have issues of resistance, toxicity, and difficult administration. In order to design new therapeutics, molecules and interactions unique to the parasite must be detailed, in hopes that some will afford suitable drug targets.

One unique process in *T. brucei* that might be targeted is RNA polymerase I-mediated transcription. *T. brucei* is unique in that RNA polymerase I not only transcribes ribosomal gene units, as in all other organisms, but is also used to transcribe gene arrays that encode its major cell surface proteins, namely the variant surface glycoprotein, or VSG in the mammalian bloodstream stage of the parasite. The importance of VSG to *T. brucei* is highlighted by the fact that interference with *VSG* mRNA rapidly halts bloodstream form culture growth and leads to the clearance of trypanosomes from infected mice. Thus, targeting proteins and interactions essential for VSG production is a valid strategy against *T. brucei*.

Chapter II details an investigation of the interaction between LC8 and a class I transcription factor A (CITFA) subunit, CITFA2, which was the focus of my thesis work. Both of these proteins, and their interaction, are essential for *VSG* transcription and trypanosome viability,

and interrupting either protein or their interaction could be a potential anti-trypanosome therapy. **Chapter III** contains the generation of a method that allows for gene silencing using heterologous sequences, which was necessary for the work in chapter II. **Chapter IV** focuses on RPB7, an RNA pol II subunit, which was published to be utilized by RNA pol I for transcription. While this finding was intriguing, it contradicted our biochemical RNA pol I characterizations. Data presented in this chapter clearly demonstrated that RPB7 is not a subunit of RNA pol I and not required for the transcription process by this polymerase.

Investigating the Functional Roles of LC8, CITFA, and RPB7 in the Multifunctional RNA Polymerase System of *Trypanosoma brucei*

Justin K. Kirkham

B.S., University of Utah, 2008

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2017

APPROVAL PAGE

Doctor of Philosophy Dissertation

Investigating the Functional Roles of LC8, CITFA, and RPB7 in the Multifunctional RNA Polymerase System of Trypanosoma brucei

Presented by Justin K. Kirkham, B.S.

Major Advisor

_____

Arthur Günzl, Ph.D.

Associate Advisor

_____

Stephen King, Ph.D.

Associate Advisor

_____

Bruce Mayer, Ph.D.

Associate Advisor

_____

Justin Radolf, M.D.

Associate Advisor

_____

Blanka Rogina, Ph.D.

University of Connecticut
2017

Acknowledgements

First and foremost, I wish to thank Arthur for his kindness, his mentorship, and his investment in me and my education. It has been a humbling experience to look back over the last few years and to realize how much of my time I have spent poorly, lazily, and ignorantly, which resulted in a lack of progress for both myself and the lab. Yet, through it all, Arthur was patient with me, and instructed me on how to change so that I might have greater success. The amount of time he devoted to me was immense and undeserved. I am now in contact with my new PI, designing a project, and I already have a series of experiments I wish to conduct, once I have my own lab – this scientific future is thanks to Arthur.

I would also like to thank Tu Nguyen, Ju Huck Lee, Bao Nguyen, Sung Hee Park, and Nitika Badjatia for all that they have taught me and done for me. Tu and Ju deserve special thanks, as they taught me almost all of the technical aspects of laboratory work. Sung Hee and Nitika have also been incredibly helpful in humoring me and my tendency to talk too much as I was struggling to understand how best to move forward.

I would also like to thank my thesis committee members: Stephen King, Bruce Mayer, Justin Radolf, and Blanka Rogina, who have been incredibly encouraging throughout the difficult process of learning how to do rigorous research. It is a credit to them that I did not fear committee meetings with them, even though they were evaluating me. Their desire to help me was so clear that I was never worried when they had input or suggestions.

I would similarly like to thank Dr. Carol Pilbeam, Dr. Barbara Kream, and all of the individuals who work to make the MD/PhD program and the graduate school a success. Many individuals met with me and offered me guidance on my research, my career, and my future.

I wish to thank my parents, Brian and Diane Kirkham, who have been so supportive and helpful while I have been away from them for my education. Even when I was not sure of myself, they have always loved me and encouraged me onward, trusting in my abilities in a way that has given me strength.

Finally, I would like to thank Lindsay and Mariah Kirkham, my amazing wife and my wonderful daughter. Lindsay has forgone a career in linguistics to devote more of her time to our family and to make up for my absence. Her dedication allowed me to focus on work while I was at work, with complete confidence that my daughter and our home were in good hands. Additionally, I would often have experiments and data on my mind even after I came home, and she was willing to bear the stresses of my education with me. She has been my strength, my friend, and my comfort for many years. Mariah, born and now 4 during my time in Arthur's lab, is my source of energy and motivation. When I come home she doesn't care much about 'the bugs that make people sick that Dada is trying to stop', instead asking me if I want to play a game as soon as I step in the door. Games with her were my therapy, and I thank her for all of them. In addition, the idea of helping other children that are as sweet and wonderful as my daughter motivated me to come back to the lab after the many days of depressing data or failed experiments.

Table of Contents

Publications and Contributions to the Thesis

Chapter II. In this study I investigated the role of LC8 and CITFA2 in transcription by RNA polymerase I. I performed all experiments, save for the following: Dr. Sung Hee Park performed the sucrose gradient sedimentation of the CITFA7-purified CITFA complex, the related EMSA, UV crosslinking analysis, and the chromatin immunoprecipitation of CITFA3 in CITFA2 knockdown cells. Dr. Tu N. Nguyen generated the plasmid for knockdown of LC8, the immunoblots monitoring the knockdown, and the original transgene expression plasmid, while Dr. Ju Huck Lee produced the LC8 recombinant protein and LC8 immune serum. Dr. Arthur Günzl performed the phylogenetic analysis. This work is in preparation to be published as *Kirkham JK, Park SH, Nguyen TN, Lee JH, Günzl A. The 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.*

Chapter III. In this study, we developed a new method for specifically silencing mRNA through targeting a fused heterologous sequence, and then used this technique to study the role of CITFA1 in transcription by RNA polymerase I. I aided Dr. Park in this work by purifying the CITFA3 antibody, and producing related immunoblots. Additionally, I performed the immunofluorescence monitoring the effect of *CITFA1* silencing, aided in plasmid generation, and produced detailed maps for the two newly generated plasmids. The result of this work was published as *Park SH, Nguyen BN, Kirkham JK, Nguyen TN, Günzl A. 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. Eukaryot Cell. 2014 Jun;13(6):785-95.*

Chapter IV. We investigated whether RPB7, a known subunit of RNA polymerase II, was being utilized by RNA pol I in *T. brucei,* a claim published by the research group of Dr. Miguel Navarro (Spanish National Research Council, Granada, Spain) (Peñate et al., 2009). I aided Dr. Park in this work by performing a tandem affinity purification of RPB7, and investigating co-purifying proteins. Additionally, I performed transient transfection and immunofluorescence to determine the degree of colocalization between RPB7 and the nucleolar protein NOP10. This work was published as *Park SH, Nguyen TN, Kirkham JK, Lee JH, Günzl A. Transcription by the multifunctional RNA polymerase I in Trypanosoma brucei functions independently of RPB7. Mol Biochem Parasitol. 2011 Nov;180(1):35-42.* A response letter to the editor discussing the discrepancy of our data and those published by the Navarro grojup was published as *Günzl A, Park SH, Nguyen TN, Kirkham JK, Lee JH (2011). Response to "Role of RPB7 in RNA pol I transcription in Trypanosoma brucei". Mol Biochem Parasitol 2011 180, 45-46.*

Chapter V. In this chapter we reviewed transcriptional regulation of RNA pol I-transcribed *VSG* genes in *T. brucei,* developing a model of how trypanosomes restrict expression of the *VSG* gene family to a single gene. I co-wrote the review with Dr. Günzl, aided by all lab members, who worked together to examine the literature regarding each regulatory protein. In addition to this, I designed and produced the figures included in the review, which was published as *Günzl A, Kirkham JK, Nguyen TN, Badjatia N, Park SH. Mono-allelic VSG expression by RNA polymerase I in Trypanosoma brucei: expression site control from both ends? Gene. 2015 Feb 1;556(1):68-73.*

List of Tables

List of Figures

Chapter I

Introduction

I-1. Kinetoplastids, trypanosomatids, and relevant human disease

Kinetoplastida comprise a group of flagellated unicellular eukaryotes which contain a uniquely organized collection of DNA in their single, large mitochondrion known as a 'kinetoplast' (Stuart et al., 2008). This phylogenetic order contains *Trypanosoma brucei, T. cruzi*, and *Leishmania* spp., parasites which cause a number of devastating diseases in both humans and livestock (Brun and Blum, 2012). *T. brucei* is the causative agent of Human African Trypanosomiasis, while *T. cruzi* causes Chagas disease in Central and South America. *Leishmania* spp. cause the leishmaniases, including visceral, mucosal, and cutaneous leishmaniasis, in 88 different tropical and subtropical countries. In total, kinetoplastids account for 17% (3 of 17) of neglected tropical diseases, according to the World Health Organization (WHO, 2015). Collectively, they currently infect 20 million people worldwide, resulting in ~100,000 deaths a year, and place half a billion people at risk for infection in over 100 countries (Bilbe, 2015; Stuart et al., 2008). While 6,000 of 8,000 protein coding genes are shared among these parasites, their vectors, mechanisms of immune evasion, and resulting infections are quite different (Bilbe, 2015; Stuart et al., 2008). Yet despite these differences, it is an obvious hope that chemotherapies designed against any one kinetoplastid parasite may be useful against all members of this order.

T. brucei consists of three subspecies: *T. brucei gambiense, T. brucei rhodesiense, and T. brucei brucei*. While closely related, *T. b. brucei* is unable to cause human infection, due to the trypanolytic effect of apolipoprotein L1 (ApoL1), a regular constituent of human serum (Pays et

al., 2006; Vanhamme et al., 2003). ApoL1, as part of a distinct class of high density lipoproteins, after endocytosis, is inserted into the membrane of the lysosome, resulting in a pore which causes the lysosome to swell, resulting in osmotic stress and cell death (Wheeler, 2010). *T. b. rhodiense* is characterized by resistance to lysis by ApoL1 due to the presence of the serum resistance associated (SRA) gene (Shiflett et al., 2007; Vanhamme et al., 2003). The SRA protein binds ApoL1 in the endosome, preventing its insertion into the lysosomal membrane, thereby blocking its trypanolytic effect. *T. b. gambiense*'s resistance to ApoL1-induced lysis, however, is primarily due to a *T. b. gambiense*-specific glycoprotein (TgsGP), which is present in parasitic cellular membranes. The insertion of this glycoprotein results in a stiffer membrane that is resistant to the effects of ApoL1, a resistance that is not mediated by a direct interaction between host and parasite proteins (Berberof et al., 2001; Uzureau et al., 2013). Though *T. b. brucei* is an important animal parasite, its inability to combat human ApoL1 has resulted in it being widely used in laboratory investigations of *T. brucei,* as it represents a safe alternative which is, genetically, almost identical to the infective subspecies.

All subspecies of *T. brucei* are introduced into a mammalian host by the bite of an infected tsetse fly (genus *Glossina*), which currently limits the range of this disease to Sub-Saharan Africa (Franco et al., 2014). Though only 1% of tsetse flies are infective for *T. brucei*, a single bite is sufficient to establish infection (Maudlin and Welburn, 1989; Thuita et al., 2008). *T. b. gambiense,* which is responsible for almost all cases (98%) of Human African Trypanosomiasis (HAT) also known as Sleeping Sickness, results in a chronic disease. After introduction of the parasite into the bloodstream, individuals experience a hemolymphatic stage of the disease. The primary symptoms during this stage are headache, pruritus, and lymphadenopathy, which are all non-specific for HAT (Brun and Blum, 2012). Upon entry of the parasite into the central nervous system (CNS),

however, the neurological deficits characteristic of the second stage of this disease are observed. These include disturbances of the sleep cycle, presented as both daytime sleep and insomnia, alterations of mood and behavior, and muscular disturbances, such as weakness and tremors (Blum et al., 2006; Brun and Blum, 2012). If untreated, infection will almost invariably lead to death after an average of 3 years (Checchi et al., 2008). Infection with *T. b. rhodesiense* results in progression through the same stages, but at a much faster rate, resulting in death after an average of 6 months. Clinically, there is a greater likelihood of symptoms related to an acute infection, including fever, malaise, abdominal discomfort, and vomiting with *T. b. rhodesiense*-related HAT (Kato et al., 2015).

I-2. Current chemotherapeutics for HAT

Though these diseases have been well documented as important to both humans and livestock for over a century, treatment options remain poor (Bilbe, 2015; Lutje et al., 2010; Steverding, 2008). Four compounds are currently used for the treatment of HAT and will be briefly reviewed. Pentamidine is effective against the hemolymphatic, or first stage, of *T. b. gambiense* infection, and is one of the easiest anti-trypanosome therapies to administer, requiring only 7-10 daily intramuscular injections (Babokhov et al., 2013). It is not effective against either stage of *T. b. rhodesiense,* nor can it be used against stage II *T. b. gambiense* infection, as it does not cross the blood brain barrier (BBB), which limits its use. Suramin is the *T. b. rhodesiense* counterpart to pentamidine, as it is effective against stage I *T. b. rhodesiense,* but not against stage I *T. b. gambiense,* and is not effective against stage II of either disease. Even though it cannot cross the BBB, it does have some synergistic effects with stage II therapies, and can be used to pre-treat stage II *T. b. rhodesiense* HAT. It requires five intravenous (IV) injections, repeated every 5-7 days, over a period of 4 weeks (Brun et al., 2010). While resistance has not been seen, the length

of treatment, the requirement for IV injection, and the lack of activity against stage II disease limit the usefulness of this drug, as well. Melarsoprol, on the other hand, is an arsenic-based compound which is widely used due to its effectiveness against stage II HAT involving either subspecies (Brun et al., 2010). Furthermore, it is the only treatment for stage II *T. b. rhodesiense* HAT. Different dosing schedules are in use, with the current recommendation being one IV injection each day for ten consecutive days (Schmid et al., 2005). These dosing schedules have resulted from attempts to alleviate some of the difficulties and dangers of melarsoprol treatment. Firstly, the compound must be dissolved in a propylene glycol solution for injection, which causes significant irritation and discomfort in recipients. Secondly, due to the arsenic group within the drug, 5-10% of patients given melarsoprol experience post-treatment reactive encephalopathy (PTRE) (Babokhov et al., 2013). PTRE has a 50% mortality rate, resulting in the death of up to 5% of patients treated with this drug. Lastly, a significant increase in melarsoprol resistance has been observed, with relapse rates rising from 17.7% to 25.4% over the course of one 3-year study (Robays et al., 2008). This has resulted in melarsoprol being used at historically low levels (Babokhov et al., 2013). The last treatment to discuss is nifurtimox-eflornithine combination therapy (NECT), which is effective against stage II *T. b. gambiense* HAT. Current recommendations are for orally active nifurtimox to be taken three times daily, for ten days. At the same time, eflornithine is given as twice daily IV infusions for seven days. While this treatment has much lower toxicity and lower rates of relapse than melarsoprol, administration of NECT requires a skilled medical team, due to the need for IV infusions, is expensive, and resistance has been seen in the laboratory setting (Babokhov et al., 2013). In summary, few compounds are currently available to treat HAT, all of which are difficult to administer, costly, or have issues of resistance. There is a great need, therefore, for the pursuit of new compounds which can treat

trypanosomiasis. This would be aided by an improved understanding of molecules essential to the parasite which could serve as drug targets (Bilbe, 2015). Although not yet approved, two promising, orally available drugs are currently in the later stages of development: fexinidazole, currently in phase III clinical trials, and oxaborole SCYX-7158, which is in phase I clinical trials (Matthews, 2015).

I-3. VSG and antigenic variation

One molecule that is absolutely essential to *T. brucei* but absent from any of its host species is its variant surface glycoprotein, or VSG. VSG is a 58-kDa glycophosphatidylinositol (GPI)-anchored glycoprotein, with most of the protein, therefore, existing extracellularly (Manna et al., 2014). A single trypanosome produces 10 million identical copies of the VSG protein from a single *VSG* gene, all of which is sent to the cell surface, which results in VSG constituting 90% of cell surface peptides. This dense coat has numerous known functions: it prevents activation of the complement system (Ferrante and Allison, 1983), it allows for the removal of bound antibody through endocytosis of antibody-bound VSG (Engstler et al., 2007), after which the VSG molecule is returned to the cell surface (Pal et al., 2003), and, most importantly, it acts as a barrier to host antibodies (Schwede et al., 2011). Specifically, it was shown that antibody raised against the N-terminus of the VSG protein, the portion that is furthest away from the cell surface, is able to bind VSG, while antibody targeting the C-terminus of VSG, which is near the cell surface, is unable to bind (Schwede et al., 2011). The net result of densely packing 10 million copies of VSG protein on the cell surface is, therefore, to block recognition of its C-terminus and other constituents of the parasite surface from immune recognition. This, by itself, does not protect the parasite from clearance by the immune system, as numerous studies have demonstrated that the immune system is capable of generating antibodies which target VSG, and that this response is capable of clearing

an infection. For one such study, see Hall et al. (2013). Trypanosomes have evolved a system of VSG switching, however, which, coupled with its dense VSG coat, allows for antigenic variation and immune evasion. *T. brucei* possesses more than 2,500 *VSG* genes and pseudogenes (Cross et al., 2014), which vary dramatically in their sequence and antigenicity, while having a remarkably conserved structure, likely through disulfide bonds (Blum et al., 1993). Every 10^2 to 10^6 cell divisions, the specific VSG expressed stochastically changes, resulting in a parasite which is antigenically different, allowing it to escape recognition by antibodies generated against the last expressed VSG (Horn, 2014; Turner, 1997; Turner and Barry, 1989). This leads to an expansion of parasites with newly-expressed VSGs. Once antibodies are generated which bind a new VSG well, the clonal expansion is halted, and parasites are rapidly cleared. Repetitions of this cycle are observed in an infected host as waves of parasitemia, which occur approximately every seven days (Ross and Thomson, 1910). The switch to an alternative *VSG* gene is not in response to immune pressure, as switching is observed with *in vitro* cultures (Doyle et al., 1980). While the antigenic variation of VSG is essential for the parasite to maintain an infection, the importance of VSG to *T. brucei* is not limited to immune evasion. *VSG* silencing experiments demonstrated that a drop in *VSG* mRNA, even without a detectable drop in VSG protein levels, causes a specific cell cycle arrest in culture and a rapid clearance of parasites from infected mice (Sheader et al., 2005).

VSG genes exist in several genetic contexts (Cross et al., 2014), yet are invariably expressed from only one of 15 specialized bloodstream expression sites (BESs) (Hertz-Fowler et al., 2008). BESs contain an array of, typically, 8–9 expression-site associated genes (ESAGs), followed by a terminal VSG gene (Hertz-Fowler et al., 2008). ESAGs are likely important for the successful infection of the mammalian host, as they encode a variant heterodimeric transferrin receptor (ESAG6 and ESAG7), whose varying affinity for the transferrins of different host species

is thought to expand the parasite's list of potential hosts (Bitter et al., 1998). These also encode adenylate cyclases (ESAG4) that inhibit the innate immune system upon trypanosome lysis (Salmon et al., 2012). Switching from the expression of one *VSG* to another can be done by gene conversion, telomere exchange, or transcriptional switching, which are represented in **Figure I-1**, as reviewed in Horn (2014), and Schwede and Carrington (2010). In the case of gene conversion, an active expression site has its *VSG* gene replaced with a *VSG* gene from another location in the genome. This leaves the promoter region, *ESAG*s, and telomere unchanged. A second method of *VSG* switching involves a larger scale recombination, which exchanges the *VSG* gene and associated telomere, while leaving the promoter and *ESAGs* intact. An additional nuance is that *VSG*s need not be exchanged as indivisible units. Recombinations involving only portions of a given *VSG* gene, resulting in the formation of a new, mosaic *VSG*, have been shown to further increase VSG diversity beyond the 2,500 *VSG* genes contained in the genome, and appears to be important for the maintenance of a persistent infection (Cross et al., 2014; Hall et al., 2013; Mugnier et al., 2015). The last type of switching involves no recombination, occurring instead through in situ [in]activation. In transcriptional switching, the active BES is down-regulated, while expression of a second, formerly silent BES increases. An additional unique aspect of *VSG* transcription in *T. brucei* is that the active *VSG* gene is located in an extranucleolar focus. This subnuclear compartment contains both the actively transcribed BES and an accumulation of RNA pol I and *VSG* mRNA, and was maintained after DNase digestion (Chaves et al., 1998; Navarro and Gull, 2001). It was termed the expression site body, or ESB. The importance of this location was confirmed by experiments in which the integration of selectable markers into two BESs allowed for the partial co-activation of two BESs within the same cell (Chaves et al., 1999). When

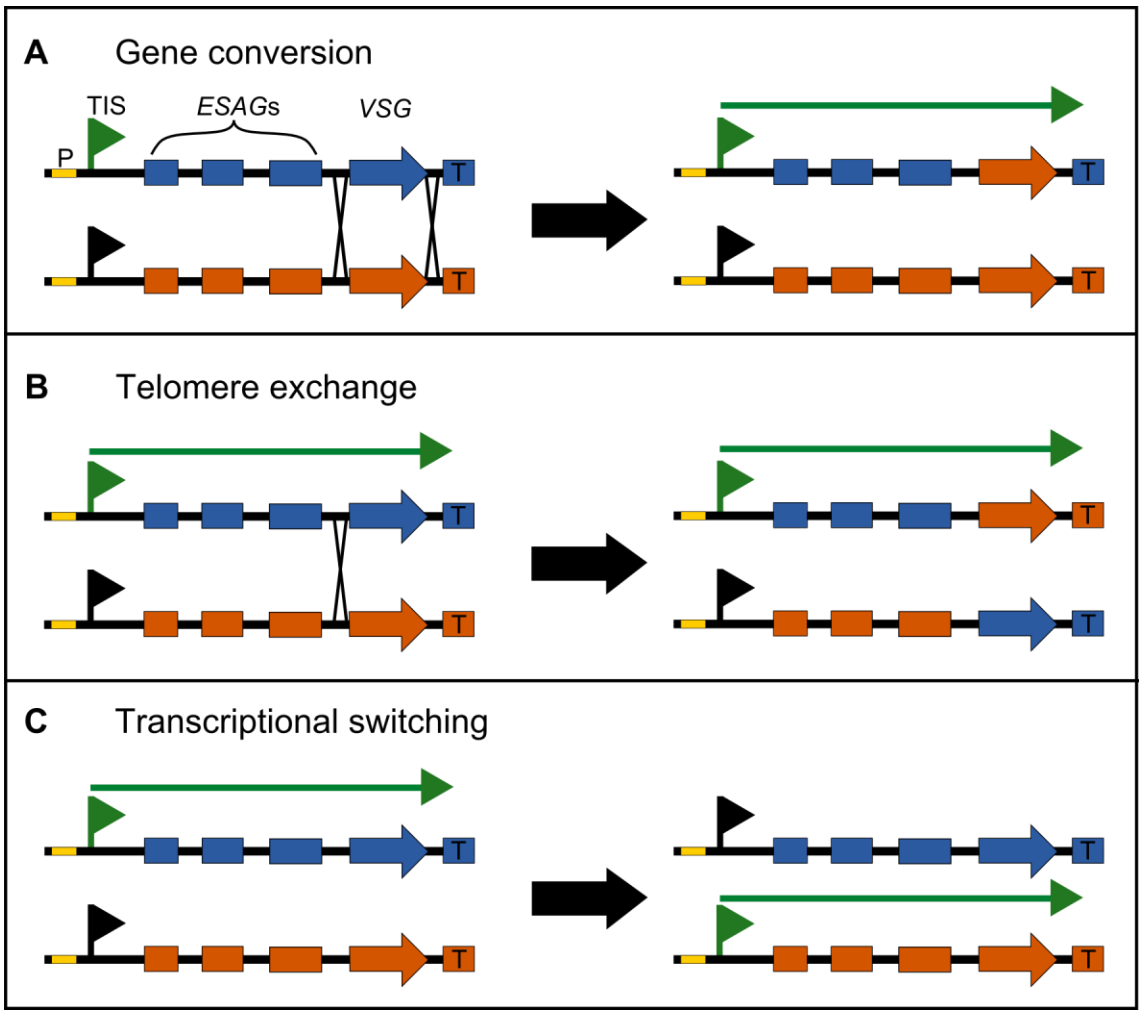


Figure I-1. Antigenic variation can occur through three different mechanisms. In bloodstream form (BF) *T. brucei*, the actively expressed *VSG* gene is located within a subtelomeric array known as a BES. Each BES contains a promoter, a variable number of *ESAG* genes, typically 8-9, and a terminal *VSG* gene near the telomere. Only one of the approximately 15 BESs within a cell is active, indicated here by a green flag, representing an active site of transcription initiation, and a green arrow, indicating transcription. *T. brucei* can switch from the actively expressed *VSG* (blue) to a silent *VSG* (orange) by three different mechanisms. Gene conversion allows for *VSG* genes from any genomic context, here shown as a silent BES, to be introduced into the active site, leaving

(Figure I-1 legend cont.) all other elements of the active BES intact (**A**). Telomere exchange, a similar process, results in the exchange of both the *VSG* gene and its associated telomere (**B**). This method of switching can only be used with a *VSG* gene adjacent to a telomere. Transcriptional switching, on the other hand, is accomplished through epigenetic mechanisms, and results in a change of expression in both *VSG* and *ESAGs* (**C**). Figure not to scale. For a scaled depiction of a BES, see **Figure V-1**.

co-selection was induced, the two BESs were found to be in close spatial proximity, and the authors concluded that some factor in the active site was essential for BES activation.

I-4. Polycistronic transcription and RNA pol I

The majority of eukaryotic genomes are arranged for monocistronic transcription, with a stoichiometry of one promoter to one gene to one transcription termination site. *T. brucei* is one notable exception to this rule, however, along with a number of other eukaryotic organisms (Lasda and Blumenthal, 2011). *T. brucei*'s genome is organized into functionally unrelated polycistronic gene arrays which, after transcription, are resolved into individual mRNAs by spliced-leader (SL) *trans*-splicing and polyadenylation (Günzl, 2010; Michaeli, 2011). BESs are one such polycistronic array, in which a single BES promoter drives the expression of 8-9 *ESAG*s, in addition to one *VSG* gene (Hertz-Fowler et al., 2008). This process of adding a 39-bp long SL cap to every mRNA in the cell by *trans*-splicing has an additional implication beyond allowing for fewer promoters and termination sites. It decouples the process of mRNA capping from transcription by RNA polymerase (pol) II (Günzl, 2010; Michaeli, 2011; Preußer et al., 2012), and has allowed for the evolution of a unique characteristic of *T. brucei*, namely the use of RNA pol I for protein coding gene expression.

The ability for RNA pol I to synthesize functional mRNA in trypanosomes was first appreciated when an *RRNA* promoter was used to express a reporter gene. In mice, the reporter gene was well transcribed, yet the resulting protein levels were very low (Grummt and Skinner, 1985). In trypanosomes, however, the *RRNA* promoter caused a high level of functional reporter protein to be produced (Rudenko et al., 1991; Zomerdijk et al., 1991a). Additional evidence for RNA pol I having a role in protein coding gene expression in the mammalian-infective

bloodstream form (BF) trypanosomes came when α-amanitin, an amatoxin found in the *Aminita* genus of mushrooms which blocks RNA pol II transcription, was shown to not block *VSG* transcription (Kooter and Borst, 1984). These implications were directly tested by silencing the largest subunit of RNA pol I, which resulted in a drop in transcription of *RRNA*, *VSG*, and procyclin, without affecting RNA pol II-mediated transcription (Günzl et al., 2003). Furthermore, it was shown that depletion of RNA pol I from extract by immunoprecipitation resulted in a drop of *in vitro* transcription from *RRNA* and BES promoters, while RNA pol II transcription was unaffected (Günzl et al., 2003). Procyclin, the major cell surface antigen in insect-stage, or procyclic form (PF) *T. brucei* is thought to protect against the digestive enzymes of its host (McConville and Ferguson, 1993). Though this is different from VSG in the type of protection it provides the parasite, it has been shown to be essential for full PF infectivity, and is somewhat considered, therefore, the PF counterpart to VSG (Ruepp et al., 1997).

I-5. The search for a class I transcription factor

After it was established that *VSG* transcription relies on RNA pol I, the search for other proteins essential to the production of VSG began. Given the high level of divergence between kinetoplastids and other eukaryotes, no class I transcription factors were annotated by sequence homology in completed trypanosome genomes (Berriman et al., 2005; Ivens et al., 2005). A first indication as to the nature of *T. brucei*'s class I factor came from *in vitro* transcription experiments using the *RRNA*, BES, and procyclin promoters. The BES promoter is short, extending 67 bp upstream of the transcription initiation site, and only contains two small required sequence elements (Pham et al., 1996; Vanhamme et al., 1995). The *RRNA* and procyclin promoters, on the other hand, are much longer, reaching ~250 bp upstream of the transcription initiation site (TIS), each containing four sequence elements (Brown et al., 1992; Janz and Clayton, 1994; Laufer and

Günzl, 2001; Sherman et al., 1991). Despite these differences, it was observed that all of these three promoters were capable of competing for a common trans-activating factor, making it likely that a common class I transcription factor was being utilized in all three cases (Laufer and Günzl, 2001). Additionally, it was shown that trypanosome nuclear extract was capable of specifically binding and shifting the BES promoter in an electrophoretic mobility shift assay (EMSA) (Pham et al., 1997). Deoxycholate treatment, which is able to disrupt protein-protein interactions, led to multiple, faster migrating bands in an EMSA, which the authors suggested were indicative that the promoter binding element consisted of more than one protein, which perhaps formed a larger complex. These two observations led to the key series of experiments that revealed the *trans*-activating factor (Günzl, 2012). Brandenburg et al. prepared crude trypanosome extract from 30 liters of trypanosome culture, and tested fractions for BES promoter binding using EMSA (Brandenburg et al., 2007). EMSA, in combination with ion exchange, heparin affinity, and DNA affinity chromatography, allowed the authors to partially purify the BES promoter binding activity. Since a ~50 kDa protein band was found to specifically UV-crosslink to the BES promoter, mass spectrometric analysis of the final eluate was concentrated on proteins migrating in SDS-PAGE at this size range. Seventeen putative BES promoter interactors were identified, most of which were without annotation. Tagging four of these proteins in individual trypanosome cell lines, combined with a promoter pull-down assay, identified one protein that bound the BES promoter while not binding an RNA pol II promoter or a non-specific DNA control. This binding required both promoter elements, and this protein was also found to bind to the *RRNA* and procyclin gene promoters, indicating that it could be the general *trans*-activating factor suggested in earlier work. Brandenburg et al. went on to show that silencing this gene by RNAi resulted in rapid cell death and a decrease in *RRNA* and *VSG* mRNA, while RNA pol II and RNA pol III transcripts were

unaffected (Brandenburg et al., 2007). *In vitro* transcription assays in the presence of an immune serum raised against this protein then unambiguously showed that it was absolutely essential for transcription from the *RRNA*, BES, and procyclin promoters.

Tagging of this protein, followed by tandem affinity purification and mass spectrometry, identified six proteins unique to kinetoplastids, which were all annotated as 'hypothetical' proteins, and a dynein light chain, termed DYNLL1 (Brandenburg et al., 2007) (Please note that due to work presented in **Chapter II** of this thesis, it has been proposed to rename DYNLL1 as LC8). Co-immunoprecipitation and sucrose gradient sedimentation confirmed that these proteins formed a complex, and the purified complex was capable of specifically binding the BES promoter in EMSAs. Finally, depleting the complex from extract via a tag abolished the ability of that extract to initiate transcription from *RRNA*, BES, or procyclin gene promoters. Adding back the purified complex to these depleted extracts was able to partially reconstitute transcriptional activity (Brandenburg et al., 2007). Since, together, these experiments identified this protein complex as the first promoter-binding factor essential for RNA pol I-mediated transcription in a kinetoplastid organism, it was termed class I transcription factor A, or CITFA, with subunits numbered according to their predicted size, from largest to smallest. Due to high sequence conservation of CITFA subunits in other kinetoplastids, and CITFA's indispensability for *RRNA* transcription, it is likely that its function in RNA pol I-mediated transcription is shared with other members of this order.

I-6. The discovery of an additional CITFA subunit

The lack of sequence homology to known transcription factors and of recognizable amino acid sequence motifs has hindered attribution of specific functions to individual CITFA subunits.

Nevertheless, additional studies built upon the above work, provided new insights into the composition of the complex, and its role in transcription by RNA pol I in *T. brucei.* Through tandem affinity purification of CITFA6, in conjunction with mass spectrometry, an additional CITFA subunit, CITFA7, was identified (Nguyen et al., 2012). In turn, co-immunoprecipitation and tandem affinity purification of a tagged CITFA7 quantitatively co-precipitated and copurified all other CITFA subunits confirming it as a *bona fide* subunit of CITFA. As expected, *CITFA7* silencing in BF *T. brucei* resulted in a drop in both *rRNA* and *VSG* mRNA, and was lethal. Furthermore, similar experiments were performed to investigate CITFA7 as had been done in CITFA2, and revealed that depletion of CITFA7 in cells or in extract greatly reduced *in vitro* transcription from both a BES and an *RRNA* promoter (Nguyen et al., 2012). In accordance with CITFA7 being a vital component of CITFA, immunofluorescence revealed that CITFA7 and RPB6z, an RNA pol I-specific subunit (Nguyen et al., 2006), colocalized in both the nucleolar periphery, where *RRNA* genes are transcribed, and in an extranucleolar foci, likely representing the ESB. These results confirmed that CITFA7 is a class I transcription factor subunit required by RNA pol I. No new members of CITFA have since been identified. The complete list of CITFA subunits, along with their predicted sizes and accession numbers, is presented in Table I-1.

I-7. A possible role for CITFA in monoallelic VSG expression

Silent BESs are not completely silent in their promoter region. A selectable marker gene inserted downstream of a silent BES is able to make trypanosomes resistant to the corresponding antibiotic at a level 100 fold below that achieved when the marker gene is inserted at the same position in the active BES (Horn and Cross, 1997). Moreover, a study using mild DNase I digestion of chromatin showed that the same DNase-sensitive sites, indicative of specific protein binding, were observed at both silent and active BES promoters, leading to the conclusion that promoters of both

active and silent BESs are occupied by the same transcription factor (Navarro and Cross, 1998). Since this factor likely was CITFA, Nguyen et al. next quantitatively investigated whether CITFA was associating with all BES promoters to the same extent, or if it was predominately associating with the promoter of the active BES. To test this, they used a previously published cell line in which two BESs had selectable markers integrated downstream of the promoter (Figueiredo et al., 2008). Through addition of antibiotics, neomycin-resistance or puromycin-resistance could be selected for. The changing of medium antibiotics did not convert the population of *T. brucei* to be resistant, rather it allowed those stochastic switchers, which had activated the alternative BES with the selectable marker through transcriptional switching (see **Figure I-1**), to be resistant to the new antibiotic.

An extensive chromatin immunuprecipitation study revealed that CITFA7 predominately associated with the active BES promoter (Nguyen et al., 2014). This biased association was maintained through consecutive rounds of antibiotic-selected transcriptional switching between the two marked BESs. In every experiment a high level of CITFA7 occupancy correlated with a high degree of promoter-proximal transcripts originated from that BES, while a low level of occupancy correlated with a low level of expression. They also showed, through ChIP experiments involving RPB6z, an RNA pol I subunit, that RNA pol I had a promoter occupancy similar to CITFA7 – it was present at the active BES, while being relatively absent from the marked silent BES. These data showed that CITFA7 and RNA pol I are both predominately associated with the active site, revealing that the activation of a single BES and the monoallelic expression of *VSG* may be regulated at the level of transcription initiation, perhaps through sequestering CITFA or RNA pol I to the active BES.

Subunit	GeneDB accession	Approximate size (kDa)
CITFA1	Tb927.11.1390	55
CITFA2	Tb927.9.12450	55
CITFA3	Tb927.11.1410	55
CITFA4	Tb927.11.8310	43
CITFA5a	Tb927.8.4030 Tb927.8.4080	28
CITFA5b	Tb927.8.4130	28
CITFA6	Tb927.5.970	23
CITFA7	Tb927.7.2600	17
LC8	Tb927.11.18680 Tb11.0845	10

Table I-1. *T. brucei* CITFA subunits, along with their GeneDB accession numbers and approximate sizes.

Analyzing *RRNA* units, a key experiment demonstrated that CITFA is directly responsible for RNA pol I occupancy of chromatin. Nguyen et al. performed ChIP experiments with RPA31, another RNA pol I subunit (Nguyen et al., 2007), in cells wherein *CITFA7* could be conditionally silenced through addition of doxycycline. While in non-induced cells RNA pol I was found to be present at both the *RRNA* promoter and the downstream 18S rDNA gene, it became relatively depleted from both of these regions upon *CITFA7* silencing, with RPA31 occupancy dropping ~80% compared to non-induced cells (Nguyen et al., 2014). This confirmed in vivo that CITFA is required for the recruitment of RNA pol I to its cognate genes.

Given that CITFA2 had previously been shown to be essential for RNA pol I-mediated transcription (Brandenburg et al., 2007), Nguyen et al. sought to determine whether CITFA2 showed the same preferential occupancy of the active BES promoter as CITFA7. Using cells with two marked BESs, as before, ChIP experiments revealed that CITFA2, like CITFA7, was predominately associated with the active BES, and that this relationship was maintained through multiple rounds of transcriptional BES switching. To further confirm the predominant association of CITFA with the actively transcribed BES, a cell line was created in which CITFA7 and RPB6z were tagged with eYFP and mCherry, respectively, allowing detection by direct fluorescence microscopy. The enhanced sensitivity of this method revealed that CITFA7 and RPB6z colocalized both in the nucleolus and the extranucleolar ESB, the sites of *RRNA* and *VSG* transcription, respectively. Given that there are approximately 15 BESs in *T. brucei* (Hertz-Fowler et al., 2008), with one being active and 14 being silent, distributed extranucleolarly throughout the nucleus, the failure to detect additional foci of CITFA and RNA pol I, beyond the ESB, confirmed that CITFA and RNA pol I are not concentrated at the silent BESs.

Through quantification of immunoblots and fluorescence images, Nguyen et al. estimated that approximately 180 molecules of CITFA7 are present in the ESB. This number was surprisingly high, since DNA-bound RNA pol I transcription factors in other systems are known to remain associated with the promoter through multiple rounds of transcription initiation and the short BES promoter could only accommodate one or two CITFA complexes at a time. To investigate whether CITFA was accumulating in the ESB through binding additional regions of the BES, beyond the promoter, CITFA7 ChIP-seq was performed. In line with their earlier CITFA7 ChIP results, ChIP-seq revealed a strong read density peak at the BES promoter, which declined rapidly both upstream and downstream of the promoter. No other peaks were observed along a BES. This confirmed that CITFA is a promoter binding factor, and that it does not travel with the polymerase during transcription. The amount of CITFA7 present in the ESB is likely the result, therefore, of a concentration of CITFA to the active BES by a DNA-independent process. Taken together, these data confirmed that CITFA is a promoter binding RNA pol I transcription factor, and show that CITFA predominantly associates with the active site. They also show that the recruitment of RNA pol I to the *RRNA* promoter requires CITFA. It is possible, therefore, that monoallelic expression of *VSG* is achieved through the association of CITFA with a single BES.

Several molecules, in addition to CITFA, have been shown to predominately associate with either active or silent BESs and appear to play a role in monoallelic *VSG* transcription. These were recently reviewed by Günzl et al. (2015), which was included as **Chapter V**.

I-8. The development of new tools required for further investigations of CITFA

During their efforts to characterize the function of individual CITFA subunits, Park et al. attempted to silence *CITFA1* by RNAi. Though different regions of CITFA1 were targeted for degradation,

none of these efforts led to a detectable reduction in *CITFA1* mRNA. In order to circumvent this problem, we developed a new system of gene silencing, which relied on targeting heterologous sequences fused to the gene of interest (Park et al., 2014). This work is covered in depth in **Chapter III**, and only a brief summary is included here. As proof of principle, the system was first developed for *CITFA7* since this subunit had been meticulously analyzed before. *CITFA7* was exclusively expressed with a 3′ UTR from a *T. brucei*-related trypanosomatid, *T. cruzi*, or with the PTP tag sequence. Expression of double-stranded RNA (dsRNA) targeting either heterologous sequence resulted in a strong reduction of *CITFA7* mRNA, and caused the same RNA pol I defects as before, when *CITFA7* was silenced by a dsRNA that targeted its coding sequence. Fusing the *T. cruzi* 3′ UTR to CITFA1 then indeed enabled a specific knockdown of this gene and caused a drop in *RRNA* and *VSG* mRNA levels similar to those observed upon silencing *CITFA7* (Park et al., 2014). *CITFA1* silencing was equally lethal to the parasite. Immunoblotting and sucrose gradient sedimentation of the CITFA complex showed that the expression of other CITFAs was not affected and the formation of the CITFA complex not disrupted by this knockdown. Similarly, immunofluoresence microscopy using a polyclonal antiCITFA3 antibody revealed that the localization of the CITFA complex was not affected by *CITFA1* silencing. Finally, ChIP experiments revealed that depletion of CITFA1 caused CITFA to no longer occupy either the *RRNA* or BES promoters, indicating a specific function of this subunit in DNA binding of the complex. The fact that CITFA remains localized to the nucleolus, and to an extranucleolar focus likely to be the ESB, upon *CITFA1* silencing indicates that CITFA does not rely on DNA binding for localization. This supports the notion that the ESB specifically contains a factor that is required for transcription of the active BES (Nguyen et al., 2014), and suggests that CITFA may be that factor.

I-9. The discovery and essential nature of a widely-conserved dynein light chain homologue

The only CITFA subunit conserved outside of kinetoplastids is the dynein light chain LC8. LC8 was originally discovered as a component of the outer arm axonemal dynein in *Chlamydomonas reinhardii* (Pfister et al., 1982). Genetic analysis suggested that this light chain was 88% identical to a homologue in *C. elegans* (King and Patel-King, 1995). Given that nematodes do not have flagella or motile cilia at any stage of their life cycle, the authors suggested that this dynein light chain homologue must function in the cytoplasm, likely as a component of cytoplasmic dynein. This was shortly thereafter confirmed, when the mammalian homologue of the dynein light chain was found to specifically co-purify with cytoplasmic dynein (King et al., 1996). Immunofluorescence microscopy revealed that this homologue co-localized with cytoplasmic dynein, but not with kinesin, providing additional evidence to the concept that this light chain homologue was shared between axonemal and cytoplasmic dynein, and that it was widely conserved among eukaryotes. Most commonly, this dynein light chain is termed DYNLL or LC8, though gene names vary based on organism and discovery history.

Cytoplasmic dynein exists as two different complexes: dynein-1, which has been better characterized, is involved in membrane trafficking, organelle dynamics, and chromosome segregation during mitosis (Paschal and Vallee, 1987), and dynein-2, which regulates retrograde intraflagellar transport (IFT), a process important to motile cilia and flagella (Criswell et al., 1996; Gibbons et al., 1994). LC8 has been shown to be a part of both of these cytoplasmic dynein complexes (Asante et al., 2014), where it binds to intermediate chains, as shown by Lo et al. (2001) and reviewed by Wu and King (2003), and appears to contribute to their structural arrangement (Makokha et al., 2002). Additional complexity became appreciated when it was demonstrated that two, nearly identical *LC8* genes exist in humans and numerous other organisms (Naisbitt et al.,

2000; Wilson et al., 2001). These two genes, termed DYNLL1 and DYNLL2, give rise to proteins that differ in only six of 89 amino acids, and many investigation of LC8 function do not allow for differentiation between these two proteins (Pfister et al., 2006). Furthermore, it has been shown that DYNLL1 and DYNLL2 are shared between dynein-1 and dynein-2, making a general differentiation of their functions difficult (Asante et al., 2014). Hereafter, DYNLL1 and DYNLL2 will be collectively referred to as LC8. *LC8* is conserved throughout eukaryotic genomes (Wickstead and Gull, 2007), and while not essential in yeast (Stuchell-Brereton et al., 2011), mutation or knockdown of *LC8* causes a mitotic block in HeLa cells (Asthana et al., 2012), and is embryonic lethal in animals (Dick et al., 1996; Goggolidou et al., 2014; Li et al., 2015; Lightcap et al., 2009), causing pleiotropic defects.

I-10. LC8 outside of the dynein motor complex

Given that LC8 is more conserved between species than other components of the dynein motor, and that LC8 is present in organisms which lack a dynein motor, it was likely that LC8 had non-dynein functions (King and Patel-King, 1995; Wickstead and Gull, 2007). LC8 has since been shown to interact with 50+ different proteins and to affect numerous cellular processes independent of its involvement with dynein, as reviewed by Barbar (2008) and Rapali et al.,(2011b). Though diverse in their function, LC8's binding partners tend to congregate into the functional categories of intracellular transport, nuclear transport, mitosis, apoptosis, and transcriptional regulation (Rapali et al., 2011b).

At physiological pH, LC8 exists almost exclusively as a dimer (Barbar et al., 2001; Benashski et al., 1997), interacting with partner proteins via two identical sites generated at the dimer interface which bind to diverse short, linear motifs (Benashski et al., 1997; Lo et al., 2001; Rapali

et al., 2011a). Structurally, each LC8 monomer possesses five β-strands and two α-helices, and LC8 dimerization requires β-strand swapping (Wang et al., 2003). LC8 promotes the dimerization of its binding partners through aligning dimerization domains present in the partner protein (Barbar, 2008; Barbar and Nyarko, 2014; Hodi et al., 2006; Wagner et al., 2006). Thus, while it was previously hypothesized that LC8 functions as a linker, allowing attachment of the dynein motor to its cargo, the emerging view is that interaction with LC8 induces homodimerization, imparting new structure and function, which is supported by detailed investigations of LC8 interactions (Barbar and Nyarko, 2015; Radnai et al., 2010; Wang et al., 2004; Williams et al., 2007). **Figure I-2** depicts a simplified current model of LC8 binding: LC8 binds a protein that is already dimerized, as depicted, or which has a propensity for self-association. Through interacting with a partner protein via a binding site created at its dimer interface, LC8 confers stability and structure, which often includes an increase in self-association. Due to the conserved nature of LC8, its diverse binding network, and its ability to modulate several aspects of cell biology, it has been termed a molecular hub (Barbar, 2008).

In addition to this improved understanding of LC8, interest in this molecule has grown as LC8 has been shown to have unique and important roles in an increasing number of human pathogens. These include viruses, such as HIV (Jayappa et al., 2015), Ebola (Luthra et al., 2015), and rabies (Tan et al., 2007), and the protistan parasite *Toxoplasma gondii* (Qureshi et al., 2013).

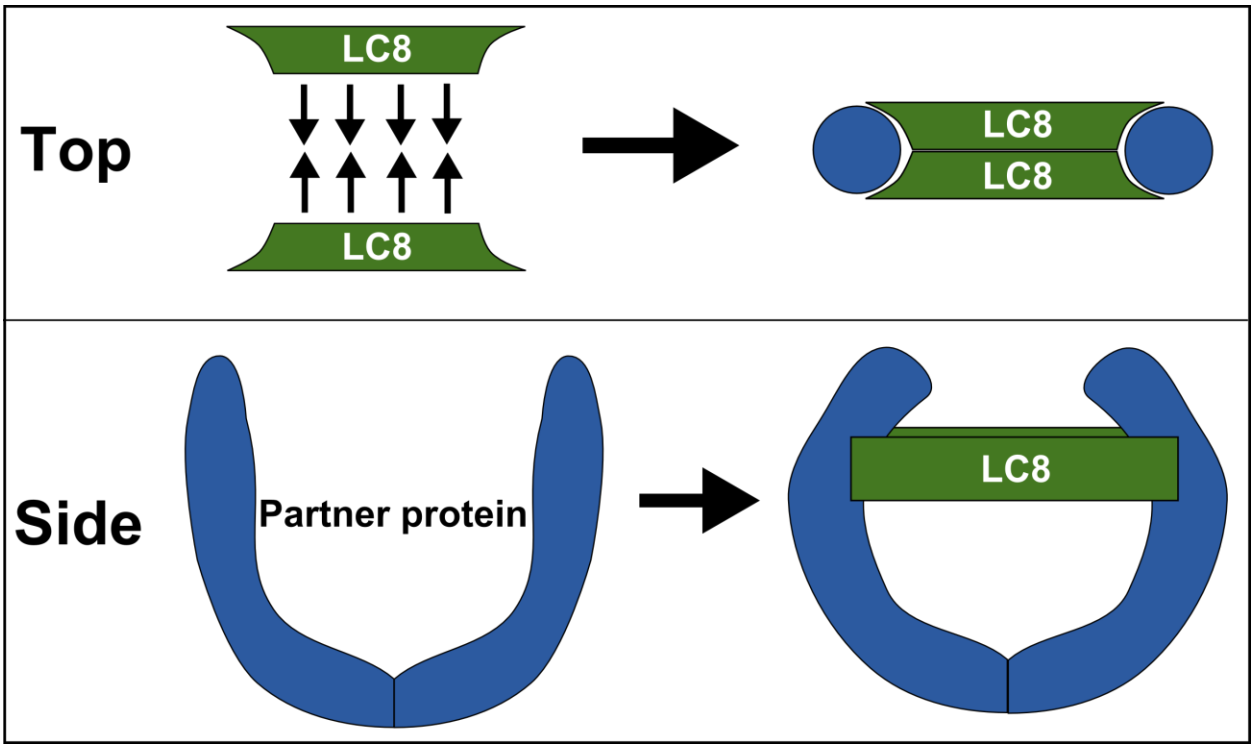


Figure I-2. Current model of LC8 interaction with a generic partner protein. LC8 exists as a dimer, which creates two identical binding sites at its dimer interface. The structure and/or dimerization state of the partner protein is often affected through interaction with LC8.

Chapter II

The 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

Abstract

Dynein light chain LC8 is highly conserved among eukaryotes and has both dynein motor and dynein-independent functions. Interestingly, LC8 was identified as a subunit of the class I transcription factor A (CITFA), which is essential for transcription by RNA polymerase (pol) I in the parasite *Trypanosoma brucei.* Given that LC8 has never been identified with a basal transcription factor and that *T. brucei* relies on RNA pol I for expressing the variant surface glycoprotein (VSG), the key protein in antigenic variation, we investigated the CITFA-specific role of LC8. Depletion of LC8 from mammalian-infective bloodstream trypanosomes affected cell cycle progression, reduced the abundances of rRNA and *VSG* mRNA, and resulted in rapid cell death. Sedimentation analysis, co-immunoprecipitation of recombinant proteins, and bioinformatic analysis revealed an LC8 binding site near the N-terminus of the subunit CITFA2. Mutation of this site prevented the formation of a CITFA2-LC8 heterotetramer and, *in vivo*, was lethal, affecting assembly of a functional CITFA complex. Gel shift assays and UV-crosslinking experiments identified CITFA2 as a promoter-binding CITFA subunit. Accordingly, silencing of *LC8* or *CITFA2* resulted in a loss of CITFA from RNA pol I promoters. Hence, we discovered an LC8 interaction that, unprecedentedly, has a basal function in transcription.

II-1. Introduction

Dynein light chain LC8 was originally discovered as a component of the outer arm axonemal dynein in *Chlamydomonas reinhardii* (Pfister et al., 1982), but was later found to also be present in cytoplasmic dyneins 1 and 2 (Asante et al., 2014; Hou and Witman, 2015; King and Patel-King, 1995; Pfister et al., 2006). LC8 is conserved throughout eukaryotic genomes (Wickstead and Gull, 2007). As a part of the dynein motor, it is important for fundamental cellular processes, such as tubulin minus end-directed intracellular transport, chromatid separation during mitosis, and nuclear migration (Fridolfsson et al., 2010), as well as flagella-specific functions, namely motility, intraflagellar transport (Pazour et al., 1998), and ciliogenesis (Goggolidou et al., 2014; Li et al., 2015). While not essential in yeast (Stuchell-Brereton et al., 2011), mutation or knockdown of *LC8* is embryonic lethal in animals (Dick et al., 1996; Goggolidou et al., 2014; Li et al., 2015; Lightcap et al., 2009). Given that LC8 is more conserved between species than other components of the dynein motor, and that LC8 is present in organisms which lack a dynein motor, it was likely that LC8 had non-dynein functions (King and Patel-King, 1995; Wickstead and Gull, 2007). LC8 has since been shown to interact with several different proteins and to affect various cellular processes, including protein localization and stability, transcription regulation, and apoptosis (Asthana et al., 2012; Barbar, 2008; Rapali et al., 2011b).

It was previously hypothesized that LC8 functions as a linker, allowing attachment of the dynein motor to its cargo. It has been shown since, however, that at physiological pH, LC8 exists almost exclusively as a dimer (Barbar et al., 2001; Benashski et al., 1997), interacting with partner proteins via two identical sites generated at the dimer interface which bind to diverse short, linear motifs (Benashski et al., 1997; Lo et al., 2001; Rapali et al., 2011a). LC8 promotes the dimerization

of its binding partners through aligning dimerization domains present in the partner protein (Barbar, 2008; Barbar and Nyarko, 2014; Hodi et al., 2006; Wagner et al., 2006). The emerging view, therefore, is that interaction with LC8 induces dimerization, imparting new structure and function, which is supported by detailed investigations of LC8 interactions (Barbar and Nyarko, 2015; Radnai et al., 2010; Wang et al., 2004; Williams et al., 2007). In addition to this improved understanding of LC8, interest in this molecule has grown as LC8 has been shown to have unique and important roles in an increasing number of human pathogens. These include viruses, such as HIV (Jayappa et al., 2015), Ebola (Luthra et al., 2015), and rabies (Tan et al., 2007), and the protistan parasite *Toxoplasma gondii* (Qureshi et al., 2013).

Trypanosoma brucei, a member of the early diverged phylogenetic order Kinetoplastida, is a vector-borne parasite that causes lethal disease in both humans and livestock (Brun and Blum, 2012). In *T. brucei*, LC8 was identified in mass spectrometry analysis of the flagellar matrix (Broadhead et al., 2006; Oberholzer et al., 2011) and, surprisingly, as a subunit of the class I transcription factor A (CITFA) (Brandenburg et al., 2007; Nguyen et al., 2012). As CITFA is a core promoter-binding factor required for initiation of RNA polymerase (pol) I-mediated transcription, this represents the first time LC8 has been found to associate with basal transcription machinery. *T. brucei* is unique in that RNA pol I not only transcribes ribosomal gene units (*RRNA*), as in all other organisms, but is also used to transcribe gene arrays that encode its major cell surface proteins, namely the variant surface glycoprotein, or VSG, in the mammalian-infective bloodstream form (BF) and procyclin in the insect-stage procyclic form (Günzl et al., 2003). This production of functional mRNA by RNA pol I is possible in *T. brucei* due to a unique mRNA processing mechanism, called spliced leader (SL) *trans*-splicing, which caps mRNA post-transcriptionally by an RNA pol II-independent process (Rudenko et al., 1991; Wirtz et al., 1994;

Zomerdijk et al., 1991a). By densely covering the cell in ~10 million copies of the same VSG protein, *T. brucei* is able to shield invariant proteins from antibody recognition (Schwede et al., 2011). The source of this massive protein expression is a single *VSG* gene located in one of ~15 bloodstream expression sites (BESs) that are monoallelically expressed (Hertz-Fowler et al., 2008). Antigenic variation of VSG, which occurs by switching to the expression of another of the ~2000 VSG genes in the trypanosome genome, allows for an indefinite infection to be maintained (Cross et al., 2014; Horn, 2014). The importance of VSG to *T. brucei* is highlighted by the fact that interference with *VSG* mRNA rapidly halts BF culture growth in the absence of immunological pressure and leads to the clearance of trypanosomes from infected mice (Sheader et al., 2005).

While it was previously shown that the CITFA complex, consisting of subunits CITFA1-7 and LC8, is essential to RNA pol I-mediated transcription and binds the BES promoter in purified form (Brandenburg et al., 2007)), the specific role of individual complex members, including knowledge of the LC8 binding partner, has remained unclear. Furthermore, given that CITFA subunits except LC8 are conserved only among kinetoplastids and are without recognizable sequence motifs, that LC8 has never been implicated in the basal process of transcription initiation, and that LC8 has not been studied in a kinetoplastid organism, we set out to understand the specific role of LC8 in RNA pol I-mediated transcription.

Kinetoplastids encode two distinct LC8 proteins, only one of which was found to be associated with CITFA. We found that this LC8, previously termed DYNLL1 (Brandenburg et al., 2007; Nguyen et al., 2012), is essential for cell viability in culture and that RNAi-mediated silencing of the gene led to defects in both cell cycle and transcription by RNA pol I. To understand LC8's specific role in the latter, we identified an LC8 binding site near the N-terminus of the essential

CITFA2 subunit. Mutation of this site was lethal to trypanosomes, preventing the incorporation of this subunit into the CITFA complex. Moreover, we show that CITFA2 directly interacts with BES promoter DNA and is required for the CITFA complex to bind to the BES promoter *in vivo*, functions that crucially depend on the CITFA2-LC8 interaction. These data revealed an essential role for LC8 in *T. brucei*, and the first evidence that LC8 is required for the formation of a transcription pre-initiation complex in any organism.

II-2. Materials and Methods

DNAs and cell lines. pT7-LC8-stl, for conditional silencing of *LC8* genes (accession numbers Tb927.11.18680 and Tb11.0845 [www.genedb.org or www.tritrypdb.org]), was generated by inserting portions of the LC8 coding region and its adjacent 3′ UTR, nucleotides +97 to +602 relative to the translation initiation codon, into the pT7-stl vector (Brandenburg et al., 2007) in a sense-stuffer-antisense arrangement, according to a published protocol (Shi et al., 2000). Transfection of SacII-linearized pT7-LC8-stl into single marker (sm)BF cells (Wirtz et al., 1999) generated smLC8 cells. pCITFA7-PTP-NEO, pCITFA4-PTP-NEO (Nguyen et al., 2012), and pPTP-CITFA2-PURO (Brandenburg et al., 2007) are plasmids that were described previously, and were used to fuse the sequence of the composite PTP tag, consisting of a tandem protein A domain (ProtA), a Tobacco etch virus cleavage site, and a protein C (ProtC) epitope, to endogenous alleles. Similarly, the gene silencing vector pT7-PTP-stl (Park et al., 2014), and pT7-CITFA2-stl (Brandenburg et al., 2007), along with accompanying cell lines, were described previously. smC2-PTP cells allow for the conditional silencing of CITFA2 through targeting the PTP tag coding sequence, and was generated in two steps: starting with a previously published smPTP cell line (Park et al., 2014), which conditionally expresses double stranded RNA targeting the PTP tag

sequence, we first used site-directed integration of SphI-linearized pPTP-CITFA2-PURO into one *CITFA2* allele to fuse the PTP tag sequence to the 5′ end of the *CITFA2* coding region. In a second step, the remaining *CITFA2* allele was replaced by a PCR product in which 100bp of *CITFA2* gene flanks surrounded the hygromycin phosphotransferase coding sequence. In order to conditionally express exogenous transgenes of *CITFA2*, pT7-trans was developed. We inserted into pT7-stl, using HindIII and XbaI restriction sites, a PCR product which contained (5′ to 3′) 490bp of the HSP70 genes 2 and 3 intergenic region, NdeI and NotI restriction sites, a hemagglutinin (HA) tag sequence ending with a stop codon, and 741bp of the β-/α-tubulin intergenic region. pT7-CITFA2-HA was generated from pT7-trans through insertion of the full coding sequence of *CITFA2* using NdeI and NotI restriction sites. pT7-NDel-HA was generated similarly, save that bases +4 through +31, corresponding to amino acids 2 through 10 of CITFA2 (PEVGTQVYW), were deleted by PCR. pT7-3Amut-HA was generated using a CITFA2 insert which had bases +16 through +24 (ACTCAAGTT, coding for amino acids TQV) replaced with GCCGCGGCA, which coded for three alanines. Transfection of these three plasmids, after linearization by EcoRV, into smC2-PTP cells generated cell lines smC2-PTP-CITFA2-HA, smC2-PTP-NDel-HA, and smC2-PTP-3Amut-HA. Transfection of these same plasmids into wild-type sm cells resulted in cell lines smC2-CITFA2-HA, smC2-NDel-HA, and smC2-3Amut-HA**.**

DNAs and recombinant protein. To generate recombinant proteins for pulldown assays and sucrose gradients, seven different recombinant protein expression plasmids were created. pCITFA2-PTH, which allowed for the expression of full-length wild-type recombinant (r)CITFA2, was produced by inserting the full coding sequence of CITFA2 into the expression vector pET100/D-TOPO (Invitrogen) using NdeI and NotI restriction sites. This resulted in rCITFA2 which had fused to its C-terminus a ProtC epitope, followed be a thrombin cleavage site,

and a 6xHis tag. Four additional plasmids were produced which only differed in the length of the integrated *CITFA2* sequence: pCITFA2-N-PTH contains sequence for only the N-terminal half of CITFA2 (bases +1 to +624), while pCITFA2-C-PTH codes for the C-terminal half (bases +621 to +1263). pCITFA2-N1/4-PTH, coding for the N-terminal quarter of CITFA2, contained bases +1 to +303, while pCITFA2-N2/4PTH, coding for the second quarter, contained bases +289 to +624. p3Amut-PTH, however, uniquely coded for mutated full-length rCITFA2 - the same triple-alanine mutation as detailed above for pT7-C2-3Amut. These vectors were transformed into BL21 *Escherichia coli*, and protein expression was induced for 15 min to 1 hour at 37°C by adding 1 mM IPTG. Shorter incubations were required for plasmids which included the C-terminus of CITFA2, as its expression appears to be toxic to *E. coli* (data not shown). Recombinant LC8 was generated by placing the entire LC8 coding sequence downstream of the glutathione S-transferase (GST) sequence in the pGEX-4T-2 vector (GE Healthcare) using BamHI and NotI restriction sites. Recombinant GST-LC8 was expressed in BL21 *E. coli* and purified by glutathione affinity chromatography (GE Healthcare) following the manufacturer's recommendations. Thrombin digest and elution was then performed, such that, in all assays, LC8 was used as an untagged, recombinant protein.

Generation of a purified anti-LC8 antibody. Immune serum against LC8 was generated by immunization of Sprague-Dawley rats with rGST-LC8, according to a standard protocol (Schimanski et al., 2006). rGST-LC8-specific antibodies were purified from serum by blot-immobolized antigen, as previously detailed (Park et al., 2014). In contrast to immune serum, the purified antibody did not detect a non-specific band present in *E. coli* that co-migrated with thrombin-digested LC8 (data not shown).

Protein analysis. Immunoblot detections were performed using polyclonal antibodies directed against CITFA2 (Brandenburg et al., 2007), CITFA6, CITFA7 (Nguyen et al., 2012), and TFIIB (Schimanski et al., 2006). The PTP tag was detected with a monoclonal anti-ProtC antibody (Roche), while HA-tagged proteins were detected with a rat monoclonal anti-HA antibody (Roche). Extract preparation and tandem affinity purifications of PTP-tagged CITFA2, CITFA4, and CITFA7 were conducted according to the standard protocol (Schimanski et al., 2005b). Crude bacterial lysates of rCITFA2-PTH-expressing BL21 *E. coli*, used in protein pulldowns with rLC8, were prepared as follows: after inducing protein expression with 1 mM IPTG IPTG for 30 min at 37°C, bacterial cultures were pelleted, and 0.2 g of the cell pellet was resuspended in 4 ml of HisTALON xTractor buffer (Clontech). 250 units of Benzonase (Sigma), 400 ng lysozyme, and 250 µl of a protease inhibitor solution, prepared by resuspending 1 tablet of protease inhibitor (Roche) in 1 ml H_20, was then added. Following a 10 min incubation, with shaking, at 4°C, the mixture was centrifuged at 3,200 g for 30 min at 4°C. The supernatant was taken and constitutes the crude bacterial lysate. For the pulldowns, rCITFA2-PTH was purified from 100 µl of the crude lysate using 20 µl equilibrated TALON metal affinity resin (Clontech), according to the manufacturer's specifications. A 40 µl binding reaction containing the CITFA2-PTH-conjugated TALON resin, 100 mM KCl, 20 mM potassium glutamate, 20 mM HEPES-KOH (pH 7.7), 3 mM $MgCl_2$, 100 ng/µl BSA, 150 mM sucrose, 2.5 µl of the protease inhibitor solution, and 100 ng rLC8 was incubated at 27°C for 1 hour, with shaking. The resin was washed 7 times with a buffer containing 400 mM KCl), 20 mM potassium glutamate, 20 mM HEPES-KOH (pH 7.7), 3 mM $MgCl_2$, 10 ng/µl BSA,, 0.5 mM dithiothreitol (DTT), and 0.1% Tween-20. Proteins were eluted using HisTALON Elution Buffer (Clontech), and investigated by immunoblotting. To analyze rCITFA2-rLC8 interactions by sedimentation, 500 ng purified rLC8 was mixed with 200µl of

rCITFA2-PTH-containing crude bacterial lysates, and incubated for 30 min at 27°C. The binding reaction was then loaded onto a 4 ml 10-40% linear sucrose gradient, ultracentrifuged, and fractionated from top to bottom, as previously described (Brandenburg et al., 2007). Co-immunoprecipitations (co-IPs) of CITFA2 were performed using trypanosome extract, as previously described (Nguyen et al., 2007). UV-crosslinking and electrophoretic mobility shift assays (EMSAs) of a radiolabeled BES promoter probe and purified CITFA, both visualized by autoradiography, were conducted as previously detailed (Brandenburg et al., 2007). Secondary structure analysis was carried out using PredictProtein (www.predictprotein.org/) (Rost et al., 2004).

RNA analysis. To analyze the effect of *LC8* silencing on transcription by RNA pol I, total RNA was prepared by the hot-phenol method, as described previously (Park et al., 2014). For the analysis of rRNA, total RNA was separated in Reliant precast 1.25% SeaKem Gold agarose RNA gels (Lonza), and rRNA was detected by ethidium bromide staining. For semi-quantitative RT-PCR analysis, total RNA was reverse transcribed with Superscript Reverse Transcriptase II (Invitrogen) using an oligo(dT) primer. Semiquantitative PCR was performed using cycle numbers that were empirically determined to be within the linear amplification range for each primer pair: 5′-GATAAGCTTACGCGTTTCAACATTGAGAAGGAT ATTGC-3′ and 5′-GATTCTAGACTCGAGTCTTTGACTCATCCGTGCTGG-3′ were used to amplify the LC8 coding sequence and 3′ UTR, while primers amplifying *α-tubulin* and *VSG2* were published previously (Nguyen et al., 2012).

Chromatin immunoprecipitation (ChIP). CITFA2-HA and CITFA3 promoter occupancy were analyzed by ChIP assays, as described previously (Park et al., 2014), using monoclonal rat anti-HA antibody (Roche) and purified polyclonal anti-CITFA3 antibody, respectively. Negative

control precipitations were carried out using affinity beads not bound to antibody. Chromatin was sonicated until fragments averaged 200-400 bp in length. The precipitated DNA was analyzed by qPCR using consensus primers for the slightly varying copies of *RRNA* and BES promoters, and a primer pair for the β-/α-tubulin intergenic region, which were specified previously (Nguyen et al., 2012; Park et al., 2014). The percent IP was calculated relative to the input material and corrected by subtracting the percent IP of the negative control ChIP. Each ChIP experiment was independently carried out 3-4 times, and statistical analyses were performed using percent IP averages. Comparisons between corrected percent precipitations of non-induced and induced cells were performed using Student's T-test. Prior to the application of the T-test, an F-test was performed to assure that the assumption of equal variance between groups was not violated. If the F-test indicated that the difference in variation was significant, then the T-test was performed with the more conservative assumption of unequal variance. An unpaired, two-tailed test was used in all cases.

Microscopy. To visualize changes in cell morphology and DNA content upon LC8 silencing, BFs were incubated with 4,6-diamidino2-phenylindol (DAPI) at a final concentration of 2 ng/µl for 45 min and imaged using a Zeiss AxioVert 200 microscope and Zeiss Axiovision 4.6.3.0 software, as described previously (Park et al., 2014). For quantification of the percentage of cells which were multiflagellated, multi-kinetoplastid, or multi-nucleated, cells were scored as abnormal if they met any one of the following criteria: ≥3 kinetoplasts, ≥3 nuclei, ≥3 flagella, or more nuclei than kinetoplasts. 150 cells were scored from both the non-induced and induced populations.

Flow Cytometry. In order to assess changes in the size and DNA content of a large number of cells upon *LC8* silencing, flow cytometry experiments were performed. Non-induced and 1 day

induced smLC8 cultures were stained with propidium iodide, and counted using an LSR II flow cytometer (BD Biosciences), according to a previously described protocol (Li et al., 2009). 30,000 cells were counted in each experiment, in triplicate, for each induction state, and ungated data was visualized using the FlowJo software package (Treestar Inc.).

II-3. Results

LC8 is essential for *T. brucei* viability, cytokinesis, and RNA pol I-mediated transcription. A survey of kinetoplastid genome data bases revealed conservation of two distinct *LC8* genes. One encodes an LC8 protein that is closely related to both human DYNLL1 and DYNLL2 (~82%/88% identity/similarity; accession number for the *T. brucei* gene is Tb927.11.1868 at www.TritrypDB.org (Aslett et al., 2010) or www.GeneDB.org (Logan-Klumpler et al., 2012)) and was previously termed DYNLL1 (Brandenburg et al., 2007). A second *LC8* is also present, yet this gene is more divergent from *LC8*s in other eukaryotes (~61%/85% identity/similarity to human DYNLL1 and 2; Tb927.11.320; **Figure II-S1**). Since a phylogenetic analysis did not reveal that these two LC8 proteins resemble the DYNLL1 and DYNLL2 dichotomy found in chordates (**Figure II-S1C** in the supplemental material), we propose to rename Tb927.11.1868 as *LC8* and Tb927.11.320 as *LC8dv* (dv for divergent). Trypsin-derived peptides of these two proteins are different except for a four amino acid-long peptide (**Figure II-S1A**). Since previous mass spectrometric analyses of purified CITFA consistently identified LC8 but never LC8dv-derived peptides (Brandenburg et al., 2007; Nguyen et al., 2012), it is highly unlikely that LC8dv is a subunit of the transcription factor complex. We therefore concentrated our analysis on LC8.

To investigate the importance of LC8 to BFs, we reduced its expression by RNA interference (RNAi), using a conditional gene silencing system which expressed double-stranded RNA

(dsRNA) targeting both the *LC8* coding and 3′ untranslated region (UTR) upon addition of doxycycline (Wirtz et al., 1999). In three different clonal cell lines, derived from smBF cells and termed smLC8, culture growth arrested within 1 day, and the majority of cells were lost by 2 days of induction (**Figure II-1A**). By 3 days of induction, no living cells could be identified by microscopic examination. RNA monitoring revealed that doxycycline induction resulted in a clear decline of *LC8* mRNA abundance after 1 day, while the levels of *LC8dv* and α-tubulin mRNA was unaffected, confirming that the knockdown was specific to *LC8* (**Figure II-1B**). To evaluate whether this rapid death phenotype would allow us to detect effects on transcription by RNA pol I, we measured levels of *VSG* mRNA derived from the active *VSG2* gene and of rRNA. While the *VSG2* mRNA level dropped considerably after 1 day of induction, the rRNA level decreased only modestly in the same experiment, possibly due to greater stability of rRNA (**Figure II-1B**). Immunoblot monitoring of *LC8*-silenced cells showed a specific reduction in LC8 protein, while CITFA6 and TFIIB, an RNA pol II-specific factor, were either increased or unchanged during this short period (**Figure II-1C**), confirming the specificity of the knockdown. These data demonstrated that LC8 is essential for trypanosome viability, and they indicated that LC8 is also important, though perhaps indirectly, for transcription by RNA pol I.

Since smBF cells and their derivatives have a doubling time of approximately 7 hours in our hands (**Figure II-1A**, -dox), propidium iodide staining, which allows for the quantification of the DNA content of individual cells, revealed a rapidly-progressing second phenotype. After only 1 day of induction, the per-cell DNA content of induced cells had approximately tripled vs non-induced cells, as revealed by flow cytometry of propidium iodide stained cultures (**Figure II-1D**).

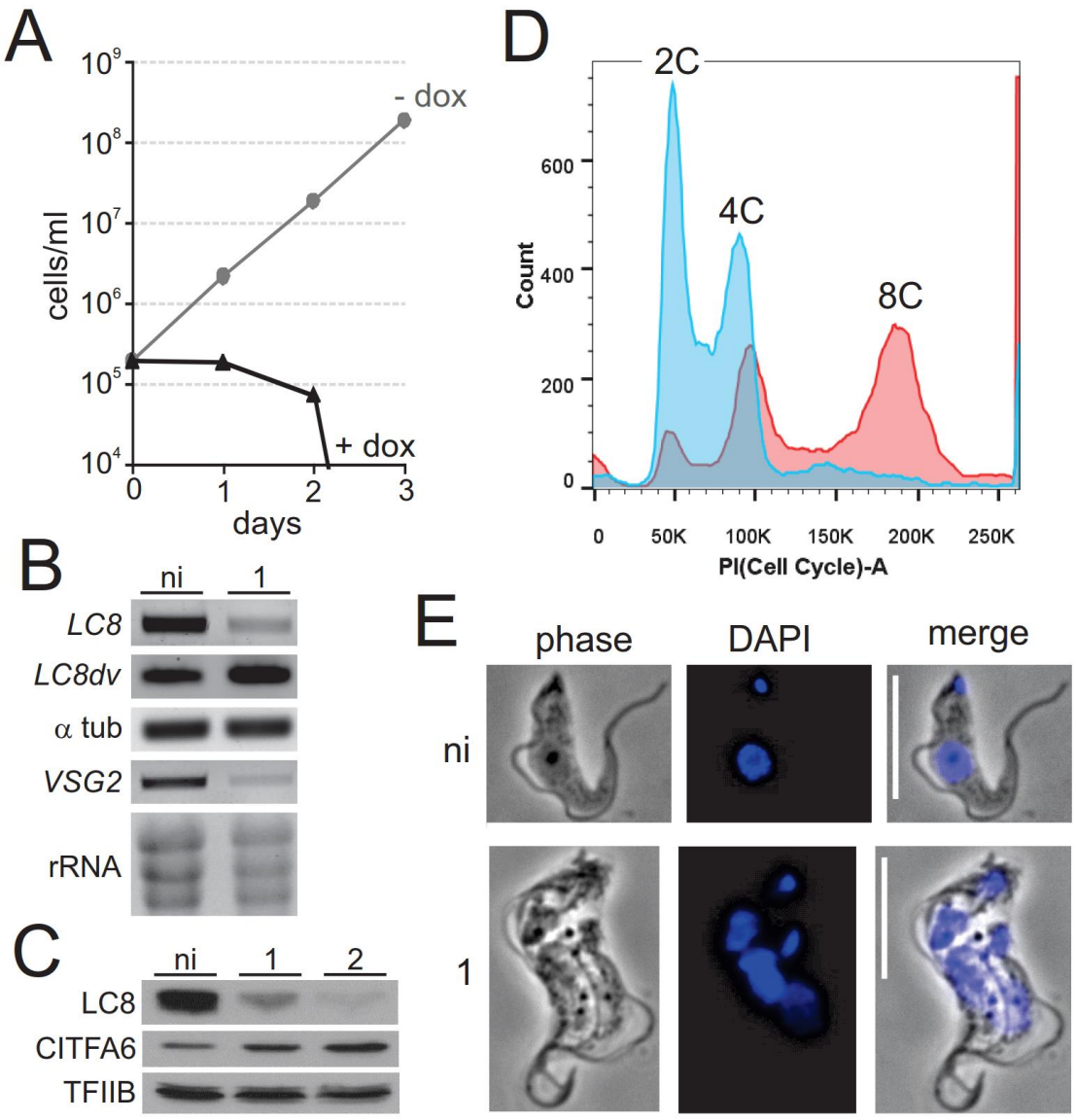


Figure II-1. *LC8* silencing has pleiotropic effects on BF trypanosomes. (**A**) Growth curve analysis of a representative clonal smLC8 BF cell line in the absence (- dox) and presence (+ dox) of the *LC8* knockdown inducing compound doxycycline. (**B**) Total RNA prepared from non-induced (ni) and 1 day induced cells was reverse transcribed with Oligo-dT and analyzed by semiquantitative PCR using *LC8*, *α-tubulin*, and *VSG2-specific oligonucleotides*. rRNA was detected by ethidium bromide staining of total RNA. (**C**) Whole cell lysates of non-induced and 1 or 2 day induced cells were analyzed by immunoblotting using LC8, CITFA6 or, as a loading control, TFIIB-specific

(**Figure II-1 legend cont.)** antibodies. (**D**) Flow cytometry analysis, without gating, of non-induced (blue) and 1 day induced cultures (red). Propidium iodide staining intensity (x-axis), which measures DNA content per cell, and cell counts (y-axis) from one representative experiment are shown. (**E**) Indirect fluorescence microscopy of non-induced and 1 day induced cells. Representative single cells from each culture were imaged using both phase contrast and DAPI fluorescence. Small and large areas of DAPI intensity represent kinetoplasts and nuclei, respectively. Scale bar is 10 μm.

Specifically, while non-induced cells demonstrated a curve with strong peaks representing cells with normal 2C and 4C DNA content, induced cultures had smaller 2C and 4C peaks, with the majority of cells displaying polyploidy (8C). This result was confirmed by DAPI fluorescence microscopy, which demonstrated that the majority of cells in culture (64%, n = 150) had increased in size, and were multi-flagellated, multi-kinetoplastid, and multi-nucleated (**Figure II-1E**), a phenotype that was observed in only two out of 150 non-induced cells. The increase in size was confirmed by flow cytometry of non-induced and *LC8*-silenced trypanosomes (**Figure II-S2**). This phenotype is not likely due to the loss of LC8 from the CITFA complex, as previously published knockdowns of other CITFA subunits failed to result in a similar phenotype (Brandenburg et al., 2007; Nguyen et al., 2014; Nguyen et al., 2012; Park et al., 2014). It would be consistent, however, with a role for LC8 in cell cycle progression, which has been reported previously in other organisms (Asthana et al., 2012; Liu et al., 2003).

LC8 binds to the N-terminus of CITFA2. In order to investigate the CITFA-specific role of LC8, we needed to determine LC8's binding partner within the CITFA complex, since disrupting this interaction would likely only interfere with LC8's RNA pol I-related function. While LC8 binding sites are conserved from yeast to humans (Wickstead and Gull, 2007), a survey did not unambiguously reveal such a site in CITFA subunit sequences (data not shown). We recently discovered in a sedimentation analysis that the peaks of CITFA7 and CITFA2 in extract differed slightly from each other, possibly indicating the existence of CITFA complexes with partially different compositions (Nguyen et al., 2012). To compare the sedimentation profile of LC8 with those of the other CITFA subunits, we tandem affinity-purified CITFA via PTP-tagged CITFA7 and sedimented the purified complex through a linear sucrose gradient (**Figure II-2A**). Fractionating the gradient from top to bottom and analyzing the protein content of each fraction

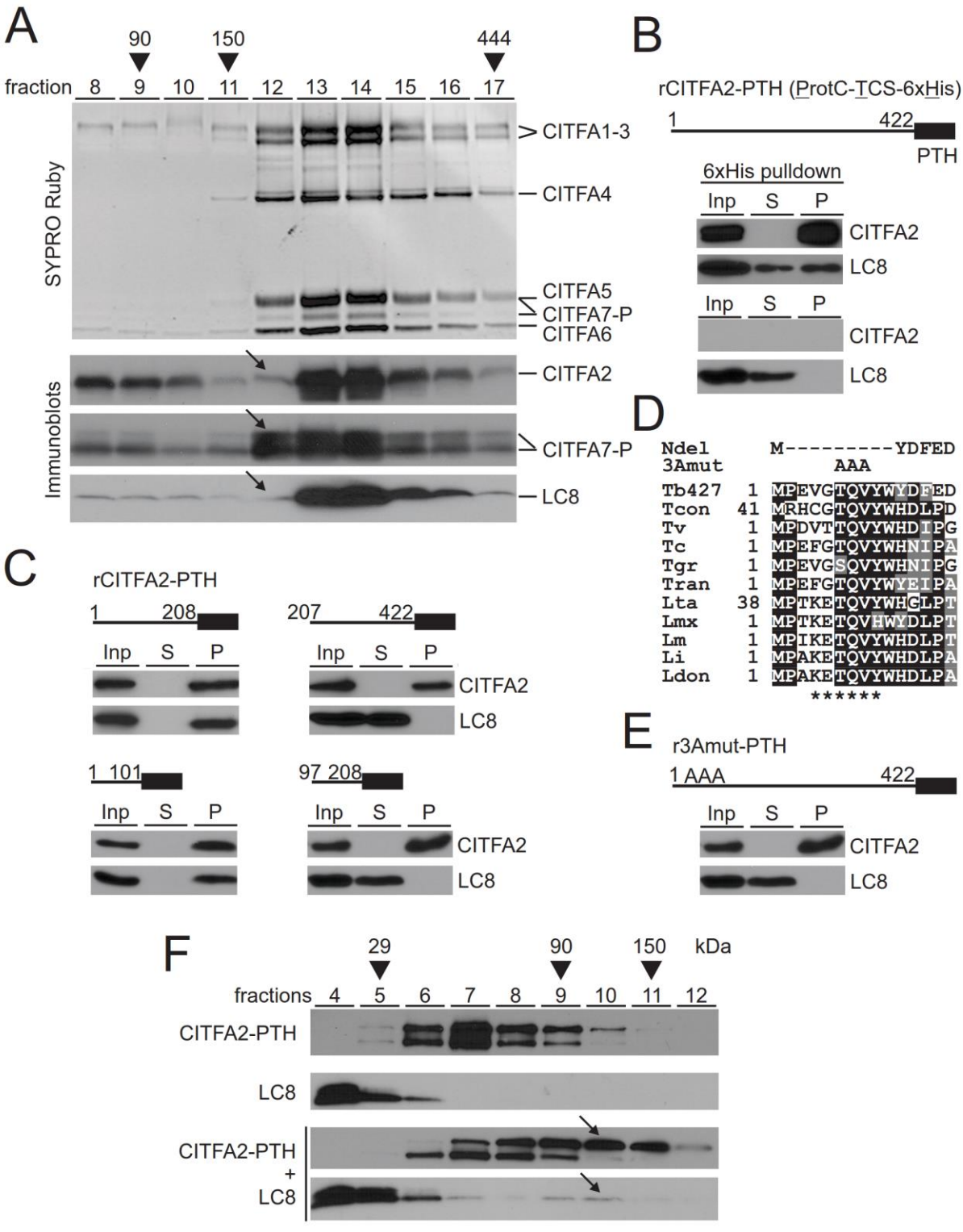


Figure II-2. LC8 binds to the N-terminus of CITFA2, promoting its dimerization. (**A**) Sedimentation of tandem affinity-purified CITFA by ultracentrifugation in a 10-40% linear

(**Figure II-2 legend cont.)** sucrose gradient. Fractions, taken from top to bottom, were separated by SDS-PAGE, stained with SYPRO Ruby and immunoblotted for specific detection of CITFA2, CITFA7, and LC8. CITFA7-P is so noted due to the presence of ProtC, which remains following TEV protease cleavage. Arrows highlight the relative absence of CITFA2 and LC8 in fraction 12. Taq DNA polymerase (95 kDa), IgG (150kDa), and apoferritin (AP) (444 kDa) were analyzed for molecular mass comparison (arrowheads). (**B**) Pulldown of recombinant, full-length (amino acid residues 1-422), wild-type CITFA2 with a C-terminal PTH tag, consisting of ProtC, a thrombin cleavage site (TCS), and the terminal 6xHis tag, in the presence of recombinant LC8. Crude extract (Inp), supernatant (S), and precipitate (P) were analyzed in relative amounts of 1:1:8 by immunoblotting, using ProtC- and LC8-specific antibodies. A negative control pulldown was conducted in the absence of CITFA2 (bottom panel). (**C**) Corresponding experiments with CITFA2-PTH fragments which are specified by residue numbers. (**D**) Alignment of the N-terminal CITFA2 sequences from *Trypanosoma brucei brucei* strain 427 (Tb427, accession numbers are listed in Table II-S1), *Trypanosoma congolense* (Tcon), *Trypanosoma vivax* (Tv), *Trypanosoma cruzi* (Tc), *Trypanosoma grayi* (Tgr), *Trypanosoma rangeli* (Tran), *Leishmania tarentolae* (Lta), *Leishmania mexicana* (Lmx), *Leishmania major* (Lm), *Leishmania infantum* (Li), and *Leishmania donovani* (Ldon). Positions with more than 50% identity or similarity are highlighted in black or gray, respectively. The proposed LC8 binding site is marked below the alignment (asterisks), while the two mutants used for further investigation, NDel and 3Amut, are indicated above. (**E**) Pulldown assay of full length rCITFA2-PTH carrying the 3Amut mutation (r3Amut-PTH). (**F**) Sucrose gradient sedimentation of rCITFA2-PTH alone, LC8 alone, or in combination, following co-incubation (bottom two panels). Arrows indicate a co-sedimentation peak in fraction 10 that is not present when either protein was analyzed on its own.

by SDS-PAGE and SYPRO Ruby staining revealed a CITFA complex that peaked in fractions 12-15 (**Figure II-2A**). The staining pattern and immunoblotting with specific antibodies, however, showed that while CITFA7 and other CITFA subunits were present in all four peak fractions, CITFA2 and LC8 were nearly absent from fraction 12, suggesting that CITFA2 may be the binding partner of LC8.

To directly test this, we generated in *E. coli* recombinant LC8 with an N-terminal GST tag and CITFA2 with a C-terminal composite PTH tag containing a ProtC epitope, a thrombin cleavage site, and six histidine residues (6xHis). Incubating purified CITFA2-PTH, immobilized on beads, with purified LC8, after removal of the GST tag, co-precipitated LC8 in a CITFA2-PTH-dependent manner, strongly indicating a direct interaction between these two proteins (**Figure II-2B**). In order to confirm this result, and better specify the site of interaction, we repeated this pulldown experiment using recombinant protein portions of CITFA2. Precipitation of the N-terminal half of CITFA2, comprising amino acids 1-208, effectively precipitated LC8 from solution, while precipitation of the C-terminal half did not (**Figure II-2C**). We further divided the N-terminal half of CITFA2, and again found that the most N-terminal portion, this time residues 1-101, precipitated LC8, while the second quarter of CITFA2 failed to precipitate LC8. We then attempted to identify the LC8 binding site within this reduced region of CITFA2 using motifs validated in other organisms (Rapali et al., 2011b), as well as a report which used a directed evolution approach to quantitatively determine the affinity preferences of the LC8 binding site (Rapali et al., 2011a). In both of these publications, the LC8 binding site almost always contained a central glutamine (Q), with threonine (T) or valine (V) in the -1 and +1 positions. Three amino acids, threonine-glutamine-valine, at the N-terminus of CITFA2 (amino acids 6-8) were identified as the likeliest site of interaction, and are almost completely conserved among kinetoplastid

CITFA2 sequences (**Figure II-2D**). In the course of this study two mutations of CITFA2 were pursued: an N-terminal deletion of amino acids 2-10 (NDel) and a replacement of amino acids 6-8 with alanines (3Amut). Precipitation of full length recombinant 3Amut-PTH failed to precipitate LC8 from solution, confirming that LC8 binds CITFA2 via this N-terminal sequence (**Figure II-2E**).

LC8 binding promotes CITFA2 dimerization. In other systems it was shown that LC8, acting as a dimer, binds to disordered regions of proteins, stabilizing their structure and allowing for areas present in the binding partner, such as coiled-coil domains, to promote dimerization, thereby resulting in the formation of a heterotetramer (Barbar and Nyarko, 2014, 2015; Rapali et al., 2011b). While secondary structure prediction software identified the N-terminus of CITFA2 as unstructured, no domains known to promote protein dimerization were recognized (data not shown). In order to investigate the possibility that LC8 binding induces dimerization of CITFA2, we performed sucrose gradient sedimentation with recombinant CITFA2-PTH and LC8. CITFA2-PTH, by itself, was found to peak in gradient fraction 7, which would be consistent with it existing as a monomer of 50 kDa (**Figure II-2F**), while the 10 kDa LC8 was found at the top of the gradient. When LC8 and CITFA2-PTH were allowed to interact before gradient sedimentation, however, both proteins exhibited a peak in fractions 9-10, consistent in size with a complex containing a 120 kDa CITFA2/LC8 heterotetramer. Serendipitously, when CITFA2-PTH was expressed in *E. coli*, immunoblots against the C-terminal ProtC tag detected its full length form and a truncated form which is missing approximately 5 kDa (**Figure II-2F**, asterisks). This putative N-terminal truncation was not shifted upon addition of LC8, indicating that LC8 cannot form a heterotetramer with the truncated CITFA2. These results strongly indicate that LC8 promotes the dimerization of CITFA2.

The CITFA2-LC8 interaction is essential for cell viability and RNA pol I-mediated transcription. In order to assess the importance of the CITFA2-LC8 interaction, we established BF cell lines in which doxycycline triggers both silencing of endogenous *CITFA2* and expression of an RNAi-resistant *CITFA2* transgene. In *T. brucei*, effective gene silencing via the RNAi pathway requires strong expression of ~500 bp-long dsRNA (Shi et al., 2000). We recently showed that targeting the heterologous PTP tag coding sequence effectively interfered with mRNAs carrying the PTP sequence, while having no deleterious off target effect in BFs (Park et al., 2014). The smPTP cell line, a derivative of the established smBF cell line for gene knockdowns (Wirtz et al., 1999), constitutively expresses the tetracycline (TET) repressor and T7 RNA pol, and, upon induction, PTP dsRNA from a TET-controlled T7 promoter (Park et al., 2014). To apply this system to *CITFA2*, we replaced one *CITFA2* allele in smPTP cells by hygromycin phosphotransferase and inserted the PTP-CITFA2-PURO plasmid into the second *CITFA2* allele to obtain cell line smC2-PTP (**Fig II-3A**, left panel). As expected, induction of PTP dsRNA led to a loss of cell viability over three days induction (**Figure II-3A**, middle panel), which matched previously performed *CITFA2* silencing experiments that targeted the CITFA2 coding sequence (Brandenburg et al., 2007). Immunoblotting of whole cell lysates confirmed that CITFA2 was exclusively expressed as a PTP fusion, and that depletion of PTP-CITFA2 was nearly complete after 1 day of induction (**Figure II-3A**, right panel). We next generated three *CITFA2* rescue constructs which differed only in respect to the LC8 binding site (**Figure II-3B**). The constructs contained the complete wild-type, 3Amut, or NDel coding region with an HA tag sequence at the 3′ end (CITFA2-HA), and were under the control of a TET-regulated T7 promoter (note that due to SL *trans* splicing, trypanosomes can utilize T7 pol for the effective production of functional mRNA (Stewart et al., 2010)). The constructs were transfected into smC2-PTP cells and targeted

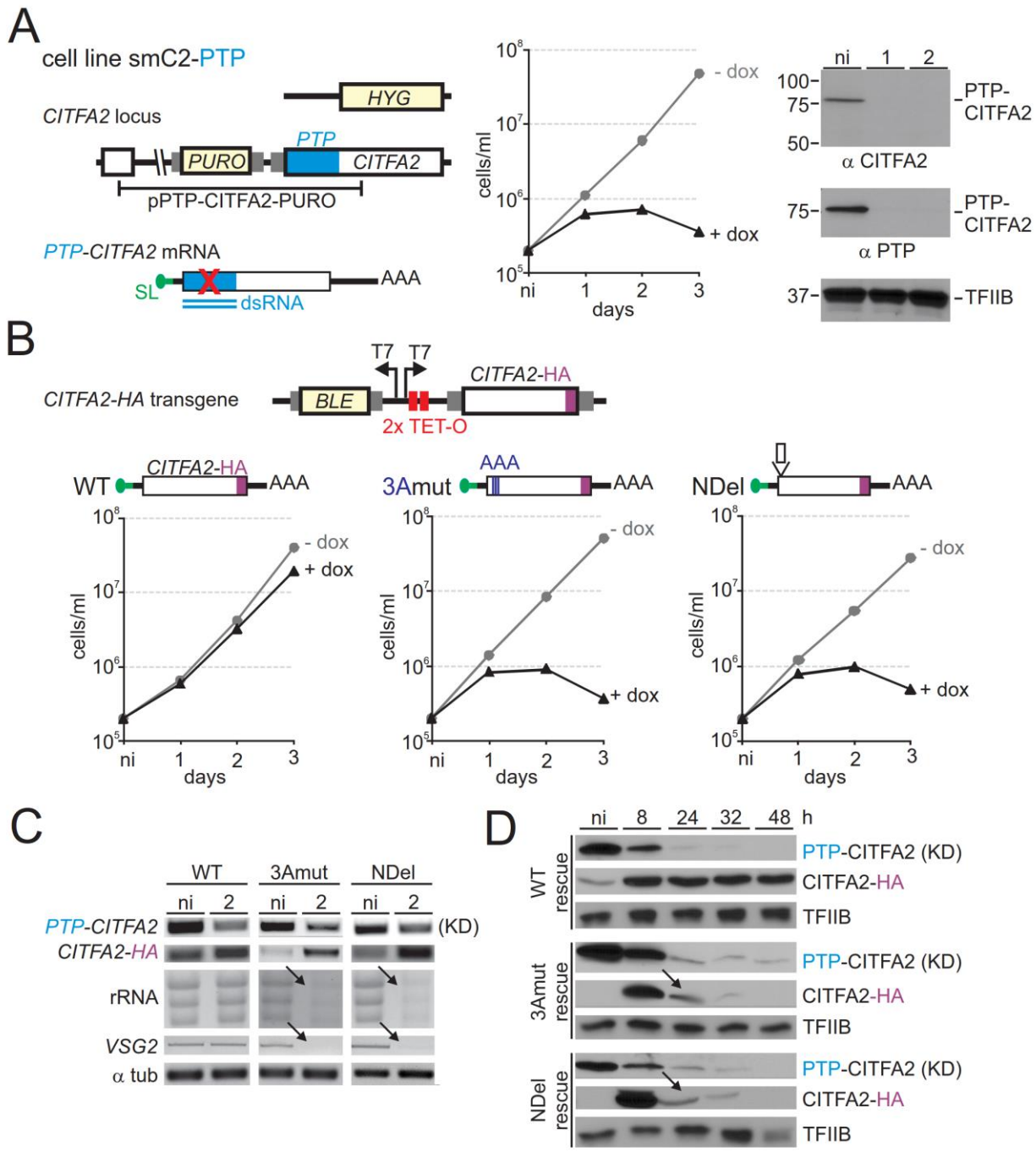


Figure II-3. Mutation of the LC8 binding site is lethal. (**A**) Left, schematic of the *CITFA2* locus (not to scale) in BF cell line smC2-PTP in which one allele has been replaced with a hygromycin resistance cassette (*HYG*), while the remaining allele has been fused to the PTP tag sequence by integration of pPTP-CITFA2-PURO, which harbors a puromycin resistance cassette (PURO).

(Figure II-3 legend cont.) Small gray rectangles indicate gene flanks with essential RNA processing signals. Doxycycline-induced expression of PTP dsRNA specifically targets *PTP-CITFA2* mRNA. Middle, smC2-PTP culture growth curve in the presence (+dox) and absence (-dox) of doxycycline. Right, immunoblot monitoring of the *CITFA2* knockdown using both anti-CITFA2 and anti-PTP antibodies, with TFIIB serving as a loading control. Note that the absence of a ~55 kDa band in anti-CITFA2 antibody probing confirms exclusive expression of PTP-tagged CITFA2 in smC2-PTP cells. (**B**) Top, schematic depiction (not to scale) of the construct that harbored the *CITFA2-HA* transgene and was targeted to the silent *RRNA* intergenic region to conditionally express RNAi-resistant CITFA2-HA mRNA and rescue the PTP-CITFA2 knockdown. Bottom, culture growth curves of representative smC2-PTP cell lines whose PTP-targeted CITFA2 knockdown was rescued with wild-type (WT), 3Amut, or NDel CITFA2-HA expression. (**C**) RNA analysis of the rescue cell lines after either no induction (ni) or two days of doxycycline. (**D**) Immunoblotting of PTP-CITFA2 and CITFA2-HA proteins during a time course of doxycyline-induced PTP-CITFA2 knockdown (KD) and CITFA2-HA expression, with TFIIB serving as a loading control. Arrows indicate co-declines of 3Amut and NDel CITFA2-HA protein with PTP-CITFA2 in the corresponding cell lines.

to the ribosomal spacer region, a silent genomic locus commonly used for integration of exogenous plasmid constructs. As expected, the expression of wild-type *CITFA2-HA* almost completely rescued for the growth defect which resulted from *PTP-CITFA2* silencing (**Figure II-3B**, left panel). In contrast, neither 3Amut nor NDel were able to rescue for the knockdown (**Figure II-3B**, middle and right panels). RNA analysis confirmed the reduction in *PTP-CITFA2* mRNA in all three cell lines, and the simultaneous expression of the HA-tagged transgenes (**Figure II-3C**). Furthermore, and consistent with the failed rescue of the *PTP-CITFA2* knockdown by both mutant *CITFA2-HA* genes, RNA pol I-derived transcripts rRNA and *VSG2* mRNA were strongly reduced in induced cells, while such defects were absent in the wild-type rescue line (**Figure II-3C**). Immunoblot monitoring confirmed the knockdown of PTP-CITFA2 after 24 hours of induction, and the expression of the three different rescue transgenes within 8 hours (**Figure II-3D**). The lethality of mutating the LC8 binding site in CITFA2 in conjunction with the specific decline of rRNA and *VSG2* transcripts strongly indicated that the interaction between LC8 and CITFA2 is essential for RNA pol I-mediated transcription and BF viability in culture.

Our next goal was to determine the specific defect in transcription upon interfering with the CITFA2-LC8 interaction. The 3Amut, and NDel rescue cell lines, however, were not informative in this regard, as immunoblots revealed that while both 3Amut and NDel mutant proteins were expressed upon induction, peak expression was transient, and protein levels fell sharply at 24 hours (**Figure II-3D**, arrows). This cannot be due to a decline in transgene mRNA, as it was found to be well expressed at 48 hours post induction (**Figure II-3C**). Given that the decline in mutant CITFA2 levels coincided with the RNAi-mediated decline of PTP-tagged CITFA2, we hypothesize that wild-type CITFA2 might be stabilizing mutant CITFA2 through weak direct interaction. This

would be consistent with the view that LC8 drives dimerization of proteins which already have a propensity towards dimerization (Barbar and Nyarko, 2014).

Mutant CITFA2 does not bind promoter DNA or interact with other CITFA subunits. We sought to obviate the problem of mutant instability by generating a cell line in which mutant CITFA2-HA and wild-type CITFA2 were expressed simultaneously. We transfected our same three *CITFA2-HA* transgene constructs into BF cells that contained two wild-type *CITFA2* alleles, generating cell lines C2-WT-HA, C2-NDel-HA, and C2-3Amut-HA. Immunoblot monitoring revealed that both wild-type and NDel CITFA2-HA were able to achieve long-term, high level expression in the presence of wild-type CITFA2, while 3Amut, though durably expressed, never reached an equal protein level (**Figure II-4A**). This expression profile was consistent in 2-4 cell lines obtained with each construct. The low expression of 3Amut could be due to a disruption in the secondary structure of CITFA2 causing instability, as the three alanine residues would strongly promote the formation of an α-helix (Pace and Scholtz, 1998) in a region predicted to form a β-strand.

Given that CITFA is a promoter-binding transcription factor, we sought to determine if the well expressed NDel mutant was recruited to promoter DNA. ChIP using anti-HA antibody, specific for the transgene-derived CITFA2-HA, revealed that while wild-type CITFA2-HA was present at both *RRNA* and *BES* promoters, NDel CITFA2-HA was not (**Figure II-4B**). This lack of promoter binding could be the result of either a loss of DNA-binding by a complete CITFA complex, or a lack of assembly of the mutant CITFA2 into the CITFA complex. To differentiate between these two possibilities, we immunoprecipitated both wild-type and NDel CITFA2-HA, and analyzed for co-IP of other CITFA subunits (**Figure II-4C**). Precipitation of the wild-type protein resulted in the co-IP of LC8, CITFA6, and CITFA7, verifying that the introduced CITFA2-

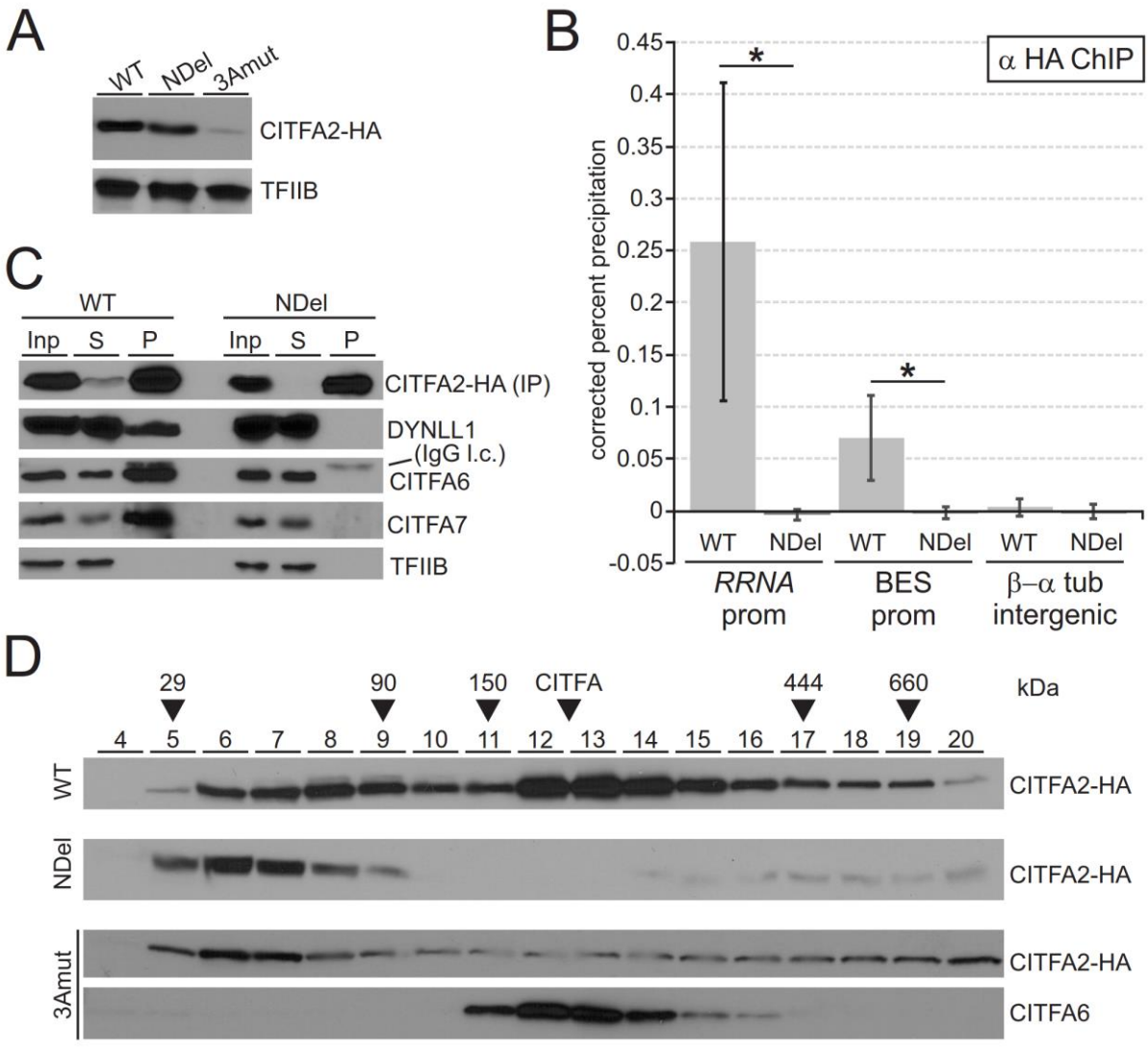


Figure II-4. Mutation of the LC8 binding site prevents recruitment of CITFA2 to promoters and its assembly into the CITFA complex. (**A**) Comparison of constitutive wild-type (WT), NDel, or 3Amut CITFA-HA expression in individual BF cell lines by immunoblotting, with TFIIB serving as a loading control. (**B**) Anti-HA ChIP assays in cell lines constitutively expressing wild-type or NDel CITFA2-HA. Precipitated DNA was analyzed using primer pairs which amplified the consensus BES promoter (BES prom), the consensus *RRNA* promoter (*RRNA* prom), and the β-/α-tubulin intergenic region. Error bars represent one standard deviation, with asterisks indicating a

(Figure II-4 legend cont.) Student's T-test *P*-value of < 0.05. (**C**) Anti-HA co-immunoprecipitation with the same cell lines. Blots monitoring wild-type (WT) or NDel CITFA2-HA immunoprecipitation were probed to detect the co-immunoprecipitation of LC8, CITFA6, CITFA7, and, as a loading control, TFIIB. Note that IgG light chain (IgG l.c.) was detected at the top of the CITFA6 immunoblot. (**D**) Sucrose gradient sedimentation of whole cell extracts, prepared from cell lines that constitutively express wild-type (WT), NDel, or 3Amut CITFA2-HA cell lines, were analyzed by immunoblotting fractions 4-20 using either anti-HA or anti-CITFA6 immune serum.

HA had assembled with other CITFA subunits. NDel CITFA2-HA, however, failed to co-IP any of these proteins, showing that the lack of promoter binding by the mutant was due to a lack of stable association with other CITFA subunits. To confirm this result, and further investigate the assembly status of both wild-type and mutant CITFA2, we performed sucrose gradient sedimentation of BF extract. While wild-type CITFA2-HA had its major sedimentation peak in fractions 12-15, which coincided with the peak of CITFA6, both 3Amut and NDel peaked in fractions 6-7, and lacked a peak in 12-15 (**Figure II-4D**). This result confirms that CITFA2 must bind LC8 for CITFA complex assembly. Given that sedimentation in fractions 6-7 would be consistent with a 50-kDa protein, it also appears likely that CITFA2 exists as a monomer in the absence of LC8 binding in our extracts. Note that wild-type CITFA2-HA has a minor peak in fractions 8-9, likely representing a CITFA2 dimer and/or a CITFA2-LC8 heterotetramer.

CITFA2 directly contacts BES promoter DNA and is required for promoter binding of CITFA *in vivo*. CITFA2 was shown to be of crucial importance to the initiation of transcription by RNA pol I from *RRNA*, BES and procyclin promoters, both *in vivo* and *in vitro* (Brandenburg et al., 2007). However, its specific function in the complex has not been determined. Early UV crosslinking of partially purified CITFA and radiolabeled BES promoter DNA resulted in a major labeled protein band of ~50 kDa (Brandenburg et al., 2007), which is the approximate size of CITFA subunits 1-3. Accordingly, depletion of CITFA1 in BFs caused a loss of CITFA3 occupancy of *RRNA* and BES promoters, indicating that CITFA1 is important for the transcription factor's ability to bind to RNA pol I promoters (Park et al., 2014). The BES promoter extends only to position -67 relative to the transcription initiation site (TIS) and harbors two distinct sequence elements (Pham et al., 1996; Vanhamme et al., 1995) both of which are required for efficient binding of CITFA to the BES promoter (Brandenburg et al., 2007). The two elements are separated

by 25 bp indicating that more than one CITFA subunit mediate DNA binding. Thus, to determine a potential DNA-binding role of CITFA2, we first used the sucrose gradient fractions of the purified CITFA complex (**Figure II-2A**), for a gel shift assay with an 82 bp-long radiolabeled BES promoter probe. Consistent with CITFA2 and LC8's near absence from fraction 12 and their sedimentation peak in fraction 13 and 14, fraction 12 shifted the promoter only faintly (**Figure II-5A**, arrow), while a strong shift signal was observed in fractions 13 and 14, indicating that CITFA2 and LC8 are important for the ability of CITFA to bind the BES promoter.

To better understand which CITFA subunits directly contact BES promoter DNA, we performed UV crosslinking using tandem affinity-purified CITFA from BFs that expressed PTP-CITFA7 (Nguyen et al., 2012). Crosslinking this high purity eluate and a radiolabeled BES promoter revealed two specific protein bands in the 50-60 kDa range and a third band of ~40 kDa (**Figure II-5**, lane 3). Since the latter band was likely CITFA4, given that no other CITFA subunit is 40 kDa, we repeated the experiment with a BF cell line that expressed CITFA4-PTP. After tandem affinity purification of a PTP-tagged protein, a ~4 kDa portion of the tag (ProtC) remains on the purified protein. Accordingly, in the CITFA4-PTP purification, the 40 kDa band shifted up (**Figure II-5B**, lane 2, CITFA4-P), unequivocally identifying CITFA4 as a direct contactor of promoter DNA. Likewise, using purified CITFA containing a PTP-tagged CITFA2 resulted in a shift of the uppermost crosslinked band, identifying CITFA2 as having direct DNA contact (**Figure II-5B**, lane 4, P-CITFA2). Since we have been unable to verify functional PTP tagging of CITFA1, we repeated this experiment with PTP-tagged CITFA3. A failure of tagged-CITFA3 to increase the size of any of the three bands (data not shown) makes it likely that CITFA1 is the third direct contactor of promoter DNA within the CITFA complex.

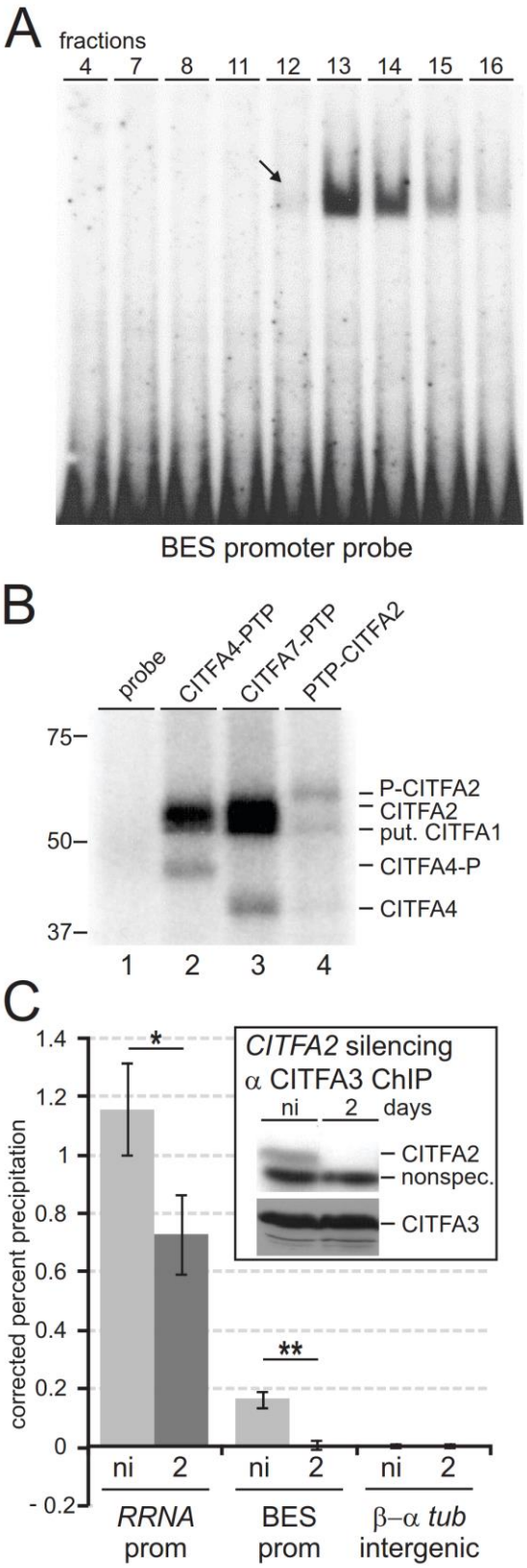


Figure II-5. CITFA2 directly contacts the BES promoter and is required for CITFA to bind to

(Figure II-5 legend cont.) RNA pol I promoters *in vivo*. (**A**) Sucrose gradient fractions of purified CITFA, shown in **Figure II-2A**, were used in an EMSA with a radiolabeled BES promoter that was visualized by autoradiography. Fraction 12, which contains minimal LC8 and CITFA and an abundance of other CITFA subunits, barely binds to the probe (arrow), while fractions 13-14, which contain an abundance of all CITFA subunits, effectively bound the promoter probe. (**B**) UV crosslinking analysis using tandem affinity-purified CITFA with radiolabeled BES promoter. After DNA digest, proteins were separated by SDS-PAGE and visualized by autoradiography. On the right, tagged and untagged CITFA subunits are identified. As explained in the text, the band that did not shift is putatively CITFA1 (put. CITFA1). (**C**) Anti-CITFA3 ChIP assay in a smCITFA2 cell line which was either not induced or in which PTP-CITFA2 was silenced for 2 days. One and two asterisks represent *P*-values <0.05, <0.01, respectively.

To confirm the importance of CITFA2 to CITFA promoter binding *in vivo*, we analyzed CITFA occupancy at *RRNA* and BES promoters in non-induced BFs and in BFs in which *CITFA2* was silenced for two days. Since CITFA2 depletion did not affect the abundance of CITFA3 (**Figure II-5C**, insert) and since absence of CITFA2 does not appear to disrupt the CITFA complex (see **Figure 2A**, fraction 12; Nguyen et al. 2012), we performed ChIP using a purified, ChIP-grade polyclonal anti-CITFA3 antibody (Park et al., 2014). Consistent with CITFA2 being a promoter binding protein, CITFA3 occupancy of BES promoters was completely lost upon CITFA2 depletion (**Figure II-5C**). Although the CITFA3 occupancy of *RRNA* promoters was significantly reduced in the same experiments, CITFA3 association with *RRNA* promoters remained substantial. This might be due to structural differences between *RRNA* and BES promoters. In contrast to the short BES promoter, the *RRNA* promoter extends to position -257 relative to the TIS, harboring four distinct promoter domains (Janz and Clayton, 1994; Schimanski et al., 2004). Thus, it is possible that additional factors present at the *RRNA* promoter interact with CITFA and stabilize it in the absence of CITFA2. In either case, these data clearly demonstrate that CITFA2 is important for CITFA binding to both *RRNA* and BES promoters.

LC8 depletion affects CITFA occupancy of *RRNA* and BES promoters. Finally, to verify that the recruitment failure of the CITFA2 NDel mutant to RNA pol I promoters is due to a loss of the CITFA2-LC8 interaction, we analyzed whether *LC8* silencing affected *CITFA* occupancy of *RRNA* and BES promoters. Using the same cell line as in **Figure II-1**, we found a highly significant reduction in binding of CITFA3 to the BES promoter, despite the fact that *LC8* silencing limited the analysis to one day of induction (**Figure II-6**). According to our observation that CITFA2 is less critical for CITFA3 occupancy of the *RRNA* promoter (see **Figure II-5C**), LC8 depletion affected *RRNA* promoter precipitation only modestly (**Figure II-6**). These results are consistent

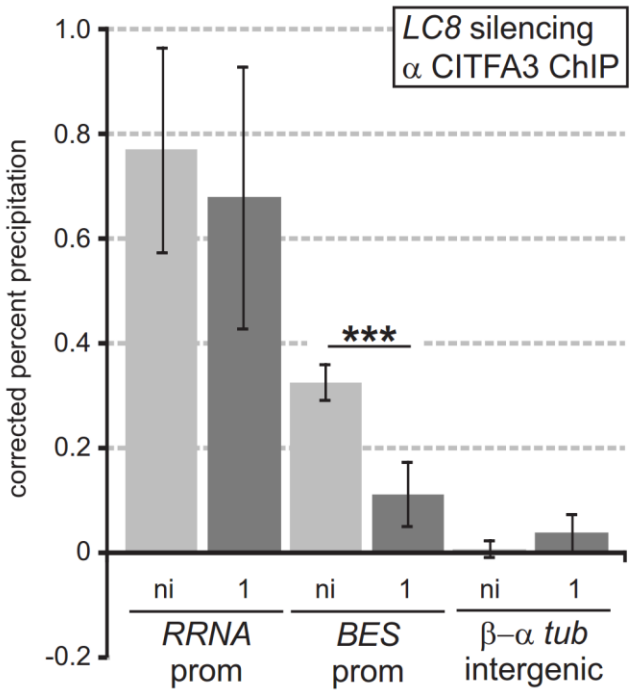


Figure II-6. LC8 is required for recruitment of CITFA to the BES promoter. Anti-CITFA3 ChIP in smLC8 BFs without induction (ni) and after 1 day of *LC8* silencing. Precipitated DNA was analyzed using the primer pairs previously noted. Three asterisks represent a *P*-value <0.001.

with the CITFA2-LC8 interaction being crucial for CITFA function and they verify LC8's important role in multifunctional RNA pol I transcription in *T. brucei*.

II-4. Discussion

Here we have shown that LC8 has at least two essential functions in *T. brucei*, namely in cell cycle progression and, as part of CITFA, in RNA pol I transcription, which was the focus of this investigation. We found that LC8 directly interacts with the N-terminus of CITFA2, requiring a conserved N-terminal TQV motif for binding. Sedimentation of recombinant CITFA2-LC8 complexes indicated that LC8 binding promotes the dimerization of CITFA2, resulting in a CITFA2-LC8 heterotetramer. Silencing of endogenous *CITFA2* in conjunction with the expression of RNAi-resistant wild-type and mutant *CITFA2* transgenes, revealed that the CITFA2-LC8 interaction is essential for trypanosome viability in culture, specifically affecting the abundance of RNA pol I transcripts. Pursuing the specific defect in the RNA pol I system, we found that CITFA2 is unable to bind promoter DNA or assemble with other CITFA subunits to form a complete CITFA complex in the absence of the CITFA2-LC8 interaction. Following up on a previous report which suggested a role for CITFA2 in promoter binding by CITFA (Brandenburg et al., 2007), we found that CITFA2, CITFA4 and, likely, CITFA1 directly contact promoter DNA. Accordingly, *CITFA2* silencing led to a defect in both *RRNA* and BES promoter binding of CITFA3, a result which was verified for the BES promoter by silencing *LC8*. These data suggest a model (**Figure II-7**) in which trypanosome LC8, by forming a dimer as in other organisms (Barbar et al., 2001; Benashski et al., 1997), binds to the N-termini of two CITFA2 molecules, promoting or stabilizing their dimerization. Since CITFA2 and LC8 remain stably associated with the CITFA complex after tandem affinity purification and sucrose gradient sedimentation, even at 400 mM KCl (**Figure II-**

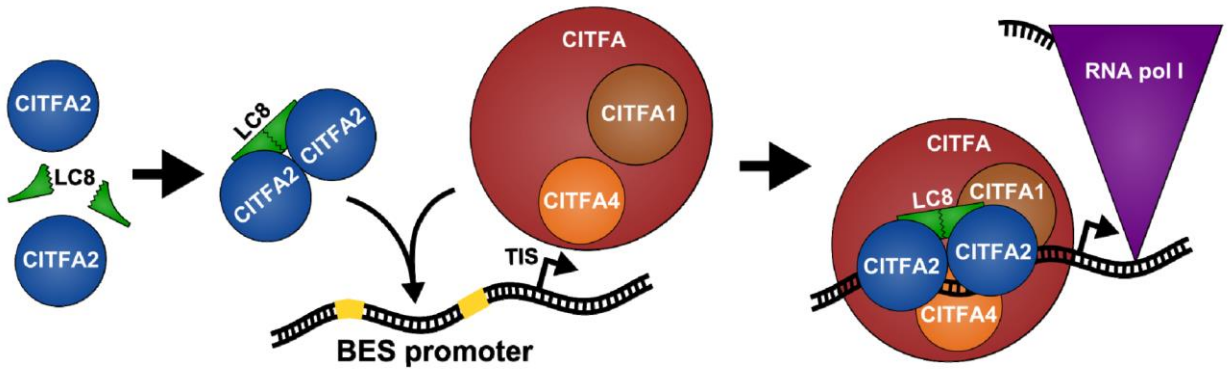


Figure II-7. Model of the CITFA2-LC8 interaction and function for BES promoter transcription. Formation of a CITFA2-LC8 heterotetramer is a prerequisite for its assembly into the CITFA complex. While a stable partial CITFA complex is formed without CITFA2-LC8, only a fully assembled CITFA complex that includes CITFA2 and LC8 is able to bind promoter DNA. The BES promoter, with its two sequence blocks essential for CITFA binding (yellow), is shown here along with the TIS. Please note that, for ease of visualization and understanding, LC8's site of interaction with CITFA2 was not shown at its dimer interface, where it actually occurs. Promoter-bound CITFA is able to, either directly or indirectly, recruit RNA pol I and enable transcription initiation.

2A; (Brandenburg et al., 2007; Nguyen et al., 2012); data not shown), it is likely that the full assembly of CITFA occurs independent of DNA. Finally, by contacting DNA directly through its CITFA2, CITFA4 and, likely, CITFA1 subunits, the transcription factor complex binds with high affinity to both elements of the BES promoter which, in turn, leads to RNA pol I recruitment and transcription initiation. Although we cannot exclude the possibility that CITFA2 dimerization leads to dimerization of the whole CITFA complex, the sedimentation profile of purified CITFA argues against this possibility. Consistent with an overall mass of CITFA of 323 kDa, which assumes a CITFA2-LC8 heterotetramer and monomers for all other subunits, full CITFA sedimented in-between the 150 and 444 kDa size markers (**Figure II-2A**).

Our data and model are consistent with the view that LC8, rather than functioning as a linker between two different proteins, instead promotes homodimerization, thereby imparting new function (Barbar, 2008; Barbar and Nyarko, 2014, 2015; King, 2008). Is LC8 essential for CITFA2 dimerization or does it rather stabilize a CITFA2 dimer by interacting with its predicted disordered N-terminus? Sedimentation analysis of recombinant CITFA2 and LC8 proteins (**Figure II-2F**) and of extracts containing NDel and 3Amut CITFA2-HA protein (**Figure II-4D**) suggested that in the absence of LC8, CITFA2 cannot form a dimer. However, when endogenous, wild-type CITFA2 was depleted in trypanosomes, then NDel and 3Amut CITFA2 were co-lost (**Figure II-3D**) whereas the same proteins could be constitutively expressed in the presence of wild-type CITFA2 (**Figure II-4B**). This stabilization of mutant CITFA2 by the wild-type protein suggests some degree of CITFA2 interaction in the absence of LC8. This finding is consistent with the demonstrated function of LC8 in structuring and stabilizing dimers of the human dynein intermediate chain (Nyarko et al., 2004; Williams et al., 2007); reviewed in (Barbar, 2008; Barbar and Nyarko, 2014), the *Drosophila* RNA-binding protein swallow (Kidane et al., 2013; Wang et

al., 2004), the human motor protein myosin Va (Hodi et al., 2006; Wagner et al., 2006), and the yeast nucleoporin Nup159 (Nyarko et al., 2013; Stelter et al., 2007).

Given that some of LC8's non-dynein binding partners, such as the human transcriptional repressor TRPS1 (Kaiser et al., 2003), estrogen receptor (Rayala et al., 2005; Rayala et al., 2006), and the NF-κB inhibitor IκBα (Jung et al., 2008) are involved in transcriptional regulation, and since proper B cell maturation was recently demonstrated to directly depend on the DYNLL1 expression level (Jurado et al., 2012), it is worth considering if LC8 might be regulating multifunctional RNA pol I transcription in *T. brucei*. In accordance with this notion, our results indicated that depletion of CITFA2 (**Figure II-5C**) or LC8 (**Figure II-6**) affected CITFA3 occupancy of the BES promoter more than that of the *RRNA* promoter, indicating that LC8 may have a specific role in BES transcription. However, CITFA2 was shown *in vitro* and *in vivo* to be essential for RNA pol I transcription initiating at *RRNA*, procyclin gene and BES promoters (Brandenburg et al., 2007). In addition, the LC8 binding site within CITFA2 is at least partially conserved among all kinetoplastids (**Figure II-2D**), including those which do not display antigenic variation and are not known to utilize RNA pol I for pre-mRNA synthesis, suggesting that if LC8 does have a regulatory role for RNA pol I transcription, it is likely a general one. Nevertheless, the CITFA2-LC8 heterotetramer appears to have a crucial role in activating the CITFA complex. As previously determined, CITFA2 in BFs is expressed at about a fivefold lower level than CITFA7, indicating that the majority of CITFA complexes in a trypanosome are inactive, requiring binding of the CITFA2-LC8 heterotetramer for activation (Brandenburg et al., 2007; Nguyen et al., 2012). Thus, it is possible that formation of a productive CITFA2-LC8 heterotetramer is quantitatively controlled in the cell. It has been shown in other systems that phosphorylation of either LC8 or its binding partner can alter their interaction, ranging from minor changes in kinetics to a complete

elimination of binding (Benison et al., 2009; Gallego et al., 2013; Lei and Davis, 2003; Song et al., 2008; Song et al., 2007). This regulatory phosphorylation occurs at S88 in human LC8, which is conserved as S89 in the genus *Trypanosoma* (**Figure II-S1B**), allowing for such a regulation to exist. Furthermore, phosphorylation of CITFA2 at amino acid T6, one of the conserved residues in the LC8 binding motif, could also be used as a means to block LC8 binding, as is the case in Nek9, a kinase involved in mitotic progression (Gallego et al., 2013).

Moreover, recent evidence suggests that CITFA is directly involved in the regulation of monoallelic *VSG* expression, which takes place outside the nucleolus in the so-called expression site body or ESB (Chaves et al., 1999; Navarro and Gull, 2001). ChIP assays consistently demonstrated that CITFA occupancy of the active BES promoter was several times higher than that of a silent BES promoter (Nguyen et al., 2014). Furthermore, the CITFA complex remained localized to the nucleolus and the ESB after depletion of CITFA1, which caused CITFA to dissociate from *RRNA* and BES promoters (Park et al., 2014). These results raised the possibility that sequestration of CITFA to the nucleolus and the ESB may be the trypanosomes' means to restrict productive RNA pol I transcription to these compartments (Günzl et al., 2015). Since LC8 has been implicated in sub-nuclear and sub-cellular localization of many of its binding partners (Barbar and Nyarko, 2014), sequestration or spatially controlled formation of CITFA2-LC8 may be a mechanism for localizing CITFA function.

Although, an initial survey of CITFA sequences did not reveal a clear LC8 binding motif as originally defined in humans and yeast (Lo et al., 2001), the reasonable conservation of the LC8 binding site within CITFA2 to the now more broadly defined consensus motifs (Rapali et al., 2011a; Rapali et al., 2011b) suggests that it may be possible to bioinformatically identify other

kinetoplastid proteins that interact with LC8, shedding light onto additional roles of LC8 in this divergent group of organisms.

Furthermore, our results add *T. brucei* to the list of pathogens that rely on LC8 for viability. *LC8* silencing caused defects in both transcription and cytokinesis, and resulted in an extremely rapid death phenotype of cultured BF *T. brucei* that is rarely observed in this system. Although, to our knowledge, small molecule inhibition of LC8-ligand interactions have not been pursued so far, this may be a worthwhile anti-pathogenic strategy since *LC8* knockdown in adult animals appears to not be lethal (Wang et al., 2014).

In summary, this work is the first investigation of LC8 in any kinetoplastid organism, and reveals a novel use of LC8 in the basal process of transcription initiation by RNA pol I. It confirms the results from studies in other organisms regarding LC8's binding motif and its role in protein dimerization, demonstrating that this function of LC8 is of ancient evolutionary origin.

Acknowledgments

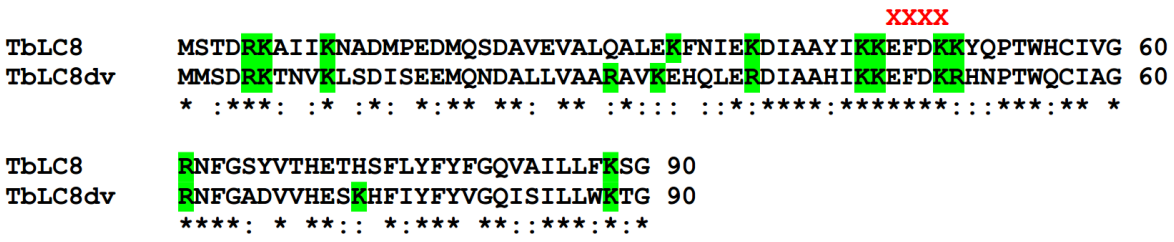
We would like to express our thanks to Dr. Evan Jellison, of the UCHC Health Flow Cytometry Core, for his help in processing our samples, to Dr. Ziyin Li, of the University of Texas, who provided us with the flow cytometry cell preparation protocol, to Dr. James Grady, of CICATS, for his insights on ChIP statistics, and to Dr. Stephen King, of the University of Connecticut Health Center, for his careful reading of this manuscript. This work was supported by grants R01 AI059377 to A.G. and F30 AI110060 to J.K.K. from the National Institute of Allergy and Infectious Diseases (NIAID) of the U.S. National Institutes of Health (NIH).

II-5. Supplemental

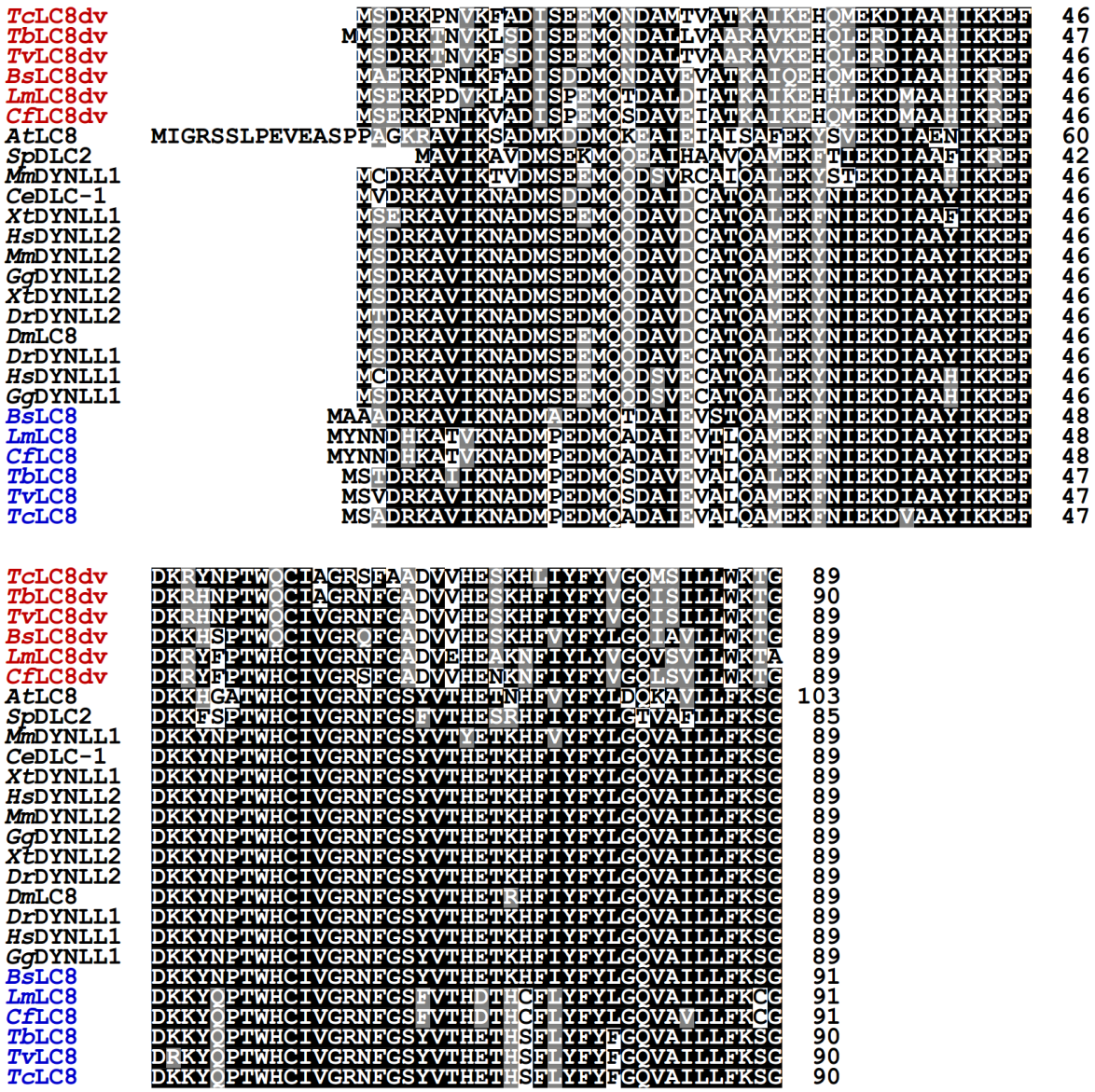
Trypanosoma brucei brucei strain 427	Tb427tmp.211.3440
Trypanosoma congolense	Tcon, TcIL3000_9_5170
Trypanosoma vivax	TvY486_0905960)
Trypanosoma cruzi	TcCLB.510741.100
Trypanosoma grayi	Tgr.1145.1010
Trypanosoma rangeli	TRSC58_05089
Leishmania tarentolae	LtaP35.3170)
Leishmania mexicana	LmxM.34.3150
Leishmania major	LmjF.35.3150
Leishmania infantum	LinJ.35.3200
Leishmania donovani	LdBPK_353200.1

Table II-S1. Accession numbers of kinetoplastid *CITFA2* genes

A


XXXX
TbLC8 MSTDRKAIIKNADMPEDMQSDAVEVALQALEKFNIEKDIAAYIKKEFDKKYQPTWHCIVG 60
TbLC8dv MMSDRKTNVKLSDISEEMQNDALLVAARAVKEHQLERDIAAHIKKEFDKRHNPTWQCIAG 60
RNFGSYVTHETHSFLYFYFGQVAILLFKSG 90
RNFGADVVHESKHFIYFYVGQISILLWKTG 90

B


TcLC8dv MSDRKPNVKFADISEEMQNDAMTVATKAIKEHQMEKDIAAHIKKEF 46
TbLC8dv MMSDRKTNVKLSDISEEMQNDALLVAARAVKEHQLERDIAAHIKKEF 47
TvLC8dv MSDRKTNVKFSDISEEMQNDALTVAARAVKEHQLERDIAAHIKKEF 46
BsLC8dv MAERKPNIKFADISDDMQNDAVEVATKAIQEHQMEKDIAAHIKREF 46
LmLC8dv MSERKPDVKLADISPEMQTDALDIATKAIKEHHLEKDMAAHIKREF 46
CfLC8dv MSERKPNIKVADISPEMQSDAVEIATKAIKEHQMEKDMAAHIKREF 46
AtLC8 MIGRSSLPEVEASPPAGKRAVIKSADMKDDMQKEAIEIAISAFEKYSVEKDIAENIKKEF 60
SpDLC2 MAVIKAVDMSEKMQQEAIHAAVQAMEKFTIEKDIAAFIKREF 42
MmDYNLL1 MCDRKAVIKTVDMSEEMQQDSVRCAIQALEKYSTEKDIAAHIKKEF 46
CeDLC-1 MVDRKAVIKNADMSDDMQQDAIDCATQALEKYNIEKDIAAYIKKEF 46
XtDYNLL1 MSERKAVIKNADMSEEMQQDAVDCATQALEKFNIEKDIAAFIKKEF 46
HsDYNLL2 MSDRKAVIKNADMSEDMQQDAVDCATQAMEKYNIEKDIAAYIKKEF 46
MmDYNLL2 MSDRKAVIKNADMSEDMQQDAVDCATQAMEKYNIEKDIAAYIKKEF 46
GgDYNLL2 MSDRKAVIKNADMSEDMQQDAVDCATQAMEKYNIEKDIAAYIKKEF 46
XtDYNLL2 MSDRKAVIKNADMSEDMQQDAVDCATQAMEKYNIEKDIAAYIKKEF 46
DrDYNLL2 MTDRKAVIKNADMSEDMQQDAVDCATQAMEKYNIEKDIAAYIKKEF 46
DmLC8 MSDRKAVIKNADMSEEMQQDAVDCATQALEKYNIEKDIAAYIKKEF 46
DrDYNLL1 MSDRKAVIKNADMSEEMQQDAVECATQALEKYNIEKDIAAYIKKEF 46
HsDYNLL1 MCDRKAVIKNADMSEEMQQDSVECATQALEKYNIEKDIAAHIKKEF 46
GgDYNLL1 MSDRKAVIKNADMSEEMQQDSVECATQALEKYNIEKDIAAHIKKEF 46
BsLC8 MAAADRKAVIKNADMAEDMQTDAIEVSTQAMEKFNIEKDIAAYIKKEF 48
LmLC8 MYNNDHKATVKNADMPEDMQADAIEVTLQAMEKFNIEKDIAAYIKKEF 48
CfLC8 MYNNDHKATVKNADMPEDMQADAIEVTLQAMEKFNIEKDIAAYIKKEF 48
TbLC8 MSTDRKAIIKNADMPEDMQSDAVEVALQALEKFNIEKDIAAYIKKEF 47
TvLC8 MSVDRKAVIKNADMPEDMQSDAIEVALQAMEKFNIEKDIAAYIKKEF 47
TcLC8 MSADRKAVIKNADMPEDMQADAIEVALQAMEKFNIEKDVAAYIKKEF 47
TcLC8dv DKRYNPTWQCIAGRSFGADVVHESKHLIYFYVGQMSILLWKTG 89
TbLC8dv DKRHNPTWQCIAGRNFGADVVHESKHFIYFYVGQISILLWKTG 90
TvLC8dv DKRHNPTWQCIVGRNFGADVVHESKHFIYFYVGQISILLWKTG 89
BsLC8dv DKKHSPTWQCIVGRQFGADVVHESKHFVYFYLGQIAVLLWKTG 89
LmLC8dv DKRYFPTWHCIVGRNFGADVEHEAKNFIYLYVGQVSVLLWKTA 89
CfLC8dv DKRYFPTWHCIVGRSFGADVVHENKNFIYFYVGQLSVLLWKTG 89
AtLC8 DKKHGATWHCIVGRNFGSYVTHETNHFVYFYLDQKAVLLFKSG 103
SpDLC2 DKKFSPTWHCIVGRNFGSFVTHESRHFIYFYLGTVAFLLFKSG 85
MmDYNLL1 DKKYNPTWHCIVGRNFGSYVTYETKHFVYFYLGQVAILLFKSG 89
CeDLC-1 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
XtDYNLL1 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
HsDYNLL2 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
MmDYNLL2 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
GgDYNLL2 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
XtDYNLL2 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
DrDYNLL2 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
DmLC8 DKKYNPTWHCIVGRNFGSYVTHETRHFIYFYLGQVAILLFKSG 89
DrDYNLL1 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
HsDYNLL1 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
GgDYNLL1 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 89
BsLC8 DKKYNPTWHCIVGRNFGSYVTHETKHFIYFYLGQVAILLFKSG 91
LmLC8 DKKYQPTWHCIVGRNFGSFVTHDTHCFLYFYLGQVAILLFKCG 91
CfLC8 DKKYQPTWHCIVGRNFGSFVTHDTHCFLYFYLGQVAVLLFKCG 91
TbLC8 DKKYQPTWHCIVGRNFGSYVTHETHSFLYFYFGQVAILLFKSG 90
TvLC8 DRKYQPTWHCIVGRNFGSYVTHETHSFLYFYFGQVAILLFKSG 90
TcLC8 DKKYQPTWHCIVGRNFGSYVTHETHSFLYFYFGQVAILLFKSG 90

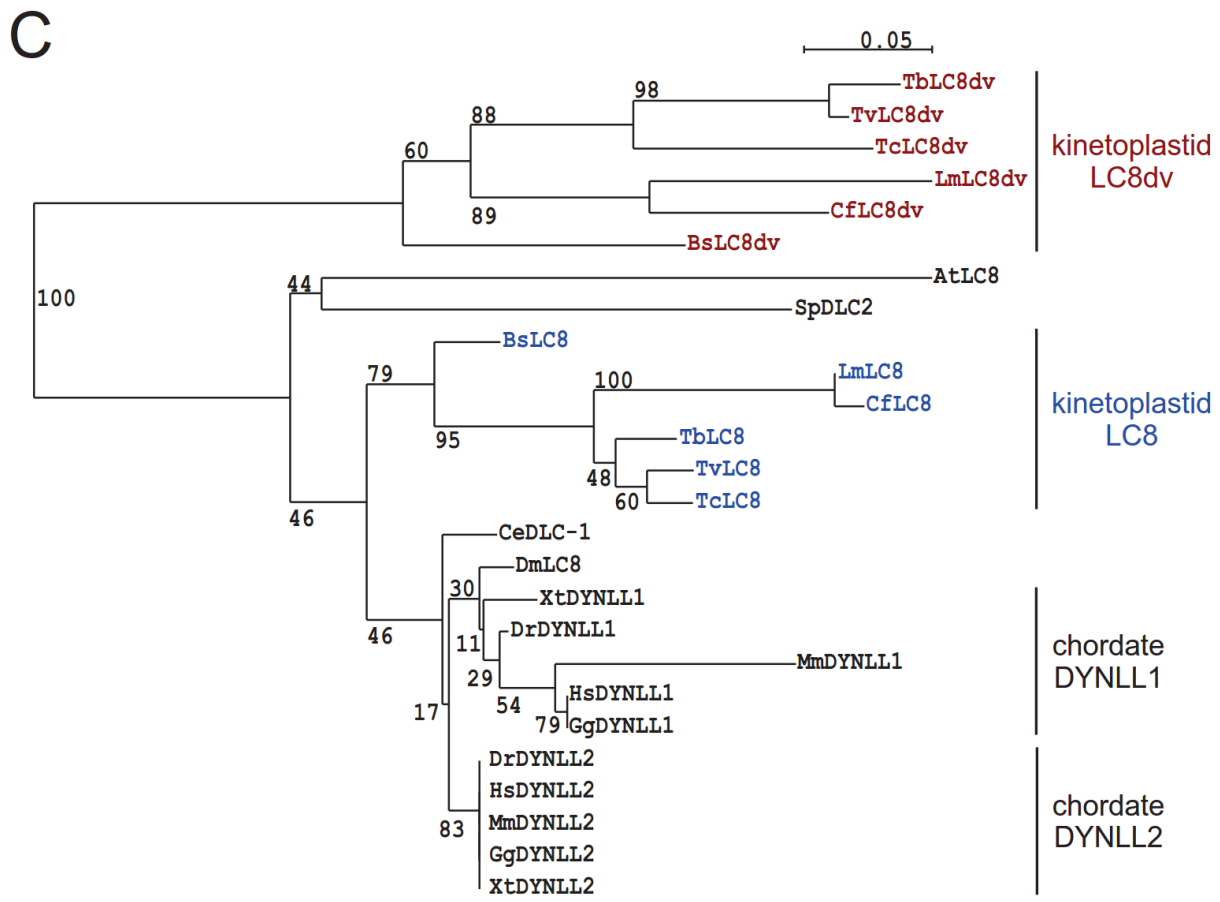


Figure II-S1. Kinetoplastids harbor two distinct, conserved *LC8* genes. (**A**) Clustal Omega alignment of amino acid sequences (Sievers et al., 2011) deduced from *TbLC8* (accession number Tb927.11.18680) and *TbLC8dv* (Tb927.11.320) coding regions. Identical and similar positions are indicated by asterisks and colons, respectively. Arginines and lysines, marking trypsin cleavage sites, are highlighted in green. The short common trypsin-derived peptide is marked by red Xs. (**B**) Multiple sequence alignment, carried out with the Clustal Omega server of the European Bioinformatics Institute (http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=clustalo) at default parameters, comprising DYNLL1 amino acid sequences from *Homo sapiens* (*Hs*DYNLL1, accession number NP_001032584), *Mus musculus* (*Mm*DYNLL1, NP_001001185), *Gallus gallus* (*Gg*DYNLL1, XP_003642263), *Xenopus tropicalis* (*Xt*DYNLL1, NP_001005077) and *Danio*

(**Figure II-S1 legend cont.)** *rerio* (*Dr*DYNLL1, NP_998189), of DYNLL2 from the same organisms (*Hs*DYNLL2, NP_542408; *Mm*DYNLL2, NP_080832; *Gg*DYNLL2, XP_004946822; *Xt*DYNLL2, NP_001165079; *Dr*DYNLL2, NP_956393), LC8 sequences from *Drosophila melanogaster* (*Dm*LC8, NP_525075), *Caenorhabditis elegans* (*Ce*DLC-1, NP_498422), *Schizosaccharomyces pombe* (*Sp*DLC2, NP_594368), *Arabidopsis thaliana* (*At*LC8, CAB46031) and from the kinetoplastids *T. brucei* (*Tb*LC8), *Trypanosoma vivax* (*Tv*LC8, TvY486_1100540 & *Tv*Y486_1100570), *Trypanosoma cruzi* (*Tc*LC8, TCDM_13942), *Leishmania major* (*Lm*LC8, LmjF.32.0230), *Crithidia fasciculata* (*Cf*LC8, CfaC1_32_0390) and *Bodo saltans* (*Bs*LC8, BS21670.1..pep & BS74770.1..pep), and divergent LC8 sequences from the same kinetoplastid organisms (*Tb*LC8dv; *Tv*LC8dv, TvY486_0034050; *Tc*LC8dv, TcCLB.504109.24; *Lm*LC8dv, LmjF.25.0260; *Cf*LC8dv, CfaC1_28_0460; *Bs*LC8dv, BS22550.1..pep). Positions with more than 50% identity or similarity are highlighted in black or gray, respectively. (**C**) Phylogenetic Tree of the shown sequence alignment using the BIONJ neighbor-joining algorithm (Gascuel, 1997) with the Seaview version 4 software package (Gouy et al., 2010). Bootstrapping was performed with 1000 replicates with values representing percentages.

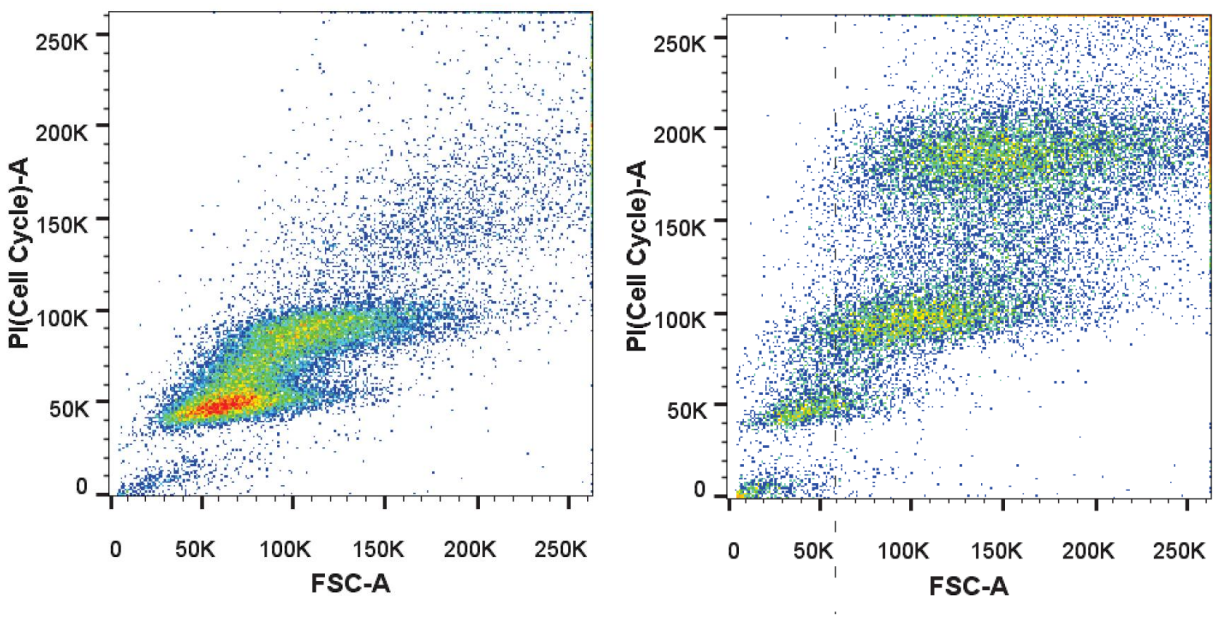


Figure II-S2. LC8 silencing results in an increase in both cell size and DNA content. Ungated count data from one of three replicate experiments comparing non-induced (left) to 1 day induced (right) LC8 knockdown cells. The y-axis represents the per-cell DNA content, as measured by propidium iodide staining, while the x-axis represents the forward scatter area, or size, of the cells. Note the appearance of a third population of cells in the induced culture which exhibits an increase in both size and DNA content. Blue represents areas of low count density, while green, yellow, and red represent increasing densities of cells with a given value.

Chapter III

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*

Abstract

Conditional gene silencing by RNA interference in *Trypanosoma brucei* can be inconclusive if knockdowns are inefficient or have off-target effects. To enable efficient, specific silencing of single copy genes in mammalian-infective, bloodstream form trypanosomes, we developed a system that targets the heterologous and functional *Trypanosoma cruzi U2AF35* 3′ UTR (Tc3) or, alternatively, the sequence of the PTP tag which can be fused to any mRNA of interest. Two cell lines were created, single marker (sm)Tc3 and smPTP, which conditionally express Tc3 and PTP dsRNA, respectively. The system depends on manipulating both alleles of the gene of interest such that cells exclusively express the target mRNA as a fusion to one of these heterologous sequences. We have generated allele integration vectors in which the C-terminal part of a gene's coding sequence can be fused to either heterologous sequence in a single cloning step. We first tested this system with *CITFA7* which encodes a well-characterized subunit of the class I transcription factor A (CITFA), an essential factor for transcription initiation by RNA polymerase I. Targeting either Tc3 or PTP fused to the *CITFA7* mRNA resulted in gene knockdowns that were as efficient and specific as targeting the endogenous *CITFA7* mRNA. Moreover, application of this system to *CITFA1*, which could not be silenced by established methods, demonstrated that this gene encodes

an essential CITFA subunit that mediates binding of the transcription factor complex to RNA polymerase I promoters.

III-1. Introduction

Among kinetoplastid organisms, the tsetse borne, lethal human parasite *Trypanosoma brucei* allows for specific gene silencing through the RNA interference (RNAi) pathway (Kolev et al., 2011; Ngo et al., 1998). This system is based on strong, conditional expression of a ~500 bp-long double-stranded (ds)RNA that targets a gene's mRNA, and is available for both the insect-stage procyclic form (PF) and the mammalian-infective bloodstream form (BF) of the parasite (Wirtz et al., 1999). Given that direct transfection of dsRNA into cells is transient in nature and does not reach all cells in a sample (Ngo et al., 1998), conditional dsRNA expression from genome-integrated vectors is the system of choice for gene silencing experiments in trypanosomes. The first step in developing such a system was the generation of trypanosomes expressing the bacterial tetracycline (Tet) repressor (TetR) that controlled tetracycline-inducible promoters (Wirtz and Clayton, 1995). Reproducible, tightly regulated Tet-inducible expression was originally established in trypanosomes that express the heterologous T7 RNA polymerase (pol) as well as TetR by a mutated T7 promoter (Wirtz et al., 1999). Two *T. brucei brucei* 427 cell lines were generated that have been widely used for RNAi experiments: the PF cell line 29-13 harboring the selectable *NEO* and *HYG* genes, which encode neomycin and hygromycin phosphotransferase, respectively, and the "single marker" (*NEO*) BF cell line (Wirtz et al., 1999).

Several genome-integration vectors have been generated for Tet-inducible expression of gene-specific dsRNA in *T. brucei*. The simplest vector type has opposing T7 promoters. Since gene

fragments can be cloned between the promoters in a single step, this type of vector, termed p2T7 (LaCount et al., 2000) or pZJM (Wang et al., 2000), enabled the generation of RNAi libraries and a first successful forward genetic RNAi screen (Morris et al., 2002). However, the T7 promoters of these vectors, which are each regulated by a single TetR binding site, the Tet operator, were partially active even in the absence of tetracycline (Wang et al., 2000; Wirtz et al., 1999), which could be lethal to transformed trypanosomes in some cases. A more tightly regulated vector has been the stem-loop vector in which a target sequence is cloned in sense and antisense directions around a stuffer fragment (Bastin et al., 2000; Shi et al., 2000). In the original vector, the stem-loop cassette is expressed from the strong EP1 procyclin gene promoter under the control of two Tet operators (Wirtz et al., 1999). Although it takes two to three cloning steps to produce a gene-specific stem-loop construct, the tight regulation of its expression proved to be very useful in PFs (Tschudi et al., 2003). Procyclin is the major cell surface antigen of PFs and procyclin genes are highly expressed by virtue of a multifunctional RNA pol I which is recruited to the procyclin gene promoter (Günzl et al., 2003). However, since procyclin is not expressed in BFs and the procyclin promoter is several-fold less active in this life cycle stage (Biebinger et al., 1996), replacement of the procyclin promoter by a T7 promoter made the "pT7-stl" stem-loop construct more suitable for gene knockdowns in BFs; two Tet operators appear to be sufficient to minimize leakiness from the strong T7 promoter (Brandenburg et al., 2007). Recently established vectors that made conditional transgene expression in trypanosomes independent of T7 RNA pol revealed effective gene knockdowns in PFs (Sunter et al., 2012). Other vector improvements simplified the cloning procedure. The pQuadra system is based on a four component ligation assay and allows the generation of a stem-loop vector in a single cloning step (Inoue et al., 2005) whereas other

strategies made both dual T7 promoter and stem-loop construction compatible with the recombination-based Gateway® cloning system (Kalidas et al., 2011; Lacomble et al., 2009).

In addition to different vector designs, the conditional gene silencing system in *T. brucei* has benefitted from several modifications. Since high T7 RNA polymerase levels appear to be toxic for trypanosomes, putting the expression of this enzyme under tetracycline control enhanced the success rate in obtaining PF cell lines that exhibited an inducible phenotype (Alibu et al., 2005). However, this modification did not generally improve the functionality of p2T7 constructs. Historically, the preferred genome integration site of inducible vectors has been the transcriptionally silent spacer of ribosomal RNA gene (*RRNA*) arrays. It appears that not all *RRNA* loci provide equal conditions for regulated dsRNA expression and, thus, a standard procedure has been to test several clonal cell lines to find the cells with the best knockdown efficiency. To avoid this position effect, Alsford *et al.* marked a specific *RRNA* locus such that vectors could be specifically and reproducibly targeted to this particular *RRNA* spacer, thereby reducing the variability of gene silencing experiments (Alsford et al., 2005). In an independent approach, Wickstead *et al.* found that targeting the p2T7 vector "p2T7-177" to a 177 bp-long, transcriptionally silent repeat region of trypanosome minichromosomes improved regulation of the T7 promoters (Wickstead et al., 2002). Nevertheless, despite these modifications and improvements, the RNAi-mediated gene knockdowns have remained inefficient in some cases. Furthermore, due to the need to express rather long dsRNA molecules, it is possible that siRNAs are produced that affect trypanosome proliferation by targeting the wrong RNA, causing a so-called off-target effect.

We have been using conditional gene silencing to characterize the *T. brucei* class I transcription factor A (CITFA) complex, which is indispensable for the multifunctional RNA pol I system in *T.*

brucei. This parasite uses RNA pol I to transcribe the *RRNA* array, as in all other eukaryotes, yet also employs it to express its major cell surface proteins - procyclins in PFs and variant surface glycoprotein (VSG) in BFs (Brandenburg et al., 2007; Nguyen et al., 2014; Nguyen et al., 2012). The latter is expressed from a single *VSG* gene, drawn from a large repertoire, in one of fifteen 40-60 kb-long BF telomeric expression sites (BESs) (Hertz-Fowler et al., 2008). While trypanosome *RRNA* transcription is localized to the nucleolus, as in other eukaryotes, the active BES is transcribed outside this compartment (Chaves et al., 1998) in the DNase I-resistant expression site body, or ESB (Navarro and Gull, 2001). CITFA consists of the subunits CITFA1 to 7, which are conserved only among kinetoplastid organisms, and the dynein light chain DYNLL1 (also known as LC8). Data obtained thus far strongly indicate that CITFA is a promoter-binding transcription initiation factor: CITFA stably bound the BES promoter in gel shift assays and required both promoter elements for efficient binding, depletion or inhibition of CITFA resulted in a loss of transcription within 121-146 bp of the transcription initiation site, *CITFA7* silencing strongly reduced promoter-proximal RNA pol I occupancy at *RRNA* repeats, and a genome-wide ChIP-seq analysis found CITFA7 occupancy within a BES to be restricted to the promoter region (Brandenburg et al., 2007; Nguyen et al., 2014; Nguyen et al., 2012). To analyze specific functions of individual subunits we have attempted to silence the expression of each subunit gene. Our previous results demonstrated that CITFA2 and CITFA7 are essential for RNA pol I transcription and trypanosome viability (Brandenburg et al., 2007; Nguyen et al., 2012). We also found that three of the eight subunits are required to maintain the integrity of the complex (TN Nguyen and A Günzl, unpublished results). However, we have been unable to generate an unambiguous *CITFA1* knockdown. Although we targeted two different regions of the *CITFA1* mRNA, the RNA levels were not significantly affected (data not shown). To circumvent this setback, we set out to

develop a generally applicable system in BFs for efficient and specific gene knockdowns that targets a heterologous sequence fused to the mRNA of interest. We herein demonstrate that the 3′ gene flank of the *Trypanosoma cruzi* U2AF35 gene (here abbreviated as Tc3) is functional in *T. brucei*, and that Tc3 can be specifically targeted when fused to a *T. brucei* mRNA. In addition, we show that the sequence of the composite PTP tag, consisting of a protein C epitope (ProtC), a TEV protease cleavage site, and tandem protein A domains (ProtA), is an equally good target for gene silencing. Employing this system, we were able to demonstrate that CITFA1 is an indispensable component of the CITFA complex and that it is required for CITFA to bind to the *RRNA* and BES promoters.

III-2. Methods and Materials

DNAs. pT7-CITFA7-stl (Nguyen et al., 2012) and pCITFA7-PTP-BLA (Nguyen et al., 2014) were described previously. For the generation of pT7-Tc3-stl, the entire 679 bp-long 3′ intergenic region of the *T. cruzi U2AF35* gene (accession number TcCLB.510943.60; note that this gene has been annotated as *U2AF26* instead of *U2AF35*), from position 703 to position 1,381 relative to the translation initiation codon, was amplified from *T. cruzi* genomic DNA and inserted into the pT7-stl vector (Brandenburg et al., 2007) in a sense-stuffer-antisense arrangement according to a published protocol (Shi et al., 2000). A T924>C point mutation was introduced to remove an XbaI restriction site which would have interfered with the stem-loop cloning strategy. pT7-PTP-stl was generated analogously using the entire PTP coding sequence (498 bp) (Schimanski et al., 2005a). The cloning strategy required the removal of two MluI restriction sites which was achieved by introducing T153>C and T327>C point mutations.

The tagging vector pCITFA7-HA-Tc3-BLA (**Figure III-S1**) has two cassettes, a C-terminal tagging module and a selectable marker cassette. It is a direct derivative of pCITFA7-HA-BLA (Nguyen et al., 2012) and was obtained by replacing the *T. brucei* RPA1 3′ gene flank with Tc3 using the vector's XhoI and ClaI restriction sites. Furthermore, our tagging vectors are derived from pBluescript II SK+ which has a T7 promoter that, after integration into an endogenous allele, could lead to overexpression of downstream genes in T7 RNA polymerase-expressing sm BF and 29-13 PF cells. We therefore removed 27 bp from pCITFA7-HA-Tc3-BLA beginning precisely at the T7 promoter and ending in the downstream KpnI restriction site. In the same way, we removed the T7 promoter from pCITFA7-PTP-BLA and termed the corresponding plasmid pCITFA7-PTP-BLAv2 (**Figure III-S2**). pCITFA1-HA-Tc3-BLA was obtained from pCITFA7-HA-Tc3-BLA by replacing the CITFA7 sequence with 696 bp of the C-terminal CITFA1 coding region (position 700 to position 1395) using the ApaI and NotI restriction sites.

The following DNA oligonucleotides were used in semi-quantitative RT-PCR: 5′-CCTACGGTGCAGCCATGCCGTTGG-3′ / 5′-TTCGGCACTGCCATATGCGAC-3′ (CITFA7 coding sequence); 5′-CAATAACAGGAACAGCTGCACCAAG-3′ / 5′-GAGGAAACTCAAGTGCATTG-3′ (Tc3); 5′-GTAGACAACAAATTCAACAAAG-3′ / 5′-ATTTAGCTTTTTAGCTTCTGC-3′ (PTP). Oligonucleotides for *VSG221*, *TFIIB* and *CITFA2* amplification were previously described (Nguyen et al., 2012). Quantitative RT-PCR of *CITFA1* mRNA was carried out either with oligonucleotide pair 5′-ATCGGATGTTGAGTCGCTGCGTTG G-3′ / 5′-AAAGTCATTCCATGCCACTGGAACC-3′ (*CITFA1* coding sequence) or pair 5′-AATACGCCAG GCAGATTGATGC-3′ / 5′-TTAAGCGTAGTCAGGTACGTCGTAAGG -3′ (*CITFA7* coding / HS). BES and *RRNA* promoter consensus oligonucleotides and oligonucleotides specific to the TFIIB gene and the β-

/α- tubulin intergenic region, used in qPCR, were previously specified (Nguyen et al., 2012; Park et al., 2011).

Cells. BFs were cultured in HMI-9 medium as specified previously (Park et al., 2011). Transfections were done with 1-2 x 10^7 BF trypanosomes using the Amaxa Basic Parasite Nucleofector kit (Lonza). Specifically, trypanosomes were pelleted at room temperature at 1,500 x g for 5 min and cell pellets were resuspended in 100 µl of *Solution 1* containing 18% *Supplement 1*. After mixing DNA (5 µg PCR product or 10 µg of plasmid DNA) into the resuspension, trypanosomes were electroporated using program X-001 on the Nucleofector 2b unit (Lonza). 500 µl of pre-warmed HMI-9 medium was added immediately to the transfected cells which were then transferred to 50 ml of ~37°C warm medium. After allowing the cells to recover for 15 min at 37°C, the cell culture was distributed into two 24-well plates. Trypanosomes were cultivated without antibiotic selection overnight after which additional medium containing selecting antibiotics was added to each well. Transfectants typically reached a transferrable cell density six to nine days later. Cells were cultured in 2.5 µg/ml G418, 1 µg/ml phleomycin and/or 2 µg/ml blasticidin.

Cell lines C7HA-Tb3 and C7HA-Tc3 were obtained by transfecting wild-type BF 427 cells with PshAI-linearized plasmids CITFA7-HA-BLA and CITFA7-HA-Tc3-BLA, respectively. The basal BF cell lines for conditional expression of Tc3 or PTP dsRNA, smTc3 and smPTP, were obtained by transfecting sm cells with the EcoRV-linearized vectors pT7-Tc3-stl and pT7-PTP-stl, respectively. These vectors were targeted to the transcriptionally silent *RRNA* spacer. smC7HA-Tc3 and smC7-PTP cell lines were generated by targeted integration of PshAI-linearized plasmids CITFA7-HA-Tc3-BLA and CITFA7-PTP-BLAv2, respectively, into the *CITFA7* locus in the first step and by replacing in the second step the remaining *CITFA7* wild-type allele with a

PCR product in which 100 bp of *CITFA7* gene flanks surrounded the hygromycin phosphotransferase coding sequence. Cell line smC1HA-Tc3 was obtained analogously using the MfeI-linearized plasmid CITFA1-HA-Tc3-BLA. Correct DNA integrations were analyzed by PCR of genomic DNA with at least one oligonucleotide placed outside the cloned or amplified sequence (data not shown).

For gene silencing experiments, dsRNA synthesis was induced with doxycycline, a more stable derivative of tetracycline, at 2 µg/ml. Cells were counted and diluted to 2 x 10^5 cells/ml daily.

Antibodies and protein analysis. For the sedimentation analysis, extract was prepared from non-induced BFs and from *CITFA1*-silenced BFs as specified previously (Park et al., 2011). 100 µl of extract was then loaded onto 4 ml 10-40% linear sucrose gradients, ultra-centrifuged and fractionated exactly as has been described (Brandenburg et al., 2007). Immunoblots of HA- and PTP-tagged proteins were probed with a rat monoclonal anti-HA antibody (Roche) and the mouse monoclonal anti-ProtC antibody HPC4 (Roche), respectively. Immune sera against CITFA7, TFIIB and U2A$^{/}$ (also known as U2-40K) were described previously (Cross et al., 1993; Nguyen et al., 2012; Schimanski et al., 2006). To obtain recombinant CITFA3 for antibody production, the entire *CITFA3* coding region was placed downstream of the glutathione *S*-transferase sequence in pGEX-4T-2 vector (GE Healthcare). Recombinant GST-CITFA3 was expressed in *Escherichia coli* strain BL21Star (DE3) and purified by glutathione affinity chromatography (GE Healthcare) following the manufacturer's recommendations. Generation of anti-CITFA3 immune serum was achieved by immunization of female Sprague Dawley rats with purified GST-CITFA3 as detailed previously (Schimanski et al., 2006). Polyclonal anti-GST-CITFA3 antibodies were purified from rat immune serum through first pre-clearing the serum of antibodies that non-specifically interact with trypanosome proteins. This was done by separating whole cell lysates of a total of 4.5 x 10^8

wild-type PF trypanosomes in nine lanes on a 12% SDS-polyacrylamide gel. The gel was transferred to a PVDF membrane and blocked in Tris-buffered saline (TBS), pH 7.5, containing 5% milk and 0.1% Tween20, for 2 hours. After washing twice in TBS the 50-55 kDa range of the membrane which contained CITFA3 was removed and the remaining membrane was incubated with 1 ml of antiserum and 9 ml of blocking solution at 4°C for 16 hours. 100 µg of recombinant GST-CITFA3 was then run on four different SDS-PAGE gels, which were then transferred and blocked as described above. The 70-80 kDa range of these four membranes, corresponding to size of GST-CITFA3, was excised, cut into small strips, and incubated with the pre-cleared antiserum and blocking solution mixture for 16 hours, at 4°C, with rotation. These membrane strips were washed three times with TBS, and then rinsed briefly one time with water. Antibodies were eluted by incubating the strips in 1 ml of 0.2 M glycine, pH 2.8, for 5 min. Eluted antibodies were then immediately quenched by adding 1 M Tris-HCl, pH 8.0, until the solution pH reached 7.5, after which bovine serum albumin was added to a final concentration of 1%. The purified antibody was aliquoted and stored at -20°C until use.

RNA analysis. Total RNA was prepared by the hot phenol method as described previously (Nguyen et al., 2007). For the analysis of ribosomal (r)RNA, total RNA was separated in Reliant pre-cast 1.25% SeaKem Gold agarose RNA gels (Lonza), and rRNA was detected by ethidium bromide staining. Relative amounts of SL RNA and U2 snRNA were determined by primer extension of 10 µg of total RNA using the $5'$-^{32}P-end-labeled oligonucleotides SL-1394 and U2f (Günzl et al., 1992) and Superscript reverse transcriptase II (Invitrogen) according to the manufacturer's specifications. Primer extension products were separated on denaturing 8% polyacrylamide-50 % urea gels and detected by autoradiography. For PCR analyses, total RNA was reverse transcribed with Superscript reverse transcriptase II and either oligonucleotide-dT or

random hexanucleotides (Roche). Semi-quantitative PCR was performed using cycle numbers which were empirically determined to be within the linear amplification range for each oligonucleotide pair. Oligonucleotide pairs used in qPCR were verified for their specificity and suitability by standard agarose gel electrophoresis. Additionally, for every round of qPCR reactions a melt curve analysis was included to ensure the amplification of only a single product, and linear regression analysis of a serial dilution of input material confirmed that the coefficient of determination (r^2) was within the 0.98 to 1.0 range.

Immunofluorescence microscopy. BF microscopy was carried out as published previously (Nguyen et al., 2012) except that coverslips with settled BF cells were incubated with a 1:100 dilution of the affinity-purified, polyclonal rat anti-CITFA3 antibody, and were then, after washing, incubated with the 1:500 diluted anti-rat IgG antibody Alexa 488 (Invitrogen) and 4,6-diamidino-2-phenylindol (DAPI) at a final concentration of 2 ng/µl. Imaging was performed using a Zeiss AxioVert 200 microscope and Zeiss Axiovision 4.6.3.0 software. CITFA3 localizations were captured using a FITC filter and a fixed exposure time of 6 s, while DAPI images were captured using a DAPI filter and a variable exposure time, which averaged 0.5 s.

Chromatin immunoprecipitation (ChIP). Anti-CITFA3 ChIP assays, using the purified polyclonal anti-CITFA3 antibody, were performed with smC1HA-Tc3 cells that were either not induced or induced by doxycyline for 42 hour as recently described (Nguyen et al., 2014). In negative controls, chromatin was immunoprecipitated with a non-specific rat immune serum. The precipitated DNA was analyzed by qPCR using consensus oligonucleotides for the slightly varying copies of *RRNA* and *BES* promoters and an oligonucleotide pair specific for the β-/α- tubulin intergenic region. The percent immunoprecipitation (IP) was calculated relative to the input material and corrected by subtracting the percent IP of the negative control assays.

III-3. Results

System outline. Our aim was to generate a system in which the same heterologous sequence can be functionally fused to any mRNA and provide a target for efficient and reproducible gene knockdown. In a first step, we would generate a single marker (sm) BF cell line that conditionally expresses dsRNA of a heterologous sequence (HS) without consequence on cell proliferation (**Figure III-1**). This smHS cell line would carry *NEO* used to generate sm cells (Wirtz et al., 1999) and the bleomycin resistance marker (*BLE*) as part of the stem-loop vector for conditional HS dsRNA expression. For a specific gene knockdown, the smHS cell line would be used in two consecutive transfections to enable specific knockdown of any single copy gene of interest, denoted here as gene *X*. In the first transfection, site-specific integration of a PCR amplification product of the hygromycin phosphotransferase coding region (*HYG*) surrounded by 5′ and 3′ flanks of gene *X* would lead to the knockout of one wild-type *X* allele while, in the consecutive transfection, targeted integration of plasmid X-HS-BLA (BLA stands for the selectable marker gene blasticidin-S deaminase) into the *X* locus would fuse the HS to gene *X*. The corresponding smX-HS cell line would exclusively express the X-HS fusion that can be targeted by doxycycline-mediated induction of HS dsRNA synthesis (**Figure III-1**).

The *Trypanosoma cruzi U2AF35* 3′ gene flank is functional in *T. brucei*. We speculated that a *Trypanosoma cruzi* 3′ gene flank could provide a HS in the form of a functional 3′ UTR because all trypanosomatids process their mRNAs by spliced leader (SL) *trans* splicing and polyadenylation (Günzl, 2010), making it likely that a *T. cruzi* gene flank is able to direct these RNA processing steps in *T. brucei*. At the same time, the *T. cruzi* intergenic sequences are divergent from their *T. brucei* counterparts and are unlikely to give rise to siRNAs that target

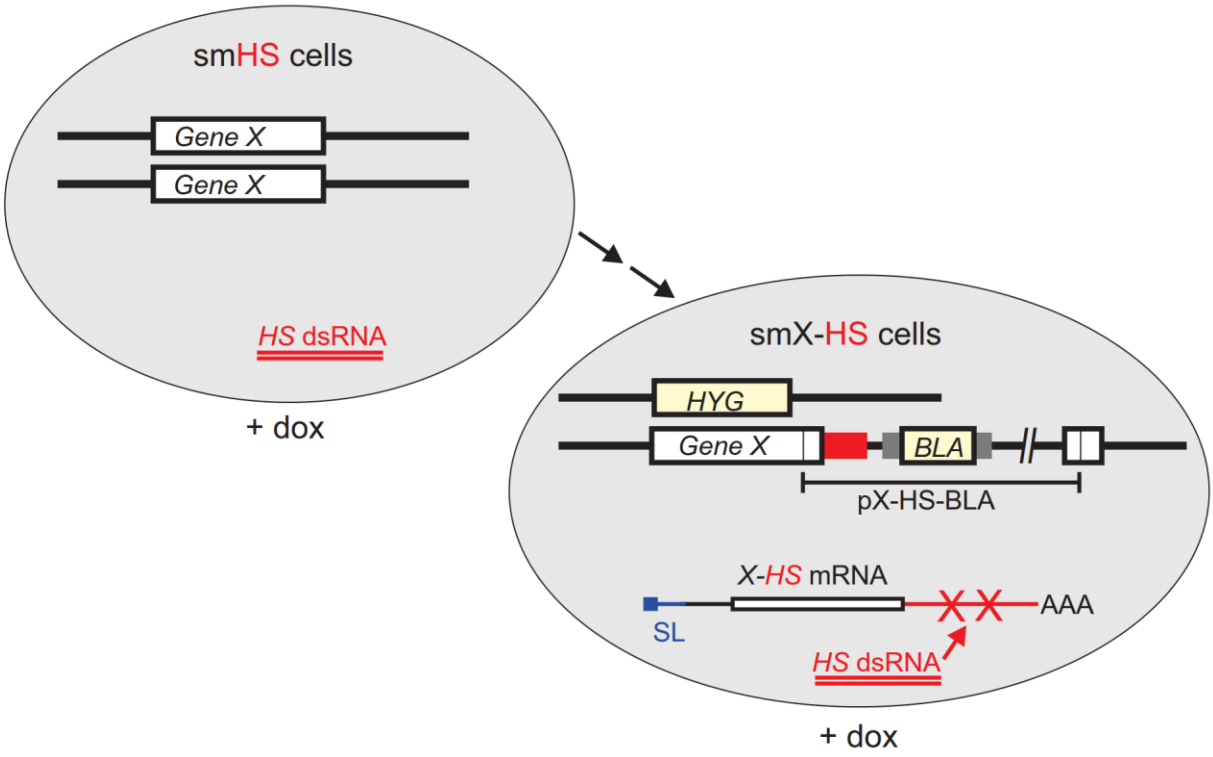


Figure III-1. Gene silencing system targeting a heterologous sequence (HS). The system is based on a single marker BF cell line that has been stably transfected with a stem-loop construct (sm-HS cell line) containing a doxycyline (dox)-inducible promoter for the expression of a heterologous dsRNA. In the absence of further genetic manipulations, induction of this dsRNA expression has no consequence on trypanosome proliferation. To enable a specific knockdown of any single copy gene *X*, two further consecutive transfections of smHS cells are necessary: one to eliminate an *X* allele by the hygromycin resistance marker (*HYG*), and one to integrate a plasmid into the remaining *X* allele which will fuse the HS to the X mRNA sequence. Doxycycline-induced (+ dox) expression of HS dsRNA will then target and destroy the *X-HS* mRNA hybrid.

T. brucei mRNAs. After comparing known *T. cruzi* 3′ UTRs (Brandao and Jiang, 2009) we chose to analyze the suitability of the 3′ UTR of the *TcU2AF35* mRNA because its complete cDNA had been characterized, and the 3′ UTR length of 390 nt appeared to be sufficiently large for an efficient gene knockdown (Vazquez et al., 2003). Moreover, the complete intergenic region between *TcU2AF35* (accession number TcCLB.510943.60, www.genedb.org) and its downstream neighbor *TcCLB.510943.50*, encoding putative delta-1-pyrroline-5-carboxylate dehydrogenase, is only 679 bp long and should harbor all necessary RNA processing signals. Finally, *U2AF35* encodes an essential RNA splicing factor that, in trypanosomes, is involved in the initial steps of the ubiquitous SL *trans* splicing process (Vazquez et al., 2009), which suggested that the *TcU2AF35* 3′ UTR could support a sufficient level of constitutive gene expression.

We first tested the functionality of the *TcU2AF35* 3′ gene flank (Tc3) in *T. brucei* with the well-characterized *CITFA7* gene. CITFA7 is an essential subunit of the CITFA complex and *CITFA7* silencing in BFs led to clear defects in RNA pol I transcription, e.g. a decrease of the RNA pol I transcripts rRNA and *VSG221* mRNA of the active *VSG* gene (Nguyen et al., 2012). Moreover, CITFA7 can be functionally tagged at the C-terminus and there is no haplo-insufficiency effect after deleting one *CITFA7* allele either in BFs or PFs (Nguyen et al., 2012). We generated two cell lines by targeting the integration of a plasmid to the endogenous *CITFA7* gene (**Figure III-2A**). In both cases, the HA tag sequence was fused 3′ to the CITFA7 coding region followed by the *T. brucei* (*Tb)RPA1* 3′ gene flank in the C7HA-Tb3 cell line or by the Tc3 gene flank in the C7HA-Tc3 cell line. RPA1 is the largest subunit of RNA pol I. The *TbRPA1* 3′ gene flank is present in all our C-terminal tagging constructs and has supported the expression of a variety of factors such that the knockout of the remaining wild-type allele did not cause haplo-

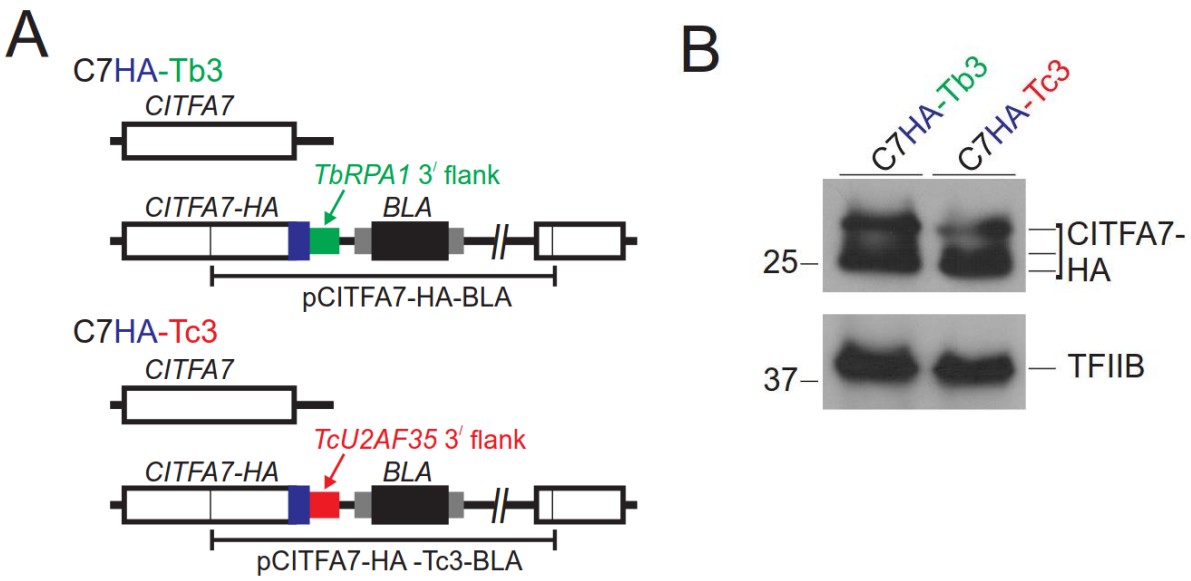


Figure III-2. The 3′ gene flank and UTR of *T. cruzi* (Tc)U2AF35 are functional in *T. brucei*. (**A**) Schematic outline (not to scale) of the *CITFA7* locus in control cell line C7HA-Tb3 and in cell line C7HA-Tc3. In both cell lines the HA tag sequence was fused to the CITFA7 coding region by targeted integration of the specified, linearized plasmid into one *CITFA7* allele. The difference is that in C7HA-Tb3 and C7HA-Tc3 cells the manipulated CITFA7 allele is under the control of the *T. brucei* RPA1 (green) and the heterologous TcU2AF35 (red) 3′ gene flanks, respectively. The CITFA7 coding region, and the HA tag and *BLA* sequences are indicated by open, blue and black boxes, respectively. The smaller gray boxes surrounding *BLA* represent *T. brucei* gene flanks providing RNA processing signals. (**B**) Immunoblot of C7HA-Tb3 and C7HA-Tc3 whole cell lysates detecting CITFA7-HA with a monoclonal anti-HA antibody and, as a loading control, the transcription factor TFIIB with an anti-TFIIB polyclonal immune serum. Note that, due to phosphorylation, CITFA7 separates in multiple bands.

insufficiency phenotypes (Schimanski et al., 2005a). Accordingly, CITFA7-HA was easily detectable in C7HA-Tb3 cell lysates in multiple bands (**Figure III-2B**) that represent the various [un]phosphorylated forms of CITFA7 (Nguyen et al., 2012). In C7HA-Tc3 cells, Tc3 supported 78% of the CITFA7-HA expression observed in C7HA-Tb3 cells (**Figure III-2B**). These results showed that, in *T. brucei*, the heterologous Tc3 sequence directed the processing of functional CITFA7-HA mRNA and the corresponding 3′ UTR supported an adequate level of *CITFA7* expression.

Expression of Tc3 dsRNA does not affect BF trypanosome proliferation. Comparing the Tc3 sequence to the *T. brucei brucei* Lister 427 genome returned a single, siRNA-sized stretch of identical sequence (27 bp) downstream of the non-syntenic gene *Tb427.10.1000*. Since this sequence motif was found only in a fraction of the heterogenous *Tb427.10.1000* 3′ UTRs (www.tritrypdb.org), we anticipated Tc3 dsRNA not to give rise to deleterious siRNAs. To test this, we inserted a stem-loop construct of the Tc3 sequence into the pT7-stl vector that has a tetracycline-inducible T7 promoter (Brandenburg et al., 2007). The vector, targeted to the ribosomal spacer, was transfected into BF sm cells which express both the tetracycline repressor and T7 RNA polymerase (Wirtz et al., 1999). Immediately after transfection, cells were cloned by limiting dilution. Three of the resulting cell lines were evaluated for their knockdown competence by transfecting them with the CITFA7-HA-Tc3-BLA construct and monitoring knockdown efficiencies on the RNA and protein levels (see below, and data not shown). Based on these results, one of the parent cell lines, termed smTc3, was chosen for all further experiments. As shown in **Figure III-3A**, induction of Tc3 dsRNA synthesis by doxycycline did not affect proliferation of smTc3 cells in culture, suggesting that no deleterious off-target effects occurred.

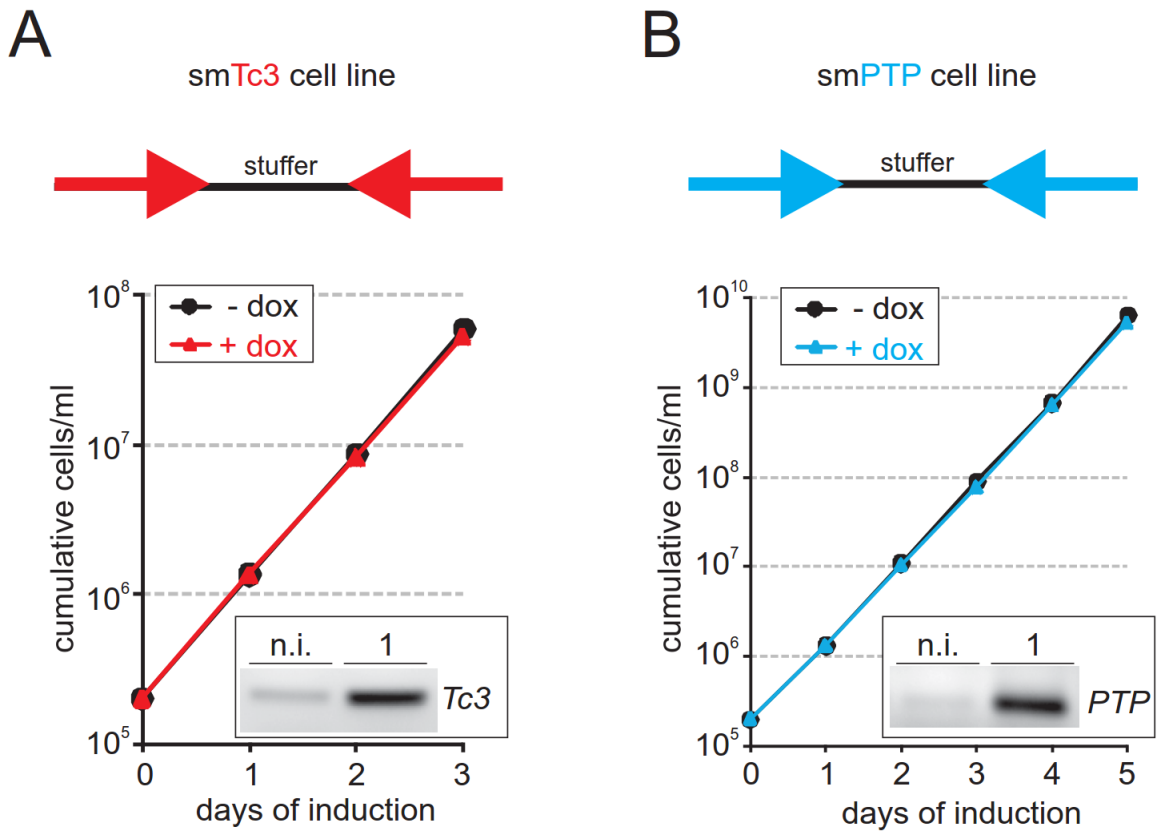


Figure III-3. Proliferation of smTc3 and smPTP cells is not affected by heterologous dsRNA expression. The smTc3 (**A**) and smPTP (**B**) cell lines are single marker BF cell lines in which stem-loop constructs were integrated into the *RRNA* spacer for inducible expression of Tc3 and PTP dsRNA, respectively. Addition of doxycycline to the medium did not inhibit cell proliferation in either cell line while it induced strong expression of Tc3 or PTP RNA, as analyzed by semi-quantitative PCR of random hexamer-derived cDNA (inserts).

Generation of an alternative smPTP cell line. At this point, we considered testing a second heterologous sequence in parallel. We decided on the large PTP tag because we have utilized it repeatedly for tandem affinity purification (Schimanski et al., 2005a), indirect fluorescence microscopy (Park et al., 2011), and chromatin immunoprecipitation experiments (Lee et al., 2010; Park et al., 2011). The composite, heterologous PTP tag, designed for a modified tandem affinity purification approach (Schimanski et al., 2005a), comprises 166 amino acids and consists of human-derived ProtC (Stearns et al., 1988), a tobacco etch virus-derived protease cleavage site, and a tandem ProtA domain of *Staphylococcus aureus* (Rigaut et al., 1999). Importantly, the PTP tagging strategy has been based on integration of a PTP plasmid into an endogenous allele as shown in **Figure III-1**. Hence, we wanted to know whether the PTP tag sequence could be used as a target for efficient gene knockdowns as well. A bioinformatic analysis could not detect any sequence match longer than 18 bp between the PTP sequence and the *T. brucei brucei* 427 genome (data not shown). Analogously to cell line smTc3, we generated clonal smPTP cell lines that conditionally express PTP dsRNA. To pick the most efficient line for subsequent experiments, we determined PTP dsRNA levels before and after doxycycline induction by reverse transcription of total RNA using random hexamers and semi-quantitative PCR of the PTP sequence (**Figure III-3B** and data not shown). Induction of PTP dsRNA by doxycycline in the chosen smPTP cell line did not affect trypanosome proliferation, again indicating that this heterologous dsRNA does not target vitally important endogenous *T. brucei* RNAs (**Figure III-3B**).

The Tc3 3′ UTR and the PTP sequence are efficient knockdown targets. To test whether targeting either of the two HSs by RNAi leads to efficient gene knockdowns, we fused both of these sequences to the *CITFA7* gene. As depicted in **Figure III-1**, this required two consecutive transfections of smTc3 and of smPTP cells so that they exclusively expressed CITFA7-HA mRNA

with the Tc3 3′ UTR and CITFA7-PTP mRNA, respectively. In the first transfection of smTc3 cells, we integrated plasmid CITFA7-HA-Tc3-BLA into one of the two *CITFA7* alleles and, in the second step, we eliminated the remaining *CITFA7* allele by transfecting a PCR product (Arhin et al., 2004) that comprised the hygromycin phosphotransferase coding region surrounded by 100 bp-long *CITFA7* gene flanks (**Figure III-4A**). Three independently derived, clonal smC7HA-Tc3-RNAi cell lines exhibited nearly identical growth defects upon induction of Tc3 dsRNA synthesis. As shown for one representative line in **Figure III-4B**, addition of doxycyline stopped culture growth after 24 hours and reduced the number of surviving trypanosomes within the next 48 hours. Since a very similar growth curve was obtained previously when *CITFA7* mRNA was targeted directly in smC7 cells (Nguyen et al., 2012), this result suggested that targeting of the Tc3 3′ UTR resulted in effective depletion of CITFA7.

Analogously to smC7HA-Tc3 cells, we generated the cell line smC7-PTP in which integration of plasmid pCITFA7-PTP-BLA fused the PTP sequence to the CITFA7 coding region in one *CITFA7* allele, and the hygromycin resistance marker replaced the remaining wild-type allele (**Figure III-4C**). Again, inducing the synthesis of PTP dsRNA stopped proliferation of smC7-PTP cells after one day and led to cell death thereafter (**Figure III-4D**). Semi-quantitative RT-PCR analyses showed that CITFA7-HA, CITFA7-PTP and CITFA7 mRNA were strongly reduced after 1 day of induction in smC7HA-Tc3, smC7-PTP and smC7 cells, respectively (**Figure III-4E**). This reduction was observed despite the fact that the level of the control TFIIB mRNA increased in each of these experiments. As we have shown previously, silencing of CITFA subunit genes rapidly decreases the levels of RNA pol I transcripts in induced cells, including ribosomal RNA, the most abundant component in total RNA preparations, which in turn leads to relative increases of RNA pol II and III transcripts (Brandenburg et al., 2007; Nguyen et al., 2012). As anticipated

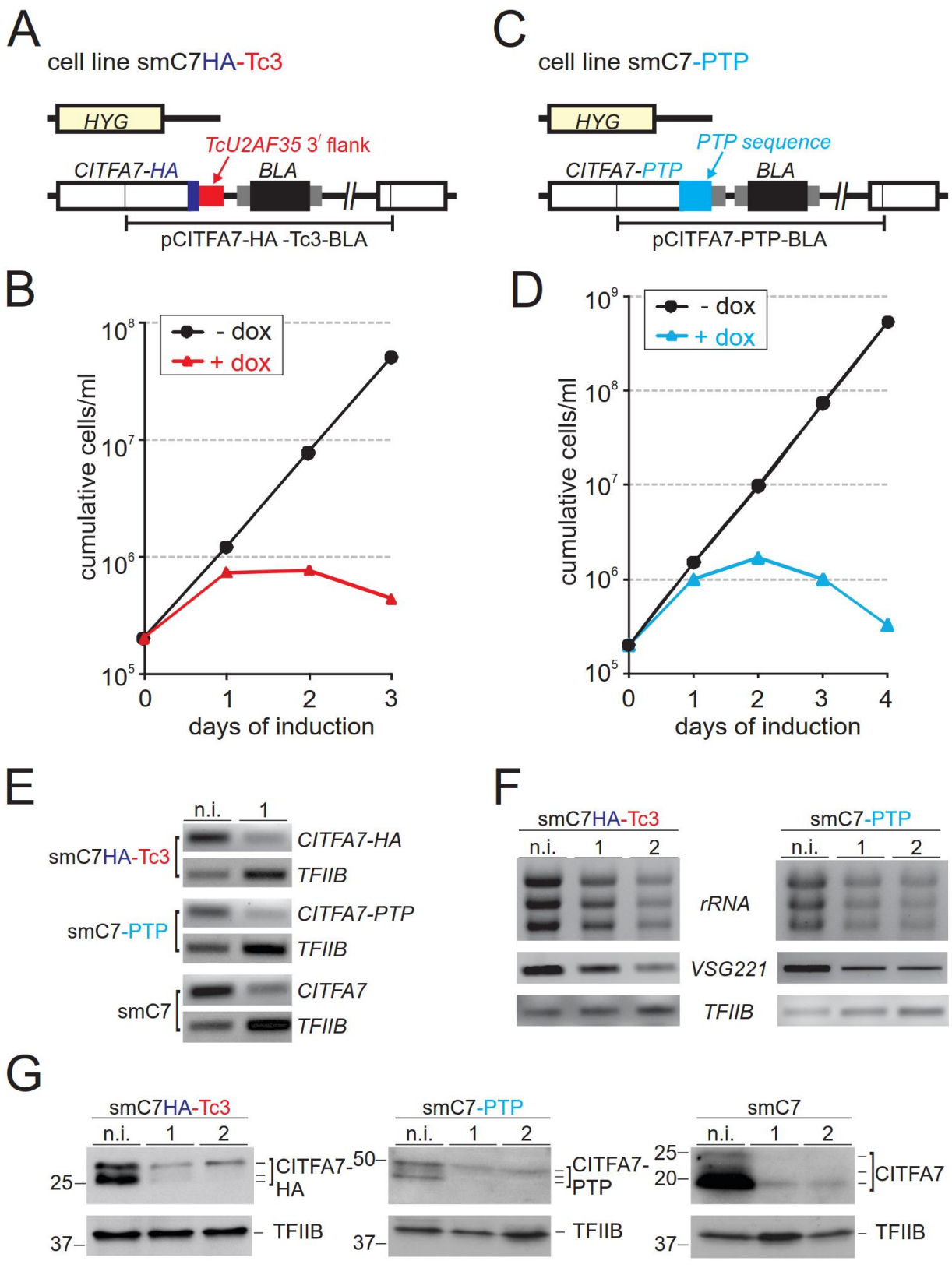


Figure III-4. Effective and specific *CITFA7* silencing by targeting heterologous sequences.

(Figure III-4 legend cont.) (**A**) Schematic outline (not to scale) of the *CITFA7* locus in smC7HA-Tc3 cells. One *CITFA7* allele was replaced by the hygromycin resistance marker (*HYG*, yellow box) and the second *CITFA7* allele modified by targeted integration of plasmid CITFA7-HA-Tc3-BLA. Coloring of boxes corresponds to the description in the legend of **Figure 2A**. (**B**) Culture growth of a representative smC7HA-Tc3 cell line in the presence and absence of doxycyline. (**C**) Schematic outline (not to scale) of smC7-PTP cells in which integration of plasmid CITFA7-PTP-BLA fused the PTP sequence (cyan box) to the 3′ end of the CITFA7 coding region. (**D**) Corresponding growth curve of a representative smC7-PTP cell line. (**E**) Semi-quantitative PCR analysis of oligo dT-primed and reverse transcribed CITFA7-HA, CITFA7-PTP, and CITFA7 mRNA in non-induced (n.i.) and one day induced trypanosomes of cell lines smC7HA-Tc3, smC7-PTP, and smC7, respectively. TFIIB mRNA was analyzed in parallel as a control. (**F**) Relative abundances of ribosomal RNA (rRNA) and of VSG221 and TFIIB mRNA were determined by ethidium bromide staining and semi-quantitative RT-PCR, respectively, in smC7HA-Tc3 and smC7-PTP cells that were non-induced or induced for one and two days. (**G**) Immunoblot detecting CITFA7-HA with an anti-HA antibody in smC7HA-Tc3 cell lysates, CITFA7-PTP with an anti-ProtC antibody in smC7-PTP cell lysates, and wild-type CITFA7 with a polyclonal immune serum in smC7 cell lysates. Detection of TFIIB on the same blots served as a loading control. Cells were analyzed in their non-induced (n.i.) state or when they were grown in the presence of doxycycline for one and two days.

from these previous results, targeting the HSs that were fused to the *CITFA7* mRNA led to a decrease of rRNA and of *VSG221* mRNA from the active *VSG* gene (**Figure III-4F**). In addition, immunoblotting showed that the gene knockdowns led to a rapid loss of CITFA7 protein (**Figure III-4G**). Together, these results clearly demonstrated that the HS-targeted *CITFA7* gene knockdowns were specific and efficient, affecting the abundance of RNA pol I transcripts in the same way as did silencing of *CITFA7* by dsRNA of endogenous sequence, as previously shown (Nguyen et al., 2012).

Tc3 3′ UTR mediated efficient *CITFA1* silencing. Next, we applied this gene knockdown system to the *CITFA1* gene, which we had not been able to silence efficiently thus far. Again, two consecutive transfections of smTc3 cells generated cell line smC1HA-Tc3 in which, after the knockout of one *CITFA1* allele, targeted integration of plasmid CITFA1-HA-Tc3-BLA fused the HA sequence and the Tc3 gene flank to the remaining *CITFA1* allele (**Figure III-5A**). In the absence of doxycyline smC1HA-Tc3 cells proliferated as fast as smTc3 cells, indicating that the Tc3 3′ UTR supported sufficient *CITFA1* expression from a single allele and that the C-terminal HA tag did not impair the functionality of CITFA1 (**Figure III-5B**). Adding doxycyline to the medium was then similarly deleterious to trypanosome proliferation and viability as in the *CITFA7* knockdown: Trypanosome proliferation was affected one day after induction and trypanosome numbers started to decline after two days of induction (**Figure III-5B**). RT-qPCR analysis revealed that the Tc3-targeted knockdown reduced *CITFA1* mRNA abundance relative to that of TFIIB mRNA by ~80% (**Figure III-5C**). Interestingly, we obtained slightly different results depending on which part of the cDNA was amplified. With oligonucleotides specific for the *CITFA1* coding region, the reduction of *CITFA1* mRNA was on average 77% whereas the reduction with an HS-

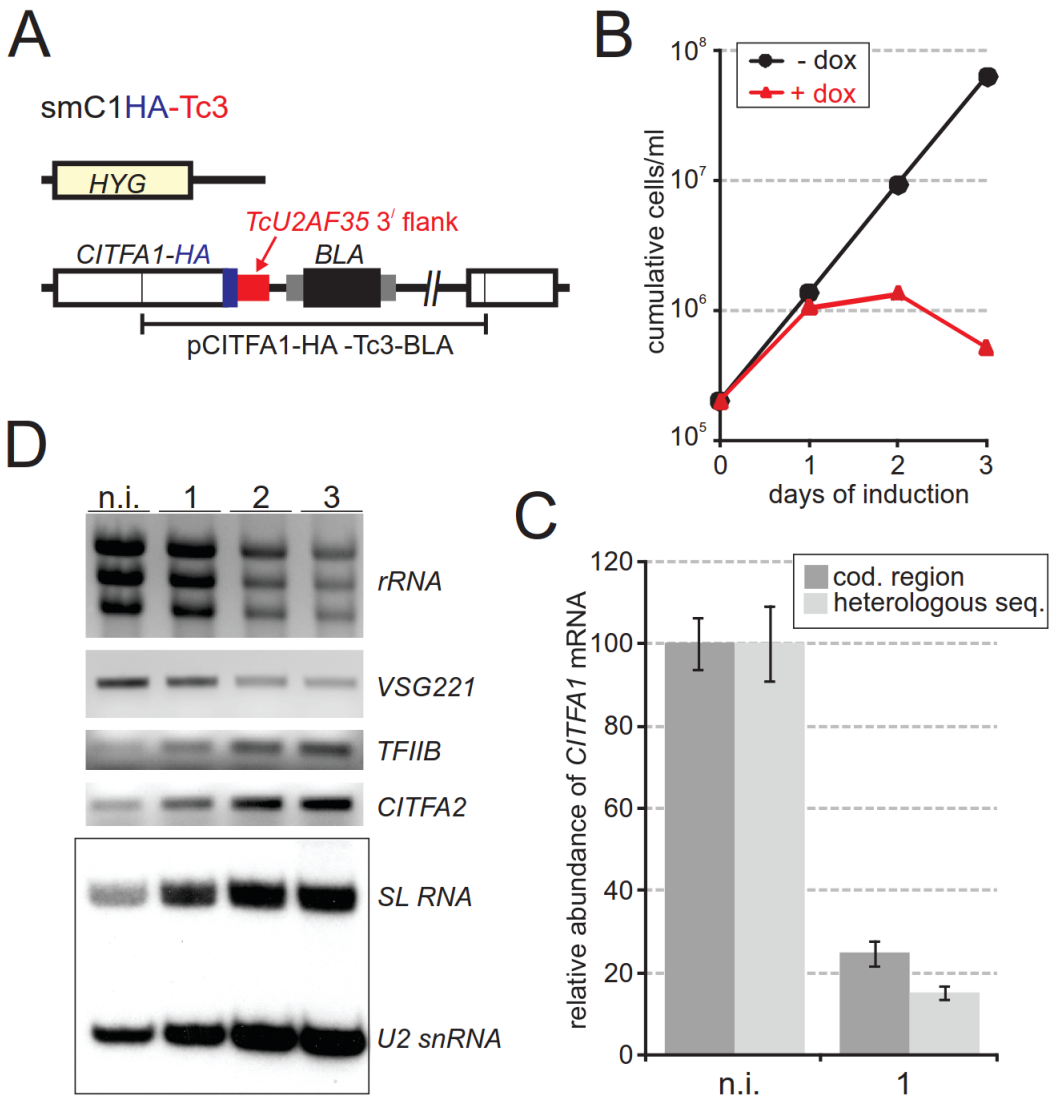


Figure III-5. Effective and specific *CITFA1* silencing by targeting the TcU2AF35 3′ UTR. (**A**) Schematic outline (not to scale) of the *CITFA1* locus in smC1HA-Tc3 cells. (**B**) Culture growth of a representative smC1HA-Tc3 cell line in the presence and absence of doxycyline. (**C**) RT-qPCR analysis of *CITFA1* mRNA in non-induced (n.i.) and one day induced trypanosomes. The oligo dT-derived cDNA was either amplified in the CITFA1 coding region (cod. region) or in the heterologous TcU2AF35 3′ UTR (heterologous seq.). CITFA1 mRNA abundance was normalized with that of TFIIB and its level in non-induced cells was set to 100 in each of three independent experiments. (**D**) Relative RNA abundances in total RNA preparations of non-induced cells and

(**Figure III-5 legend cont.)** of cells induced for 1, 2, or 3 days were analyzed by ethidium bromide staining (rRNA), semi-quantitative RT-PCR (VSG221, TFIIB and CITFA2 mRNA), or a primer extension assay (SL RNA, U2 snRNA).

specific oligonucleotide was 83%. Although this difference seems not that dramatic, we would like to point out that we have made similar observations with other gene knockdowns. Amplification of different parts of the reverse-transcribed mRNA revealed different knockdown efficiencies, with the targeted region typically resulting in the greatest reduction of mRNA abundance (AG laboratory, unpublished results). Further RNA analysis showed that, as expected, *CITFA1* silencing decreased the abundance of the major RNA pol I transcripts rRNA and *VSG221* mRNA and, consequently, elevated the relative abundances of RNA pol II-synthesized TFIIB mRNA, CITFA2 mRNA and SL RNA, as well as that of the RNA pol III transcript U2 snRNA (**Figure III-5D**). Together, these results demonstrated that *CITFA1* was efficiently and specifically silenced by targeting the Tc3 3′ UTR, identified CITFA1 as the third CITFA subunit that is essential for trypanosome viability in culture, and indicated that CITFA1 has an essential function in trypanosome RNA pol I transcription.

CITFA1 is required for binding of the transcription factor complex to RNA pol I promoters. The analysis of specific functions of individual CITFA subunits has been hampered by the fact that amino acid sequences of CITFA subunits have not revealed functional motifs such as DNA binding domains. In addition, procedures intended to break up the complex into smaller functional units have failed thus far (data not shown). However, we recently found that depletion of CITFA7 from cells resulted in the concomitant, rapid loss of other CITFA subunits within 2 days of *CITFA7* silencing, indicating that CITFA7 has a scaffold function and that the stability of CITFA subunits in trypanosomes depends on the integrity of the transcription factor complex (TN Nguyen and A Günzl, unpublished results). An immunoblot analysis of whole cell lysates showed that CITFA1-HA was strongly depleted after 1 and 2 days of *CITFA1* silencing (**Figure III-6A**). To assess the abundance of other CITFA subunits in these samples we used previously established immune sera

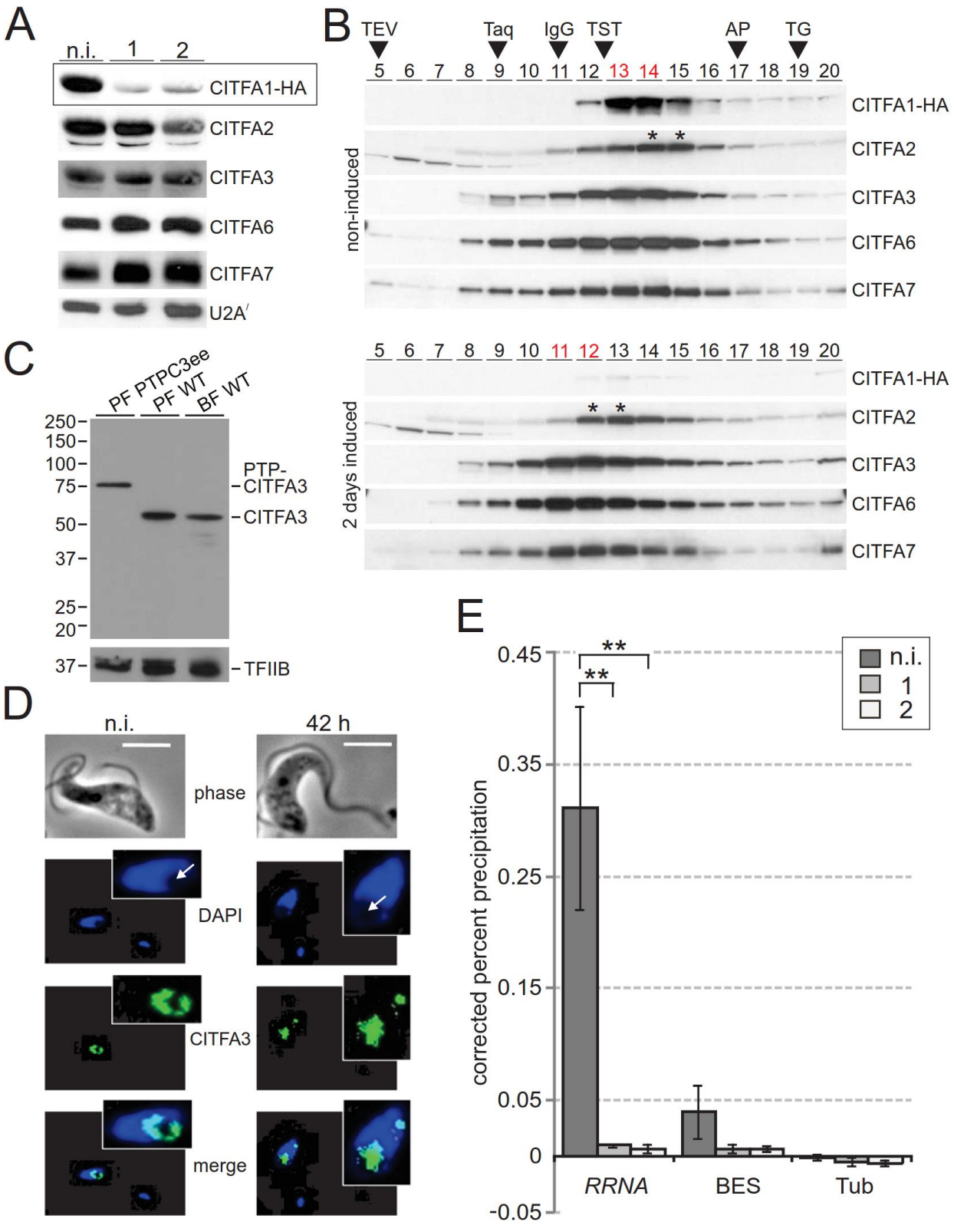


Figure III-6. CITFA1 is required for CITFA promoter binding *in vivo*.

(Figure III-6 legend cont.) (**A**) Immunoblot detecting the indicated CITFA subunits in whole cell lysates of non-induced cells (n.i.) and in cells in which CITFA1-HA was silenced for 1 or 2 days (boxed panel) on the same blot. Detection of the spliceosomal U2A$^{/}$ protein served as a loading control. (**B**) Sedimentation of extract by ultracentrifugation in a 10 to 40% linear sucrose gradient. Fractions 5 to 20, taken from top to bottom, were analyzed by immunoblotting. Note that the two CITFA1 blots were co-developed. For comparison, sedimentations of TEV protease (29 kDa), *Taq* DNA polymerase (95 kDa), IgG (150 kDa, 6.6S), the TRF4-SNAPc-TFIIA transcription factor complex (TST, 230 kDa), apoferritin (AP, 444 kDa, 17S), and thyroglobin (TG, 660 kDa, 19S) were analyzed in parallel gradients (arrowheads). Fractions with CITFA3, 6 and 7 co-sedimentation peaks are indicated by red lettering and the shifted CITFA2 peaks by asterisks. (**C**) Immunoblot using the purified, polyclonal anti-CITFA3 antibody for detection of PTP-CITFA3 in PF PTPC3ee cells and untagged CITFA3 in PF and BF wild-type (WT) whole cell lysates. Loading was controlled by the detection of the RNA pol II transcription factor TFIIB. (**D**) Indirect immunofluorescence microscopy of CITFA3 in non-induced cells and cells in which *CITFA1* was silenced for 42 hours. Nucleolar areas are indicated by white arrows. For the induced cells, an example was chosen in which the putative ESB was detected as an additional spot outside the nucleolus, bar: 5 µm. (**E**) Anti-CITFA3 ChIP experiments with non-induced cells or with cells in which *CITFA1* was silenced for one or two days. Occupancy by CITFA3 was determined by qPCR at *RRNA* and BES promoters and, as a control, at the β-/α-tubulin intergenic region (Tub). The percent precipitation was corrected by subtracting the percent precipitation from negative control precipitations using a comparable, non-specific immune serum. Each experiment including the negative control was carried out three times independently and differences in occupancy were statistically analyzed by a two-tailed, student's

(Figure III-6 legend cont.) *t*-test assuming equal variance. Two asterisks indicate *P*-values that are <0.01.

against CITFA2, CITFA6 and CITFA7, and a newly generated polyclonal anti-CITFA3 antibody (see **Figure III-6**). The results demonstrate that, besides a minor reduction of CITFA2 on day 2, the CITFA subunits analyzed exhibited robust expression during the two-day-experiment suggesting that CITFA1, in contrast to CITFA7, is not required for overall complex integrity. To substantiate this notion, we analyzed the CITFA complex upon CITFA1 depletion by sucrose gradient sedimentation in which fractions were taken from top to bottom of the gradient (**Figure III-6B**). In non-induced cells the sedimentation profile was exactly as determined previously (Brandenburg et al., 2007; Nguyen et al., 2012): CITFA1, 3, 6, and 7 exhibited sedimentation peaks in fractions 13 and 14 while the less abundant CITFA2 peaked in fractions 14 and 15 shifting part of the complex one fraction down the gradient. Loss of CITFA1 decreased the sedimentation of CITFA2, 3, 6 and 7 by ~2 fractions (note that the sedimentation peak of the minor amount of detectable CITFA1 remained in fraction 13). However, the CITFA subunits still co-sedimented between the 6.6 S IgG marker (150 kDa) and the 230 kDa-large trypanosome TRF4-SNAPc-TFIIA transcription factor complex (Schimanski et al., 2005a). Since the CITFA complex without CITFA1 and without CITFA1 and 2 has a calculated mass of 221 kDa and 173 kDa, respectively, this result strongly indicates that CITFA1 depletion did not affect the integrity of the CITFA complex. Accordingly, a pull-down of CITFA3 in extract efficiently co-precipitated CITFA2, 6 and 7 in both non-induced and CITFA1-depleted cells (**Figure III-S3**).

The polyclonal anti-CITFA3 antibody was obtained by raising an immune serum in rats against a recombinant GST-CITFA3 fusion protein that was expressed in *Escherichia coli* and purified by glutathione affinity chromatography (data not shown), and by affinity purifying the antibody from serum with immobilized antigen. While the calculated molecular weight of CITFA3 is 47 kDa, the antibody recognized a single band of ~55 kDa in both BF and PF whole cell lysates (**Figure III-**

6C). This band is the correct band because in cell lysates of the PF line PTPC3ee, expressing exclusively CITFA3 with an N-terminally fused PTP tag, the band shifted up by ~20 kDa, the size of the tag. Hence, the polyclonal antibody detected CITFA3 with high specificity. We therefore used this antibody for two further assays. Firstly, we analyzed CITFA3 localization in non-induced cells and in cells in which *CITFA1* was silenced for 42 hours because there was a possibility that CITFA1 directs the CITFA complex to the nucleus or, within the nucleus, to the nucleolus and the ESB. We analyzed 70 randomly selected non-induced cells and 83 *CITFA1*-silenced cells in detail. In all cells CITFA3 exhibited subnuclear localization. In DAPI staining, the nucleolus becomes clearly visible as a spherical area of low DNA density resulting in a fainter DAPI stain (Daniels et al., 2012). In 93% of non-induced cells and in 86% of *CITFA1*-silenced cells, the CITFA3 signal was confined to the nucleolus (**Figure III-6D**). We also detected an additional, smaller extranucleolar spot in 11% and 6% of non-induced and induced cells, respectively (**Figure III-6D**). This spot is likely the ESB, since we have shown previously that CITFA7 reliably colocalized with RNA pol I in an extranucleolar compartment of similar size and signal intensity (Nguyen et al., 2014; Nguyen et al., 2012). Since the vast majority of cells in this analysis exhibited CITFA3 localization in the nucleolus independent of the CITFA1 knockdown, it appears that CITFA1 is not required for localizing the complex to the sites of RNA pol I transcription.

Secondly, we conducted an anti-CITFA3 ChIP assay in non-induced and *CITFA1*-silenced cells (**Figure III-6E**). For the PCR analysis of precipitated DNA, we used consensus oligonucleotide pairs that recognize either all copies of the *RRNA* promoter or all BES promoters (Park et al., 2011). The antibody effectively precipitated *RRNA* promoter DNA whereas BES promoter DNA was enriched 7.8 fold less than the *RRNA* promoter. This was expected because we have recently obtained a similar result in anti-CITFA7 ChIPs due to the fact that CITFA

predominantly binds the promoter of the active BES and occupies promoters of silent BESs to a much lesser extent. This is in contrast to *RRNA* promoters which, according to a ChIP-seq analysis, appear to be generally occupied by CITFA (Nguyen et al., 2014). Hence, in this assay, most if not all *RRNA* promoters were precipitated while BES promoter enrichment was mainly restricted to the active BES. Independent of the enrichment efficiency, *CITFA1* silencing reduced CITFA3 occupancy of both promoter types. For the efficient *RRNA* precipitation, this reduction was highly significant (**Figure III-6E**). Since we showed that CITFA1 has no role in the formation of a stable CITFA complex or in localizing CITFA to the nucleolus, these results strongly indicate that CITFA1 has a specific function in binding of the complex to *RRNA* and BES promoters.

III-4. Discussion

We have established a system in BF *T. brucei* for specific and efficient gene silencing that is based on targeting a HS that is fused to the mRNA of a gene of interest. We have established two such HSs, namely the 3′ UTR of the *T. cruzi* U2AF35 gene and the coding sequence of the large PTP tag. The system is based on BF cell lines smTc3 and smPTP which inducibly express Tc3 and PTP hairpin RNAs, respectively. Two consecutive transfections are required to fuse the HS to the gene of interest in one allele and to eliminate the remaining allele. Although these transfections are time consuming, this system comes with distinct benefits. Since doxycycline did not alter the rate of proliferation of smTc3 and smPTP cell lines (**Figure III-3**) or affect trypanosome morphology as observed by light microscopy, it can be inferred that Tc3 and PTP dsRNA-derived siRNAs do not target genes that are important for trypanosome culture growth. Furthermore, since both of these cell lines have the regulatable stem-loop vector already integrated, an *RRNA*-specific position effect is highly unlikely. Moreover, the vectors pCITFA7-HA-Tc3-BLA (**Figure III-S1**) and

pCITFA7-PTP-BLAv2 (**Figure III-S2**) offer a straightforward cloning strategy for fusing the HS to the gene of interest. C-terminal coding sequences can be amplified and inserted into the ApaI and NotI restriction sites in a single step. The only requirement for successful targeting of the plasmid to an endogenous allele is a restriction site within the gene coding sequence that is surrounded by at least 100 bp of coding sequence on either side and can be used to linearize the plasmid. Since this system is based on targeting the same sequence independent of the gene of interest, we anticipate that it can provide unambiguous gene silencing data in those cases where targeting endogenous sequences was not successful or inefficient. Finally, it is likely that the system can be implemented in PF 29-13 cells as well. However, the knockout of wild-type alleles in 29-13 cells would have to be accomplished with the *PURO* marker instead of the *HYG* marker as shown in **Figure III-1**, because 29-13 cells already harbor *HYG*.

The system also has its limitations, however. Firstly, it can only be applied to single copy genes because it is obligatory to produce a cell line that exclusively expresses the mRNA of interest as a HS fusion and no wild-type mRNA. Secondly, the system, as presented here, will not function if the expression of both alleles is required for trypanosome viability and proliferation. It should be noted, though, that we have so far generated many viable BF and PF cell lines which expressed various essential nuclear proteins exclusively as PTP fusions from a single allele. This has included RNA pol subunits, various transcription and RNA splicing factors, and kinases, suggesting that haplo-insufficiency rarely affects trypanosome viability in culture (see for example references (Badjatia et al., 2013; Lee et al., 2010; Li et al., 2008; Luz Ambrosio et al., 2009; Nguyen et al., 2007). If haplo-insufficiency is a concern, it may be possible to integrate one HS plasmid into each endogenous allele instead of knocking one allele out. For this, the blasticidin marker gene has to be replaced with the puromycin marker which can be achieved by a single PCR amplification and

cloning step, since either the selectable marker cassettes or the coding region can be excised by restriction digests (**Figures III-S1** and **III-S2**). Thirdly, the HS may negatively affect the expression of the mRNA or the functionality of the resulting protein. If the tag interferes with the function of the protein there are two options. The PTP tag can be fused to the N-terminus with our published pN-PURO-PTP vector (Schimanski et al., 2005a). Although in the N-terminal PTP tag the ProtA and ProtC domains have a different N- to C-terminal sequence order, the PTP dsRNA produced in smPTP cells effectively silenced a *CITFA2* gene in which the PTP tag sequence was inserted after the initiation codon (JK Kirkham and A Günzl, unpublished data). Alternatively, Tc3 can be employed without the HA tag, which was used here to enable specific detection of the protein translated from the fusion mRNA. In this case the 27 bp-long HA coding sequence needs to be removed from pCITFA7-HA-Tc3-BLA. Furthermore, if the gene of interest produces a mRNA whose regulation through its 3′ UTR is critical for cell viability, then replacing it either with Tc3 or with the *T. brucei* RPA1 3′ UTR in pC-PTP-BLA may be deleterious. In this case tagging the protein N-terminally with the PTP tag or replacing the RPA1 3′ gene flank with that of the gene of interest in pC-PTP-BLA may be a solution.

Recently, a conditional gene knockout system has been established in *T. brucei* that employs Cre recombinase and loxP sites (Kim et al., 2013a; Kim et al., 2013b). As with our approach, this system requires the manipulation of both gene alleles, e.g. it also depends on two consecutive transfection steps. This system offers instantaneous removal of the gene of interest in the genome and an unambiguous assessment of gene essentiality. However, in contrast to our system, this approach is irreversible, preventing the analysis of temporary gene knockdowns or the titration of the gene silencing level.

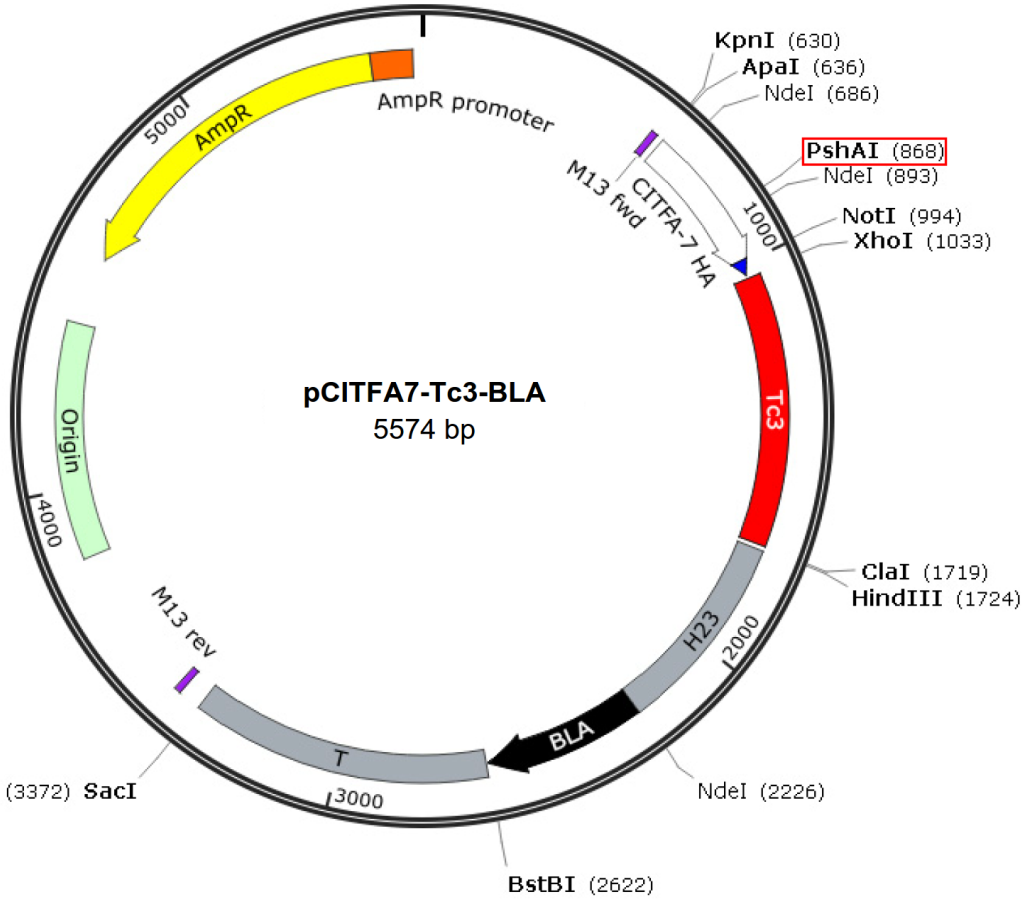
Finally, we have successfully employed this new gene silencing approach to evaluate CITFA1's role in the CITFA complex. We could unambiguously show that CITFA1 is essential for BF viability in culture and, as expected, has a vital role in rRNA and VSG mRNA expression. After CITFA2 and CITFA7, this is the third CITFA subunit whose knockdown led to rapid trypanosome death in culture, underscoring the indispensability of this transcription initiation factor for trypanosome viability. Furthermore, we showed that CITFA1 depletion strongly reduced CITFA occupancy at *RRNA* and BES promoters. Since we could not detect mislocalization of CITFA3 or the loss of CITFA subunits that are indicative of complex disruption as a consequence of *CITFA1* silencing, it appears that CITFA1 has a direct role in binding to promoter DNA. Most interestingly, CITFA3 remained localized to the nucleolus/ESB after CITFA1 knockdown, which suggests that subnuclear CITFA localization is not mediated by its binding to DNA. This finding further supports a recently described model in which DNA-independent concentration and confinement of CITFA to the nucleolus and ESB restricts maximal transcription initiation by RNA pol I to these compartments (Nguyen et al., 2014).

Acknowledgements

We are grateful to Dr. Barbara Burleigh (Harvard University) for providing *T. cruzi* genomic DNA. This work was supported by grants R01 AI059377 and R01 AI AI073300 to A.G. from the National Institute of Allergy and Infectious Diseases (NIAID) of the U.S. National Institutes of Health (NIH).

III-5. Supplemental

A



B

CACCTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAA
ATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAA
AGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTT
TTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAAC
GTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA
CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCG
GGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCA
CGACGTTGTAAAACGACGGCCAGTGAATTGGGTACCGGGCCCAGCAACCTTGCCTGCTATATTACTTGATACCTGGTGATGACGT
GTCGCATATGGCAGTGCCGAACGCTCGCGGGGAGTACCCGAACCTCATCCCTGGCATGCCCTACCTTGAGGAACTGTTACTGGAG
TTAGACAATGAGGAGGAAATTGAGGACAACGCCGTGCTTCCGTTGGTAGAATACGCACCGCCACCGCTCACGGCCGACGCAGGAG
TGCAAGTTGAGTGGACCGACGTCCCCATCCCCCCGCCACCGCATATGATGCGATATGTGAAACCATTGACGGAAACGACACAATC
AGTTACAGGGAGTGATTGCTCACCCTCTACAACACAGGAGTTTGGTGGTGAGGACGTGCGGCCGCTGTACCCTTACGACGTACCT
GACTACGCTTAACTCGAGCGTGCGGGGGAACAAAAGGCAAAAGTAACAACAAGAAGGATTTATTTTAAAAATCGGAGTGTGTTTT
TTCCCAATAACAGGAACAGCTGCACCAAGAAAAGGAAACGCCTTAATGTTTACGCTGTGGACATGTGAGCATGACCCCGTTGATG
AAACGTTATTTCCAATGCACTTGAGTTTCCTCAATGGGGAGGATGCATGAGTGCTTATTATTGGCAGTCTAGAAAGGGGGCTCTC
ACACGCTTCCAAATTCAATTCTGCAATCTGCTTTCTGTAACAATGGCAGAAAGAAAAAAAAAAAAATGATTCTGATACTGTTGAT
GGAATAAATGAGAGGCTGTTCGTAAATAAAGACTTTGCAGATCTTCTTCTTTTCAAGGGGCCTGGGAGAGGCAAAAAAAAAAAA

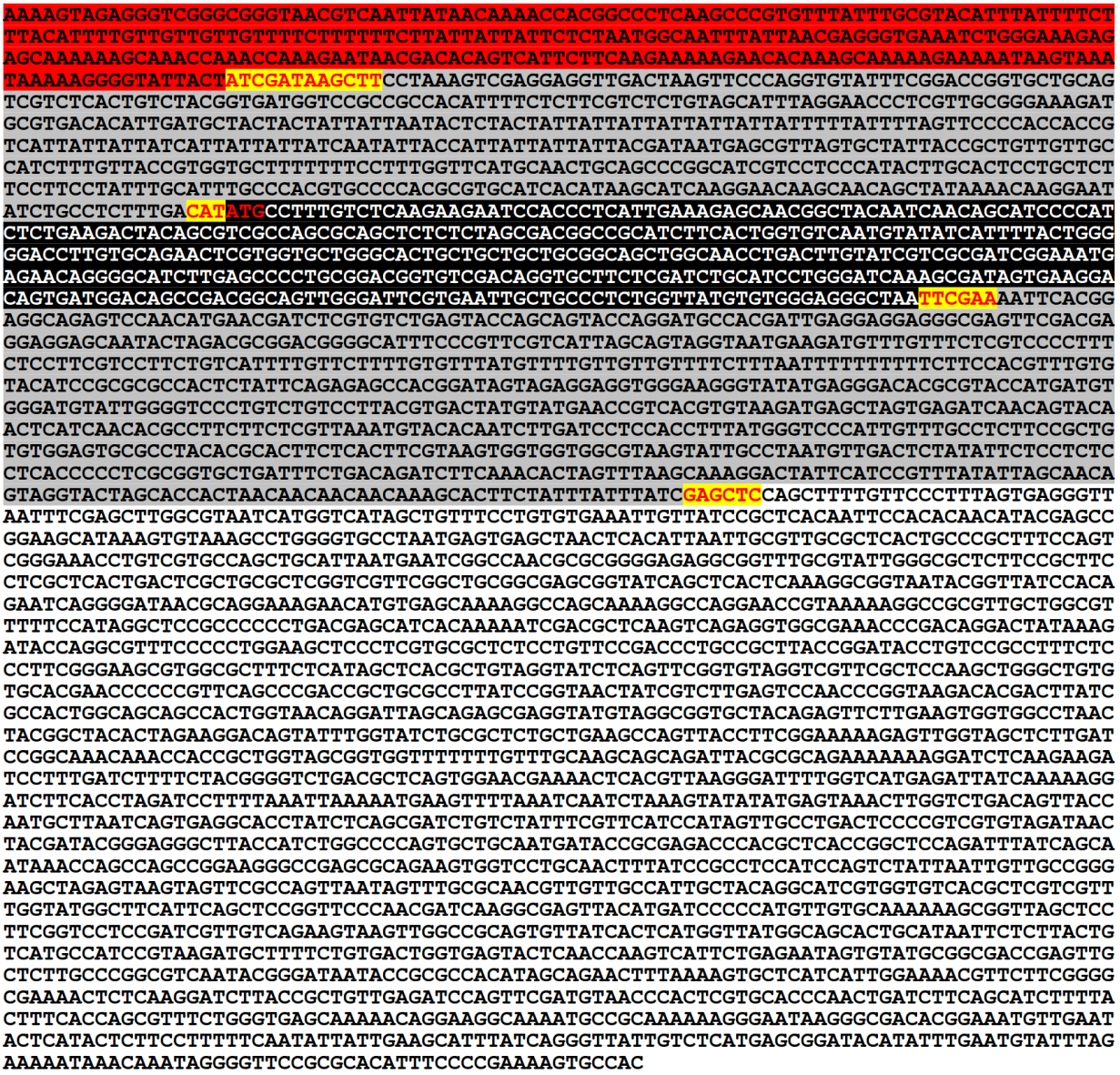
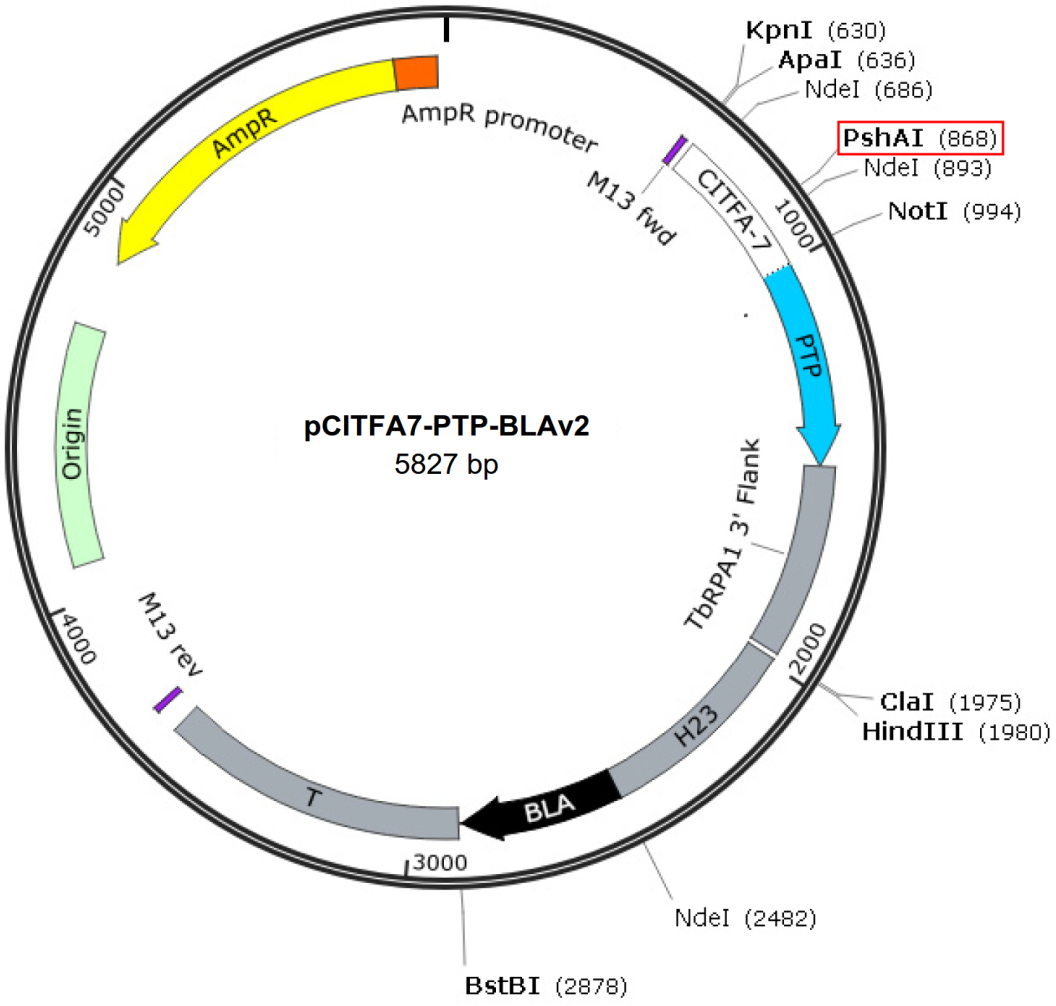


Figure III-S1. Plasmid CITFA7-HA-Tc3-BLA. (**A**) Plasmid map drawn to scale. The plasmid is a derivative of pBluescript II SK+. Two gene cassettes were introduced in tandem. The first cassette comprises the *Trypanosoma brucei* CITFA7 C-terminal coding region (white box) fused to the HA tag (blue), followed by a stop codon and the *Trypanosoma cruzi* U2AF35 3′ gene flank (Tc3; red). The second cassette harbors the coding region of blasticidin-S deaminase (BLA, black) flanked by the intergenic region of the *T. brucei* heat shock protein 70 genes 2 and 3 (H23, gray) (Lee, 1996) and the *T. brucei* β-/α-tubulin intergenic region (T, gray). Note that the vector was

targeted for integration into the CITFA7 gene by linearizing it inside the CITFA7 sequence with PshAI. To use this vector for other genes, the CITFA7 sequence needs to be replaced with the C-terminal coding sequence of a gene of interest using the KpnI/ApaI and NotI restriction sites. This sequence must also have a linearization site which should be surrounded by a minimum of 100 bp of gene sequence on either side. Furthermore, if the new sequence does not contain an NdeI restriction site, as the CITFA7 sequence does, then, if desired, the BLA resistance marker may be replaced using the NdeI and BstBI restriction sites. (**B**) Nucleotide sequence of the plasmid. Highlighting of inserted sequences corresponds to the map. The CITFA7 sequence is italicized. Important restriction sites and the stop codon following the HA sequence were highlighted in yellow and pink, respectively.

A



B

CACCTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAA
ATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAA
AGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTT
TTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAAC
GTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCA
CACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCG
GGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCA
CGACGTTGTAAAACGACGGCCAGTGAATTG**GGTACCGGGCCC***AGCAACCTTGCCTGCTATATTACTTGATACCTGGTGATGACGT*
GTCGCATATGGCAGTGCCGAACGCTCGCGGGGAGTACCCGAACCTCATCCCTGGCATGCCCTACCTTGAGGAACTGTTACTGGAG
TTAGACAATGAGGAGGAAATTGAGGACAACGCCGTGCTTCCGTTGGTAGAATACGCACCGCCACCGCTCACGGCCGACGCAGGAG
TGCAAGTTGAGTGGACCGACGTCCCCATCCCCCCGCCACCGCATATGATGCGATATGTGAAACCATTGACGGAAACGACACAATC
*AGTTACAGGGAGTGATTGCTCACCCTCTACAACACAGGAGTTTGGTGGTGAGGAC*GT**GCGGCCGC**TGGAAGATCAGGTGGATCCT
CGTCTTATTGATGGGAAATATGATATTCCAACTACTGCTAGCGAGAATTTGTATTTTCAGGGTGAGCTCAAAACCGCGGCTCTTG
CGCAACACGATGAAGCCGTGGACAACAAATTCAACAAAGAACAACAAAACGCGTTCTATGAGATCTTACATTTACCTAACTTAAA
CGAAGAACAACGAAACGCCTTCATCCAAAGTTTAAAAGATGACCCAAGCCAAAGCGCTAACCTTTTAGCAGAAGCTAAAAAGCTA
AATGATGCTCAGGCGCCGAAAGTAGACAACAAATTCAACAAAGAACAACAAAACGCGTTCTATGAGATCTTACATTTACCTAACT
TAAACGAAGAACAACGAAACGCCTTCATCCAAAGTTTAAAAGATGACCCAAGCCAAAGCGCTAACCTTTTAGCAGAAGCTAAAAA
GCTAAATGGTGCTCAGGCGCCGAAAGTAGACGCGAATTCCGCGGGGAAGTCAACC**TGA**TCGCCTGAACCGCTATTTTAAGTATAT
ATATATGTATATTACCTTTTTTTGTTTTTTCTTTGCGCTTTTACCTCCTCTTTTGTCTCAACTTTTTTTAATTTTATGAGAGGCA
CTTGTGTGAAATTGGACACCATAAAGGGCTGAGGAGGGAAGCAGAGAGCAAATGAAAATATGGGAGGGTGAAAAATACATACATA
TATATTTGTGTTTGTTCGTATTTGGAGGGATACAAATACGATATATGAAGATATCAACAGTTTTTCACCTGTGGGGGAAGCGTGA
TTATTATCAGTGAGTGGATGGACTGAAGAACGAATAAATGTGTAACAATTTCGCCCTTCATTTTTTCTTCGGTGGCAGACGGGTT

Figure III-S2. Plasmid CITFA7-PTP-BLAv2. (**A**) Plasmid map drawn to scale. The plasmid corresponds to pCITFA7-HA-Tc3-BLA described above except for the tagging cassette in which the CITFA7 sequence (white) is fused to the PTP tag sequence (cyan) via a NotI restriction site, followed by a stop codon and the *T. brucei* RPA1 (accession number Tb927.8.5090 at www.TriTrypDB.org) 3/ gene flank (gray). (**B**) Nucleotide sequence of the plasmid. The RPA1 sequence is presented in white lettering with dark gray highlighting to distinguish it from the H23 and T sequences.

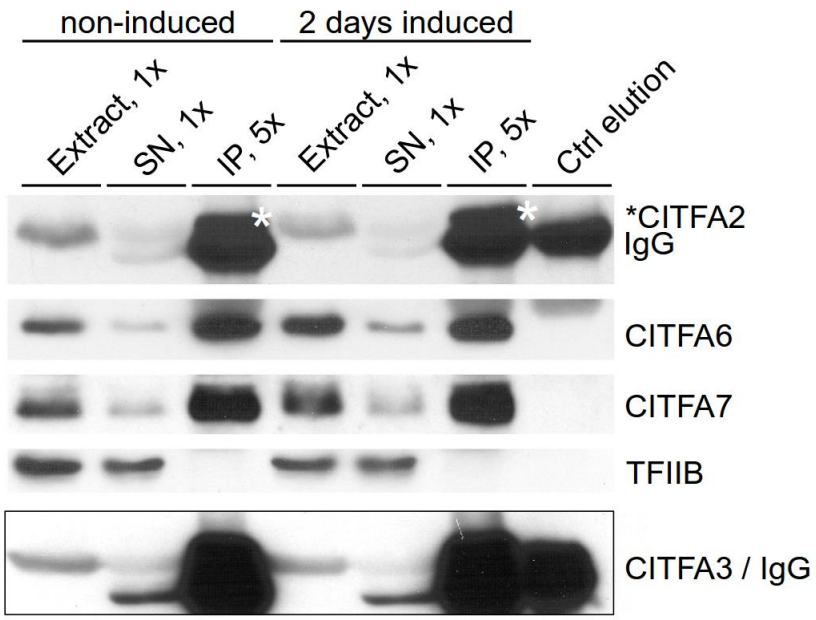


Figure III-S3. Anti-CITFA3 co-immunoprecipitation. CITFA3 was immunoprecipitated in extract prepared from non-induced bloodstream form trypanosomes or from cells in which *CITFA1* was silenced for 2 days. Extract, supernatant (SN) and immunoprecipitate (IP) were analyzed by immunoblotting. x-Values indicate relative amounts loaded. A control (Ctrl) elution of protein G bead-coupled antibodies was carried out to detect IgG contamination. CITFA2, CITFA6, CITFA7 and, as a negative control, TFIIB were detected with polyclonal immune sera on the same blot. Asterisks indicate the CITFA2 band just above the IgG contamination. Although CITFA3 was detected in a different gel that better separates proteins in the 50 kDa range, the CITFA3 signal in the precipitate was not distinguishable from that of the IgG heavy chain. Nonetheless, this experiment demonstrates that *CITFA1* silencing did not affect the association of CITFA2, CITFA6, and CITFA7 with CITFA3 indicating that the CITFA complex remained intact upon CITFA1 depletion.

Chapter IV

Transcription by the multifunctional RNA polymerase I in *Trypanosoma brucei* functions independently of RPB7

Abstract

Trypanosoma brucei has a multifunctional RNA polymerase (pol) I that transcribes ribosomal gene units (*RRNA*) and units encoding its major cell surface proteins variant surface glycoprotein (VSG) and procyclin. Previous analysis of tandem affinity-purified, transcriptionally active RNA pol I identified ten subunits including an apparently trypanosomatid-specific protein termed RPA31. Another ortholog was identified *in silico*. No orthologs of the yeast subunit doublet RPA43/RPA14 have been identified yet. Instead, a recent report presented evidence that RPB7, the RNA pol II paralog of RPA43, is an RNA pol I subunit and essential for *RRNA* and *VSG* transcription in bloodstream form trypanosomes (Penate *et al.*, 2009, EMBO Rep. 10:252-257). Revisiting this attractive hypothesis, we were unable to detect a stable interaction between RPB7 and RNA pol I in either reciprocal co-immunoprecipitation or tandem affinity purification. Furthermore, immunodepletion of RPB7 from extract virtually abolished RNA pol II transcription *in vitro* but had no effect on *RRNA* or *VSG* ES promoter transcription in the same reactions. Accordingly, chromatin immunoprecipitation analysis revealed cross-linking of RPB7 to known RNA pol II transcription units but not to the *VSG* ES promoter or to the 18S rRNA coding region. Interestingly, RPB7 did crosslink to the *RRNA* promoter but so did the RNA pol II-specific subunit RPB9 suggesting that RNA pol II is recruited to this promoter. Overall, our data led to the conclusion that RNA pol I transcription in *T. brucei* does not require the RNA pol II subunit RPB7.

IV-1. Introduction

Trypanosoma brucei has a multifunctional RNA polymerase (pol) I that transcribes the large ribosomal (*RRNA*) gene unit in all life cycle stages, and, uniquely, the telomeric expression site encoding the active variant surface glycoprotein (VSG) gene in its bloodstream form (BF) as well as gene units encoding procyclin in its procyclic form (PF). BFs evade the mammalian immune response by antigenic variation of their cell surface coat consisting of ~10 million identical VSG molecules expressed from a single gene. The variation occurs when parasites switch to the expression of a different *VSG* gene drawn from a large *VSG* gene repertoire (recently reviewed in (Horn and McCulloch, 2010)). Expression of the whole VSG coat from a single gene requires extremely high expression levels and it has been determined that the transcription rate of the active *VSG* gene is approximately 50 times higher than that of a β tubulin gene (Ehlers et al., 1987). Such high rates are the hallmark of RNA pol I transcription which in general accounts for more than 50% of the transcriptional activity in a eukaryotic cell (Russell and Zomerdijk, 2005; White, 2008). However, utilization of RNA pol I for protein coding gene expression requires a deviating mode of gene expression. In the mouse, strong *RRNA* promoter-driven transcription of a reporter construct, led to reporter enzyme activities which were 20-50fold lower than in a control experiment with an RNA pol II promoter (Grummt and Skinner, 1985). Conversely, *RRNA* promoter-driven expression of a selectable marker gene in *T. brucei* increased parasite resistance around 30fold over RNA pol II-mediated expression of the same gene demonstrating that, in trypanosomes, RNA pol I can efficiently synthesize functional mRNA (Rudenko et al., 1991; Zomerdijk et al., 1991a). The opposite outcomes in these studies are most likely due to different modes of mRNA capping. In the budding yeast *Saccharomyces cerevisiae* and higher eukaryotes, RNA capping is co-transcriptional and specifically linked to RNA pol II because the capping enzymes bind to the phosphorylated carboxy-terminal domain (CTD) of the largest RNA pol II

subunit RPB1 (Bentley, 2005). Conversely, in trypanosomes and related organisms, protein-coding gene transcription is polycistronic and individual mRNAs are processed from precursors by spliced leader (SL) *trans* splicing and polyadenylation. Since in *trans* splicing the capped SL, comprising the 5′-terminal part of the small nuclear SL RNA, is transferred to the 5′ end of each mRNA, this process constitutes a post-transcriptional capping mechanism that decouples capping from RNA pol II transcription (Günzl, 2010).

Eukaryotic RNA pols I-III consist of twelve subunits which are either shared or paralogous to each other. In addition, yeast RNA pol I contains two RNA pol I-specific subunits, RPA49 and RPA34, which are not essential for yeast proliferation. The multifunctional nature of trypanosome RNA pol I has spurred the investigation of this enzyme. While the largest two subunits, RPA1 and RPA2, were discovered first (Jess et al., 1989; Schimanski et al., 2003; Smith et al., 1989), eight of the remaining ten core subunits could be identified bioinformatically after the *T. brucei* genome was completed (Kelly et al., 2005). The missing subunits were the orthologs of yeast RPA43 and RPA14. Since these two subunits form a functional doublet in yeast and humans, it was proposed that trypanosome RNA pol I may assemble the paralogous RNA pol II subunits RPB7 (GeneDB/TritrypDB accession number Tb11.01.6090) and RPB4 (Tb927.3.5270) instead (Kelly et al., 2005), possibly to aid this polymerase in the synthesis of functional mRNA. This was an attractive hypothesis because the respective genes were readily detected in trypanosomatid genomes and because RPB4/7 analysis in other systems indicated specific functions of this protein doublet in RNA synthesis such as binding of an RNA processing factor (Mitsuzawa et al., 2003), direct RNA interaction (Újvári and Luse, 2005), and linking mRNA synthesis to mRNA decay (Lotan et al., 2007). Moreover and most recently, the RPB4/7 doublet was characterized as a potential “mRNA coordinator” in yeast that can be deposited on mRNA facilitating efficient

translation and thereby linking expression outcomes from gene transcription to translation (Harel-Sharvit et al., 2010).

Based on co-immunoprecipitation assays, *RPB7* expression silencing experiments, co-localization of RPB7 and the RNA pol I-specific subunit RPB6z (Tb11.03.0935), and *in vitro* transcription assays in BFs, Penate *et al.* recently concluded in their publication title that "*RNA pol II subunit RPB7 is required for RNA pol I-mediated transcription in* Trypanosoma brucei" (Peñate et al., 2009). However, tandem affinity purification of trypanosome RNA pol I in two different laboratories did not identify RPB7 as a co-purifying subunit (Nguyen et al., 2007; Nguyen et al., 2006; Walgraffe et al., 2005). Moreover, isolation of RNA pol I from PF extract that was active in both non-specific and promoter-dependent transcription assays did not reveal a protein band of ~20 kDa which is the apparent size of *T. brucei* RPB7 (Nguyen et al., 2007). This was of particular concern because RPB7 and its RNA pol I and III paralogues RPA43 and RPC25 were shown in yeast to be essential for promoter-dependent transcription initiation (Edwards et al., 1991; Peyroche et al., 2000; Zaros and Thuriaux, 2004). We therefore revisited the role of *T. brucei* RPB7 in RNA pol I transcription omitting *RPB7* expression silencing which may affect gene expression independently of RNA pol function or may rapidly lead to secondary defects through a general shut-down of RNA pol II transcription.

IV-2. Methods

DNAs

For the generation of BF cell lines that exclusively expressed RPB7 or RPB9 with a C-terminal fusion of the composite PTP tag sequence, encoding tandem protein A (ProtA) domains,

a tobacco etch virus (TEV) protease site and the protein C epitope (ProtC), plasmids RPB7-PTP-NEO and RPB9-PTP-NEO were first integrated into endogenous alleles. The plasmids were generated by inserting, respectively, 549 bp and 385 bp of the *RPB7* and *RPB9* (Tb11.02.5180) C-terminal coding regions into pC-PTP-NEO (Schimanski et al., 2005b) using ApaI and NotI restriction sites. For pRPB9-PTP-NEO linearization, an AflII restriction site was engineered into the RPB9 sequence creating a silent mutation in *RPB9* codon 58. In a second round of stable transfections, the remaining wild-type *RPB7* and *RPB9* alleles were knocked out by transfection of 10 µg of a chimeric linear DNA in which 100 bp-long gene flanks were fused to the coding region of the hygromycin phosphotransferase by standard PCR as described previously (Arhin et al., 2004). pRPB7-HA-BLA was created by pasting the *RPB7* target sequence from pRPB7-PTP-NEO into the published plasmid pRPA31-HA-BLA (Nguyen et al., 2007). For transfections, pRPB7-PTP-NEO/pRPB7-HA-BLA, pRPB9-PTP-NEO, and pPTP-RPB6z-PURO were linearized with restriction enzymes PmeI, AflII, and BmgBI, respectively. pPURO-PTP-RPB6z and transcription templates Rib-trm, VSG-trm, and SLins19 were described previously (Laufer et al., 1999; Nguyen et al., 2007).

The following oligonucleotide pairs were used in semi-quantitative and quantitative (q)PCR analyses of DNA obtained in chromatin immunoprecipitation (ChIP) assays: consensus *VSG* ES promoter, 5′-TCTAAAAGAATCATATCC-3′/5′-AAGCGTAGATGAGATTAAAGTC-3′; consensus *RRNA* promoter, 5′-AATACAACACACAATAGG-3′/5′-GTCTGAGAGCGGTCAGTTGC-3′; 18S rRNA coding region; 5′-TCATCAAACTGTGCCGATTAC-3′/5′-CTATTGAAGCAATATCGG-3′; *SLRNA* promoter, 5′-CTACCGACACATTTCTGGC-3′/5′-GCTGCTACTGGGAGCTTCTCATACC-3′; β-α tubulin

intergenic region, 5′-GCTGATTTCTGACAGATCTTCAAAC-3′/5′-GTGGATGCAGATAGCCTCACGCATG-3′.

Cells

PF and BF of *Trypanosoma brucei brucei* strain 427 were cultured as described previously (Günzl et al., 2000; Hirumi and Hirumi, 1989). Transfected cells were cloned by limiting dilution immediately after electroporation. BFs were selected with 2.5 µg/ml of G418, 1 µg/ml of hygromycin, and 2 µg/ml of blasticidin whereas PFs were selected with increasing antibiotic concentrations ranging from 15 to 40 µ/ml of G418 and from 10 to 20 µg/ml of hygromycin. DNA integrations were confirmed in each clonal cell line by PCR of total DNA using oligonucleotides that hybridize outside the cloned region. Protein tagging for each clonal cell line was confirmed by immunoblotting.

Protein Analysis

Tandem affinity purification of PTP-tagged RPB7 (RPB7-PTP) was carried out exactly as specified in the standard protocol (Schimanski et al., 2005b). Purified proteins were separated on a 10–20 % SDS–polyacrylamide gradient gel and detected with Pierce Gelcode Coomassie blue stain. The same protocol was used for the preparation of active RNA pol II, with the PTP tag fused C-terminally to subunit RPB9. However, in this purification the enzyme was kept on the anti-ProtC beads and not eluted. The beads were stored at -20°C until used in *in vitro* transcription assays.

Co-immunoprecipitation assays were carried out with 100 µl of crude cell extract corresponding to 4×10^8 BF cells. PTP-RPB6z was precipitated with 40 µl settled volume of human IgG beads (GE Healthcare) which bind the ProtA domains of the PTP tag. RPB7-HA was precipitated with a rat monoclonal anti-HA antibody (Roche) that we bound to paramagnetic

protein G beads (GE Healthcare) to avoid a non-specific interaction of this antibody with the ProtA domains. After one hour on ice, the supernatant was taken off and the precipitate washed seven times with 700 µl of TET100 buffer (100 mM NaCl, 20 mM Tris pH 8.0, 3 mM $MgCl_2$, 0.1% Tween 20). Precipitated PTP-RPB6z was eluted by resuspending the beads in 45 µl of TEV protease buffer (150 mM KCl, 20 mM Tris pH 7.7, 3 mM 1M MgCl2, 0.5 mM EDTA, 1 mM dithiothreitol [DTT], 0.1 % Tween 20) containing 40 units of AcTEV protease (Invitrogen) and by an incubation of the protease digest at 28°C for 30 min. Eluted proteins were mixed with 15 µl of 4x standard SDS loading buffer and boiled for 10 min. Anti-RPB7-HA precipitates were directly released into SDS loading buffer.

In immunoblots, PTP-tagged proteins were probed with the mouse monoclonal anti-ProtC antibodyHPC4 (Roche) in the presence of 1 mM $CaCl_2$ and HA-tagged proteins with a rat monoclonal anti-HA antibody (Roche). Protein detection was achieved with peroxidase-labeled anti-mouse or anti-rat IgG secondary antibody (Vector Laboratories) in combination with the BM chemiluminescence blotting substrate (Roche) according to the manufacturer's protocol. Previously published antisera were used to detect the two largest RNA pol I subunits RPA1 and RPA2 (Schimanski et al., 2003) whereas for RPB1 detection a polyclonal anti-RPB1 immune serum was raised in rat against an *Eschericha coli*-expressed GST fusion protein comprising RPB1 residues 1424 to 1766 according to a published procedure (Brandenburg et al., 2007).

Chromatin Immunoprecipitation

For each ChIP experiment, 1 x 10^8 bloodstream trypanosomes were fixed with formaldehyde and washed as published (Lee et al., 2006). Cells were sonicated in a Bioruptor UCD-200 (Diagenode) for a total of 25 min (30 s on/30 s off) at 4 °C, with pulse setting on “high”. Following

a 10-min centrifugation at 25,000g and 4 °C to pellet cell debris, the supernatant was pooled and chromatin pre-cleared with antibody-free, bovine serum albumin-blocked ProtA Dynabeads (Invitrogen) for 1 hour at 4 °C while rotating. For immunoprecipitation of PTP-tagged proteins, 0.75 ml of chromatin solution was incubated overnight at 4 °C with either a rabbit polyclonal anti-ProtA antibody (Sigma) or a nonspecific rabbit immune serum as negative control. Sodium chloride was then added to a final concentration of 350 mM and the chromatin was captured with ProtA Dynabeads for 40 min at 4 °C. Subsequent washing of the beads and DNA preparation through RNase and proteinase K treatments were carried out as detailed previously (Lee et al., 2010). In a final step, the DNA was purified using a MinElute kit (Qiagen) and eluted in 20 µl of water.

The DNA was analyzed by standard qPCR assays using the SsoFast EvaGreen Supermix (BioRad) on a CFX96 cycler (BioRad) according to the manufacturer's recommendations. Between two and five independent ChIP experiments were conducted for each cell line. For each amplification and ChIP experiment, triplicate qPCR samples were analyzed using the Bio-Rad CFX Manager software package. Each amplification product was analyzed for specificity by both agarose gel electrophoresis and melt curve analysis. Standard curves for oligonucleotide pairs were derived from input DNA dilution series and ranged in their r^2 value from 0.98 to 1.0. Fold enrichment values were calculated as the ratio between starting quantities of positive and control precipitations.

In vitro *transcription*

The bloodstream *in vitro* transcription system will be described elsewhere in detail (manuscript in preparation). Briefly, BF cells were harvested from 6 liter of culture that was grown

to a cell density of 2 x 10^6 cells/ml. Cells were washed twice with Tryp wash solution (100 mM NaCl, 3 mM $MgCl_2$, 20 mM Tris–HCl, pH 7.5), equilibrated in transcription buffer (150 mM sucrose, 20 mM potassium l-glutamate, 3 mM $MgCl_2$, 20 mM HEPES–KOH, pH 7.7), and resuspended in 1.5 times the packed cell volume of transcription buffer containing 1 mM DTT, 10 µg/ml leupeptin, and 10 µg/ml aprotinin. After adding 20 µl of phosphatase inhibitor cocktail (Sigma) and ~200 µl settled volume of 800 µm low binding silica beads (OPS Diagnostics), cells were broken by five cycles of shock freezing in liquid nitrogen, thawing and vortexing for 2 min in a 4°C cold room. Broken cells were extracted by mixing the sample quickly with one tenth volume of buffer C (1500 mM KCl, 20 mM HEPES-KOH pH 7.7, 3mM $MgCl_2$) and by incubating the mix for 20 min on ice. After centrifugation at 25,000g and 2 °C for 10 min, the extract was separated from insoluble cell debris, diluted with 0.5 volumes of ice-cold transcription buffer, concentrated ~5x in a centricon-10 centrifugal filter device (Millipore), aliquoted, shock-frozen in liquid nitrogen, and stored at -80°C. For mock- or immuno-depletion of RPB7-PTP, 50 µl of BF transcription extract was mixed with transcription buffer-equilibrated protein G beads or IgG beads (settled volume of 30 µl), respectively, and incubated on ice for 1.5 hours. Beads were pelleted for 2 min. at 3000 g and 2 °C and extract was separated from beads, shock-frozen, and stored at -80 °C.

Standard transcription reactions were carried out in a volume of 20 µl containing 4 µl of cell-free extract, 20 mM potassium L-glutamate, 20 mM KCl, 3 mM $MgCl_2$, 20 mM HEPES-KOH, pH 7.7, 20 mM creatine phosphate, 0.48 mg/ml of creatine kinase, 2.5% polyethylene glycol, 0.2 mM EDTA, 0.5 mM EGTA, 4 mM DTT, 10 µg/ml leupeptin, 10 µg/ml aprotinin, 20 µg/ml RNA pol I promoter template (Rib-trm or VSG-trm), 7.5 µg/ml SLins19 template, 12.5 µg/ml unrelated plasmid DNA, 10 µg/ml leupeptin, and 10 µg/ml aprotinin. The reactions were pre-incubated on

ice for 10 min in the absence of nucleoside triphosphates (NTPs). After adding the NTPs to a final concentration of 0.5 mM, the reaction was incubated for 1 h at 27°C and stopped by adding Trizol solution (Invitrogen). Total RNA was prepared as described previously (Laufer et al., 1999). Newly synthesized RNAs were detected by primer extension of ^{32}P-end-labeled oligonucleotides Tag-PE and SLtag which hybridize to unrelated oligonucleotide tags of VSG-trm/Rib-trm and SLins19 RNAs, respectively. Primer extension products were resolved on 6% polyacrylamide-50% urea gels and visualized by autoradiography. In RNA pol II reconstitution assays, tandem affinity-purified RNA pol II, bound to anti-protein C matrix and equilibrated in transcription buffer, was added back to the reactions. The amounts corresponded to 0.35% and 0.7% of a standard PTP tandem affinity purification (Nguyen et al., 2007; Schimanski et al., 2005b).

Indirect immunofluorescence light microscopy

Immunolocalizations of procyclic cells were carried out as described previously (Luz Ambrosio et al., 2009).

IV-3. Results and Discussion

Interaction of RPB7 with RNA polymerases I and II in extract

In a first step, we wanted to confirm an interaction between RPB7 and the RNA pol I subunit RPB6z in a BF extract that was active in both accurate RNA pol I and II transcription (Laufer and Günzl, 2001; Laufer et al., 1999). To avoid any cross-reactivity of polyclonal antisera, we generated a cell line in which RPB6z was N-terminally tagged with the composite PTP tag (PTP-RPB6z; (Schimanski et al., 2005b)) and RPB7 C-terminally with the HA tag (RPB7-HA). These

proteins were then precipitated with antibodies directed against their tags. Since we have previously shown that PTP-RPB6z is fully functional (Nguyen et al., 2007) and since BF and PF cell lines which exclusively expressed C-terminally tagged RBP7 did not exhibit any growth defect (data not shown), we concluded that tagging of these two subunits did not interfere with their function. When we immunoprecipitated PTP-RPB6z from extract at low stringency conditions, the RNA pol I subunits RPA1 and RPA2 were efficiently co-precipitated whereas RPB7-HA, RPB1 and a non-specific control were not (**Figure IV-1A**, left panels). Since there was the possibility that only a small amount of RNA pol I was active and bound to RPB7, we increased the relative amount of precipitate from 4x to 50x but still could not detect a signal above the background found with the spliceosomal U2-40K protein (**Figure IV-1A**, lower left panels). Similarly, when we precipitated RPB7-HA with a monoclonal anti-HA antibody, we saw efficient co-precipitation of RPB1 but not of the RNA pol I subunits RPA1, RPA2 and PTP-RPB6z (**Figure IV-1A**, right panels). Only when the precipitate was increased to 50x, we detected a faint PTP-RPB6z signal but a band of similar strength was detected with the U2-40K control suggesting that this was the background level due to the low stringency conditions applied. But even when the background signal was not taken into account, densitometry revealed that RPB7 is bound to RNA pol I in transcriptionally active extract to less than 1%.

In a different approach investigating the interaction of RPB7 with RNA pols I and II, we tandem affinity-purified RPB7 complexes from extract prepared from PFs that exclusively expressed RPB7 as a C-terminal PTP tag fusion. Although protein tags can change the affinities of protein-protein interactions, the rationale here was that if RPB7 is an essential subunit for RNA

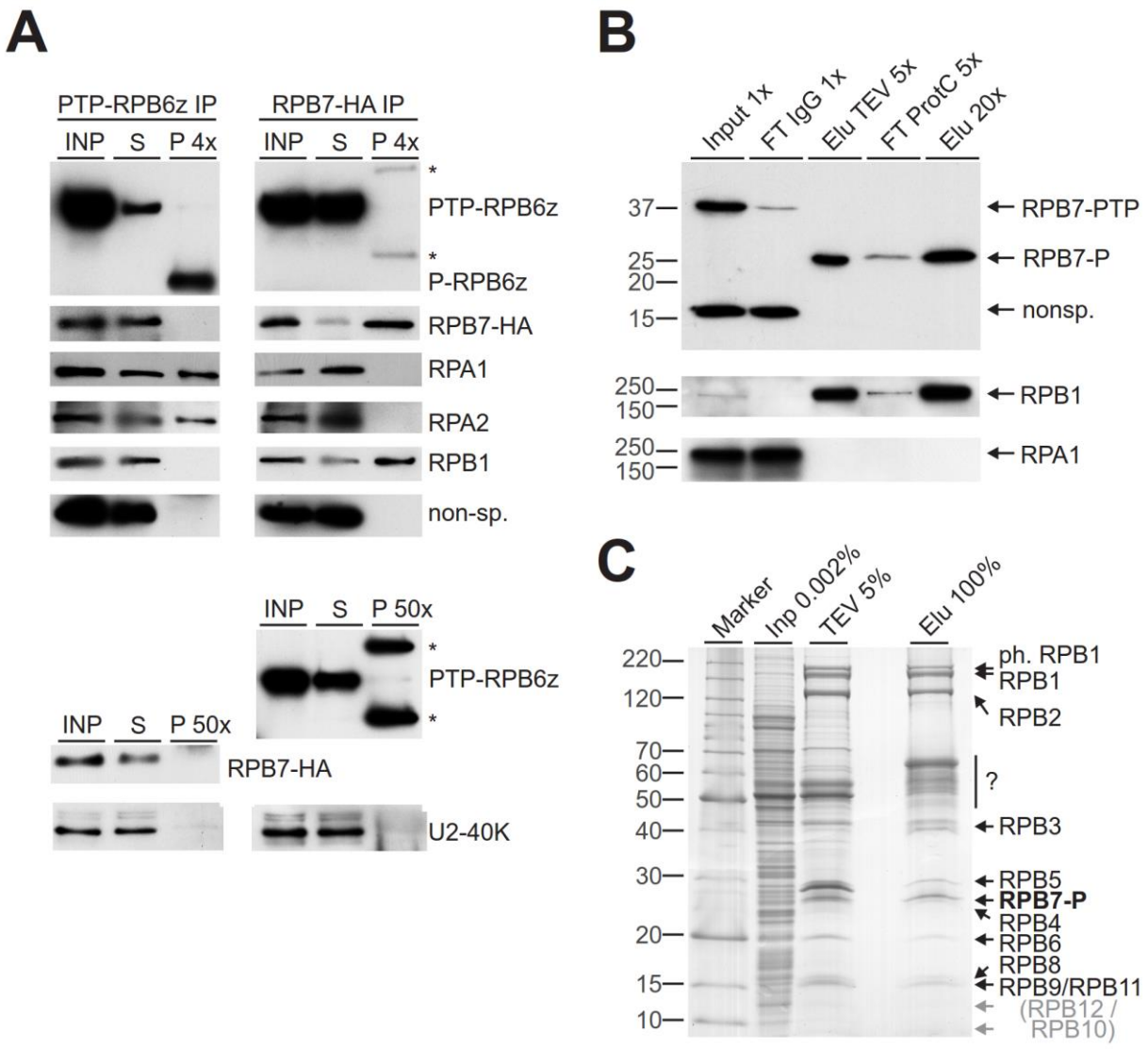


Figure IV-1. RPB7 does not detectably interact with RNA pol I in transcriptionally active extract. (**A**) Immunoblot analysis of PTP-RPB6z immunoprecipitation (PTP-RPB6z IP; left panels) and of RPB7-HA immunoprecipitation (RPB7-HA IP, right panels). For each immunoprecipitation, equivalent amounts of crude extract (Inp) and of supernatant (S) were analyzed and four or fifty times the relative amount of the precipitate (P). Asterisks indicate IgG heavy and light chain signals. Note the reduction in size of PTP-RPB6z to P-RPB6z upon TEV protease digest in the PTP-RPB6z immunoprecipitation. (**B**) Immunoblot analysis of TbRPB7-PTP purification. Crude extract (Inp), flow-through of the IgG sepharose column (FT-IgG), TEV protease eluate (TEV),

(Figure IV-1 legend cont.) flow-through of the anti-protein C matrix (FT-ProtC), and final eluate (Elu) were analyzed in relative amounts as indicated. Tagged RPB7 and endogenous RPB1 and RPA1 were successively detected on the same blot. (**C**) RPB7 co-purified proteins. The complete final eluate (Elu) of a standard RPB7-PTP purification was separated on a 10–20% SDS/polyacrylamide gradient gel and stained with Coomassie blue. For comparison, small aliquots of crude extract (Inp) and TEV protease eluate were co-analyzed. The band assignments on the right were according to previous RNA pol II purifications (Das et al., 2006; Devaux et al., 2006; Lee et al., 2009) and our own immunoblot results. RPB1 occurs in two bands, the upper one being phosphorylated (ph. RPB1). The question mark indicates several co-purified bands which have not been identified yet. The band pattern of RPB12 and RPB10 has not been determined yet.

pol I transcription then it should interact in a similar manner with this enzyme as it does with RNA pol II. The tandem affinity purification of RPB7-PTP by IgG chromatography and by anti-ProtC affinity chromatography efficiently concentrated RPB7 complexes in the final eluate (**Figure IV**-1B, top panel). RPB1 was enriched throughout the purification whereas RPA1 was only detected in crude extract and the flow-through of the IgG column demonstrating that RPB7 is a *bona fide* RNA pol II subunit but not assembled into a stable RNA pol I complex to a detectable extent (**Figure IV-1B**, middle and bottom panels, respectively). Accordingly, when the final eluate was separated by SDS-PAGE and stained with Coomassie blue, an RNA pol II-specific band pattern (Das et al., 2006; Devaux et al., 2006) was detected and not bands characteristic for RNA pol I (**Figure IV-1C**). We therefore concluded that in both BF and PF extracts that are active in RNA pol I and II transcription, tagged RPB7 was quantitatively and stably associated with RNA pol II but not with RNA pol I at our detection level.

3.2. Chromatin immunoprecipitation of RPB7

Although the RPB4/7 doublet can be deposited on mRNA (Harel-Sharvit et al., 2010), the genome-wide occupancy profile of RPB7 in yeast was congruent with that of the RNA pol II subunit RPB3 (Jasiak et al., 2008). Hence, If RPB7 is an essential component of RNA pol I, it should be found with RPB6z at RNA pol I promoters in ChIP experiments because it is unlikely that major amounts of inactive polymerase that lack RPB7 are recruited to these promoters. Since we had previously shown that a polyclonal anti-ProtA antibody is of ChIP grade that can enrich DNA-bound transcription factors more than 50fold over control precipitations with a non-specific immune serum (Lee et al., 2010) and since we wanted to directly compare occupancies of RPB6z, RPB7 and the essential, RNA pol II-specific subunit RPB9 with each other, we generated clonal BF cell lines that exclusively expressed these three essential proteins as PTP fusions (data not

shown). As expected, RPB6z was clearly enriched at *VSG* ES and *RRNA* promoters that recruit RNA pol I (**Figure IV-2**). The weaker enrichment of RPB6z at *VSG* ES promoters in comparison to *RRNA* promoters is most likely due to the fact that *RRNA* transcription is spread over several repeat units whereas only one of fifteen *VSG* ESs is actively transcribed in an extranucleoloar compartment (Navarro and Gull, 2001). RPB7 and RPB9 did not significantly cross-link to *VSG* ES promoters. In contrast, occupancy of RPB7 was clearly established at *RRNA* promoters albeit to a lesser extent than for RPB6z. Since the *RRNA* promoter was also occupied by RPB9 (**Figure IV-2**), this finding suggests that RNA pol II is recruited to this site rather than RPB7 being a subunit of RNA pol I. This notion was strengthened by high RPB6z occupancy and concomitant absence of RPB7 and RPB9 in the 18S rRNA coding region downstream of the *RRNA* promoter. The presence of RNA pol II at the *RRNA* promoter correlates well with the previously reported finding that the most distal domain of the *T. brucei RRNA* promoter resembles the bipartite upstream sequence element (USE) of the *SLRNA* promoter in opposite orientation and that this ribosomal USE, as its *SLRNA* counterpart, bound the small nuclear RNA-activating protein (SNAP)50 *in vitro* (Schimanski et al., 2004). Therefore, it is possible that binding of the SNAP complex to the distal domain of the *RRNA* promoter nucleates an RNA pol II transcription pre-initiation complex that recruits RNA pol II. This recruitment may be of functional significance because it was recently shown in the human system that noncoding RNA, derived from the *RRNA* promoter sequence, induces epigenetic regulation of *RRNA* transcription (Schmitz et al., 2010).

The absence or low occupancies observed with RPB7 and RPB9 at RNA pol I transcription units cannot be due to inefficient cross-linking to DNA because both proteins were nearly 60fold enriched at *SLRNA* promoters and clearly detectable in the β-/α- tubulin intergenic region (**Figure IV-2**). In sum, these ChIP experiments demonstrated that RPB7 occupancy closely paralleled the

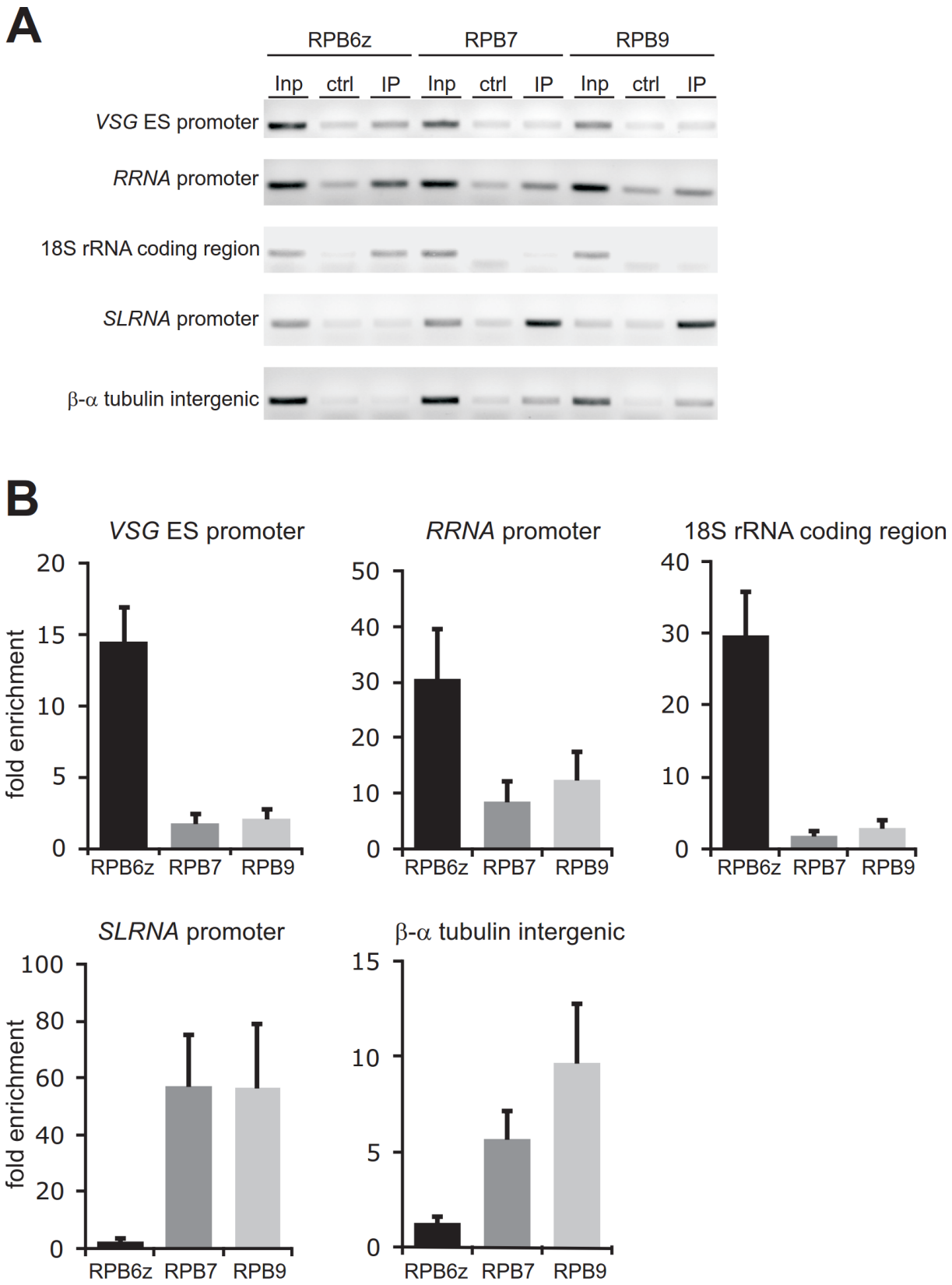


Figure IV-2. ChIP analysis of RPB6z, RPB7 and RPB9. (**A**) Chromatin of clonal BF cells which

(**Figure IV-2 legend cont.)** exclusively express either PTP-RPB6z, RPB7-PTP or RPB9-PTP was precipitated with a rabbit polyclonal anti-ProtA antibody (IP) and analyzed by semi-quantitative PCR. In negative control reactions, chromatin was precipitated with a non-specific rabbit immune serum and for a positive PCR control, a small aliquot of DNA isolated from total chromatin (Inp) was amplified. The occupancy of these subunits were determined for the consensus *VSG* ES and *RRNA* promoters, the 18S rRNA coding region, the SL RNA promoter and the β-/α- tubulin intergenic region. (**B**) Corresponding qPCR analysis shown as the fold enrichment over the negative immunopreciptation. These results are based on at least two independent ChIP experiments each of which was analyzed by three qPCR reactions. Note the different scales on the y axes.

occupancy of the RNA pol II subunit RPB9 at both RNA pol I and II transcription units whereas it did not correlate with the occupancy of RPB6z at these sites. These findings do not support an essential role of RPB7 in RNA pol I transcription.

RPB7 depletion from extracts specifically affects RNA pol II transcription

Previously, we have developed a homologous *in vitro* transcription system in procyclic *T. brucei* that was active in RNA pol I transcription and in RNA pol II-mediated *SLRNA* transcription (Laufer and Günzl, 2001; Laufer et al., 1999). However, to compare our results with those of Penate et al. (Peñate et al., 2009), we needed to establish the system in BF extract. We have been able to develop a small scale extract procedure that was applicable to BFs. We prepared extract from cells that exclusively expressed RPB7-PTP and depleted the extract of ~90% of RPB7 using PTP-binding IgG beads (**Figure IV-3A**). We then analyzed RNA pol I and II activity in assays in which the *SLRNA* promoter template SLins19 was co-transcribed either with the *VSG* ES promoter template VSG-trm or the *RRNA* promoter template Rib-trm. In previous work, we have unambiguously shown that SLins19 transcription is mediated by RNA pol II and VSG-trm and Rib-trm transcription by RNA pol I (Günzl et al., 2003). In comparison to mock-depleted extract, RPB7 depletion virtually abolished *SLRNA* transcription whereas it had no effect on VSG-trm and Rib-trm transcription (**Figure IV-3B**, compare lanes 1 and 2 and lanes 5 and 6). The defect on *SLRNA* transcription was not a non-specific artifact because adding back tandem affinity-purified and active RNA pol II (Das et al., 2006) partially restored transcriptional activity from the *SLRNA* promoter in a dose-dependent manner (lanes 3 and 4, and lanes 7 and 8). We therefore concluded that RPB7 cannot have an indispensable role in promoter-dependent transcription of RNA pol I.

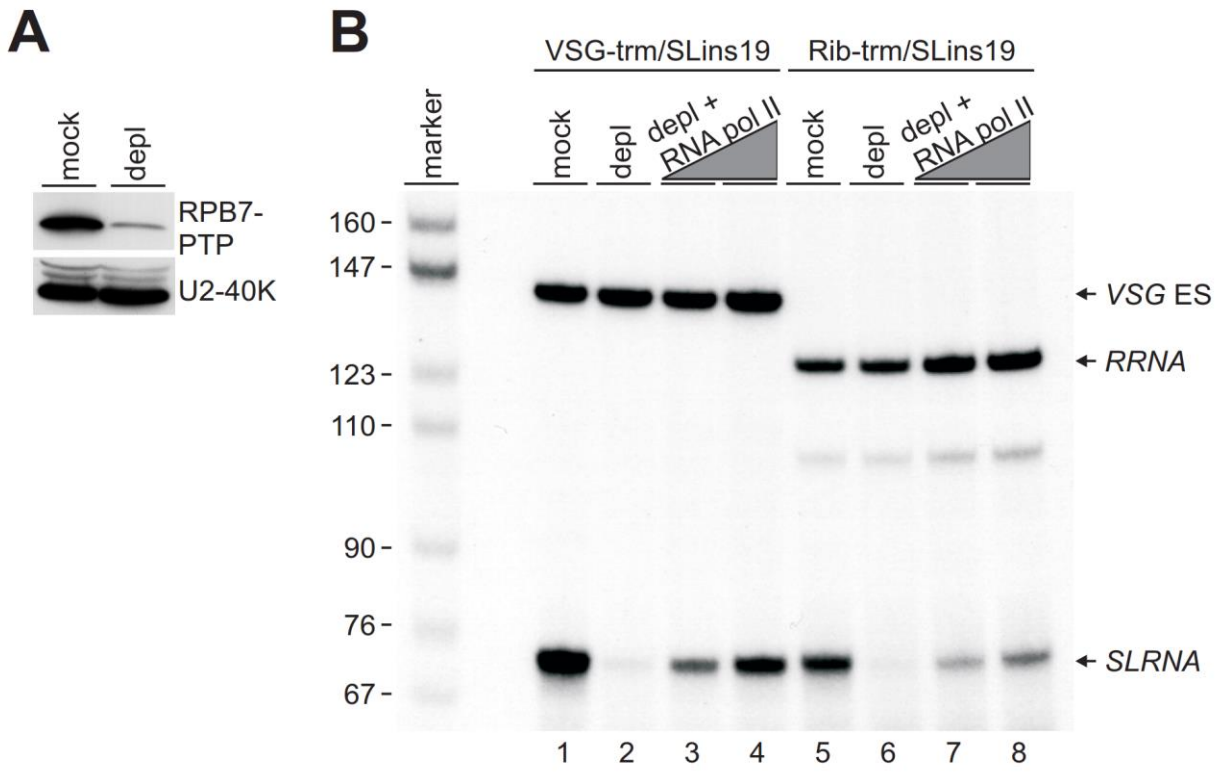


Figure IV-3. *In vitro* transcription analysis of RPB7-depleted extracts. (**A**) Immunoblot of BF extract prepared from cells exclusively expressing RPB7-PTP that was either mock-depleted or RPB7-depleted by means of ProtA-interacting IgG beads. Detection of U2-40k served as a loading control. (**B**) Co-transcription reactions of the SLRNA promoter template SLins19 with the VSG ES promoter template VSG-trm or the *RRNA* promoter template Rib-trm. Reactions were carried out with either mock or RPB7-depleted (depl) extracts. Extracts were reconstituted with tandem affinity-purified RNA pol II. Transcription signals were obtained through primer extension of radio-labeled oligonucleotides that hybridize to unique tag insertions in the SLins19, VSG-trm and Rib-trm RNAs. Primer extension signals were separated on a 6% polyacrylamide-50% urea gel and visualized by autoradiography. On the right, signals for correctly initiated transcription from *VSG* ES, *RRNA* and *SLRNA* promoters are indicated. Marker, MspI-digested pBR322.

Penate *et al*. (Peñate et al., 2009) presented *in vitro* transcription results which indicated that RPB7 is important for *VSG* ES promoter-dependent transcription. However, they used a G-less cassette approach which did not discriminate between accurately initiated transcription and non-specific read-through transcription raising the possibility that their transcription signals were generated in a promoter-independent manner. On the other hand, they did show that *VSG* ES promoter deletion in the template construct resulted in a loss of transcription signal, a finding which cannot be resolved in the light of our results.

RPB7 does not detectably localize to the nucleolus.

RNA pol I has been localized either throughout the nucleolus (Navarro and Gull, 2001) or to the nucleolar periphery (Landeira and Navarro, 2007; Peñate et al., 2009). However, in the latter case the signal was clearly inside the spherical structure of low DNA density that marks the nucleolus in a DAPI stain. Penate *et al.* showed a co-localization mask of RPB7 and RPA1, and argued that in the nucleolar periphery, which is the site of RNA pol I transcription (Landeira and Navarro, 2007), there is significant co-localization. However, their RPB7 localization did not reveal a clear nucleolar or ESB signal and it is possible that the co-localization of RPB7 and RPA1 was derived from adjacent signals that could not be resolved in their analysis. To assess whether RPB7 can be found in the nucleolus we transiently transfected procyclic cells that expressed RPB7-PTP with a construct harboring a NOP10-GFP fusion gene (Ruan et al., 2007). NOP10 is a component of small nucleolar ribonucleoprotein particles and localized in the central part of the nucleolus (Boisvert et al., 2007). In the representative cell shown in **Figure IV-4**, DAPI staining revealed a large and clear nucleolus. RPB7-PTP was found in most parts of the nucleus closely following the perimeter of the nucleolus but not detectably extending into the nucleolus. As expected, in the same cell NOP10-GFP was found in a central nucleolar position.

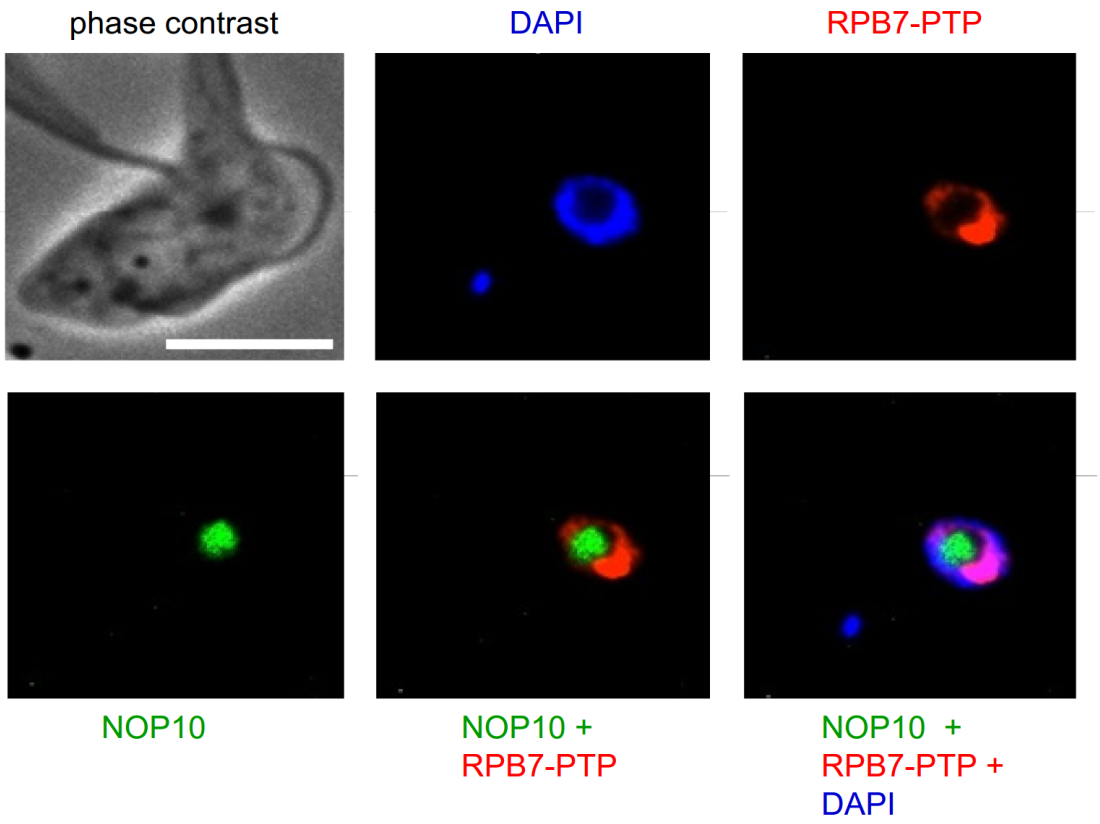


Figure IV-4. Immunofluorescence light microscopy of procyclic trypanosomes showing constitutively expressed RPB7-PTP in red, transiently expressed Nop10-GFP in green, and DNA in blue. Scale bar is 6.5 µm.

Importantly, the co-localization demonstrates a clear gap between Nop10 and RPB7 representing the nucleolar periphery. The gap suggests that RPB7, in contrast to previous localizations of RNA pol I, does not localize to the nucleolus in detectable amounts and is present only at the nucleolar perimeter. This, however, is not surprising because *RRNA* repeats are tightly flanked by RNA pol II transcribed loci, and in procyclic cells, an apparent RNA pol II transcription unit partially overlaps with a procyclin gene unit (Liniger et al., 2001) suggesting that on DNA, RNA pol II-based RPB7 should be present in close proximity to RNA pol I.

IV-4. Conclusion

In this study, we have revisited the role of RPB7 in *T. brucei* RNA pol I transcription. We did not find detectable interactions of RPB7 with RNA pol I in reciprocal co-immunoprecipitations or in tandem affinity purifications with RPB7 as bait. Moreover, RBP7 did not cross-link to RNA pol I transcription units independent of the RNA pol II subunit RPB9, and immunodepletion of RPB7 from BF extracts did not affect *VSG* ES and *RRNA* promoter transcription whereas it did abolish *SLRNA* transcription by RNA pol II. Unlike RNAi-mediated expression silencing of *RPB7*, these approaches avoid potential secondary effects of an RNA pol II shut-down. While our data do not exclude a potential role of RPB7 in *VSG* expression, they show that RPB7 is not a *bona fide* subunit of RNA pol I and has no indispensable role in RNA pol I transcription in *T. brucei*.

It should be noted here that RPA31 could be the missing RNA pol I paralog of RPB7 because it was essential for promoter-dependent RNA pol I transcription and found to be associated with RNA pol I only in the presence of RPB6z (Nguyen et al., 2007; Nguyen et al., 2006), the subunit ortholog that recruits RPA43 in yeast (Schimanski et al., 2006). Although RPA31 was described as a “novel” subunit because its sequence is only conserved among trypanosomatids (Nguyen et

al., 2007), recent characterizations of *T. brucei* TFIIB (Lee et al., 2009; Palenchar et al., 2006), TFIIH (Peyroche et al., 2002) and, in particular, mediator (Lee et al., 2010) revealed that trypanosomatid transcription factor sequences are extremely divergent to those of other eukaryotes and, therefore, may not reveal the identity of a protein.

Acknowledgments

We are grateful to Dr. Christian Tschudi (Yale University) for providing plasmid NOP10-GFP. This work was funded by NIH grant R01 AI059377 to A.G.

IV-5. Response to "Role of RPB7 in RNA pol I transcription in Trypanosoma brucei"

In their letter, Navarro *et al.* (Navarro et al., 2011) criticize our approaches and conclusion that RPB7 is not required for RNA polymerase (pol) I transcription in *Trypanosoma brucei* (Park et al., 2011). Our results are in sharp contrast to their previously published work which claimed that RPB7 is essential for RNA pol I transcription (Peñate et al., 2009). We used several independent criteria to evaluate the functional role of RPB7 in RNA pol I transcription, namely reciprocal co-immunoprecipitation (co-IP), tandem affinity purification (TAP), chromatin immunoprecipitation (ChIP), *in vitro* transcription assays and indirect immunofluorescence (IF) microscopy. In no case did we find evidence that RPB7 has an RNA pol II-independent function in RNA pol I transcription or that this protein is stably associated with RNA pol I. We had hoped that the letter by Navarro *et al.* would help to resolve the discrepancy between our two studies; instead we had to read some rather far-fetched arguments why our results could be artifacts.

Navarro *et al.* describe RPB7 to be essential for RNA pol I transcription and, at the same time, not to be a *bona fide* RNA pol subunit (Navarro et al., 2011). Conversely, though dissociable after transcription initiation (Harel-Sharvit et al., 2010), RPB7 has been described as an essential RNA pol II subunit that is required for promoter-dependent transcription initiation (Edwards et al., 1991) and ChIP assays demonstrated that RPB7 occupancy is congruent to that of the non-dissociable subunit RPB3 (Jasiak et al., 2008). Accordingly, we readily found *T. brucei* RPB7 as part of RNA pol II complexes and cross-linked to RNA pol II transcription units while no such link to RNA pol I was detectable (Park et al., 2011). Navarro *et al.* also state that they have "*previously described a low-affinity interaction of RPB7 with RPA1*" (Navarro et al., 2011). The term "*low affinity*",

however, was not used in their original publication (Peñate et al., 2009) and their interpretation of low affinity remains unclear.

The most obvious discrepancy between the two studies resides in the *in vitro* transcription outcomes. In our assays, RPB7 depletion of extract nearly abolished RNA pol II transcription but left RNA pol I transcription unaffected (Park et al., 2011) whereas a similar approach by Penate *et al.* interfered with RNA pol I transcription (Peñate et al., 2009). It should be noted that our system allows co-transcription of RNA pol I and II promoter templates and, due to primer extension-derived transcription signals, monitors accurate transcription initiation. In comparison, the transcription signals provided by Penate *et al*. (Peñate et al., 2009) depended on a G-less cassette that was introduced downstream of the transcription initiation site. Hence, there is no discrimination between accurately initiated and non-specific read-through transcription. In addition, individual reactions were not internally controlled. This is particularly bothersome because their analysis was not quantitatively assessed, e.g. based on signal quantification of repeated experiments.

According to Navarro *et al.* the lack of an effect on RNA pol I transcription in our assays could stem from residual amounts of RPB7 left in the extract (Navarro et al., 2011). Since our extracts were prepared from cells that exclusively expressed the essential RPB7 as a TAP tag fusion, this would mean that the tandem protein A domains of tagged RPB7 were masked in the RNA pol I complex but accessible in the RNA pol II complex. Having precipitated and purified more than 30 different TAP-tagged proteins (see for example (Brandenburg et al., 2007; Lee et al., 2010; Lee et al., 2009)), we have not come across a single case in which the large tandem protein A domains of the TAP tag were inaccessible to IgG beads or antibodies.

Navarro *et al.* stated that in our co-IP analysis 25% of RPB7 interacted with RPB1 and only 1% of RPB7 interacted with RPA1 (Navarro et al., 2011). This is not true. In fact, we did not find any interaction between RPB7 and the RNA pol I-specific subunit RPB6z in both reciprocal co-IP and RPB7 TAP whereas in both assays RPB7 was stably associated with RNA pol II. We did, however, acknowledge that, due to the ambiguousness of quantifying large amounts of immunoprecipitate, we could not exclude the possibility that a very small amount of RPB7 (<1%) was associated with RPB6z.

In regard to the amount of active RNA pol I in extracts, a study in yeast found that up to 98% of RNA pol I was inactive. However, the inactivity was due to the dissociation of the transcription factor RRN3 from RNA pol I and not to the lack of an RNA pol I subunit, e.g. the RNA pol I paralog of RPB7 termed RPA43 (Milkereit and Tschochner, 1998).

Navarro *et al.* criticize our ChIP approach because we, like many others, compare the enrichment of precipitated DNA over a negative control immunoprecipitation in which we use a comparable, non-specific immune serum. While the suggested normalization to input amount is valid, too, this kind of analysis promotes the usage of extremely low ratios of chromatin input and antibody beads. More importantly, our ChIP analysis demonstrated that identically tagged RPB7, RPB6z and RPB9 could be effectively and specifically cross-linked to chromatin. In all cases, we found congruency of RPB7 and RPB9 (RNA pol II) occupancies that did not correlate with RPB6z (RNA pol I) occupancy. In this regard, we find it inappropriate that Navarro *et al.* mention unpublished data in their support which cannot be evaluated yet.

Concerning IF microscopy, we maintain that the co-localization mask provided by Penate *et al.* (Peñate et al., 2009) cannot discriminate between adjacent signals and true co-localization. Since they concluded that RPB7 is generally required for RNA pol I including transcription of

ribosomal RNA genes, our approach to localize RPB7-PTP in nucleoli of transiently transfected procyclics is valid.

Importantly, a major difference between our study and that of Penate *et al*. was that we used functional protein tags (exclusive expression of essential, tagged proteins) and they used a polyclonal anti-RPB7 antiserum. Since it is very likely that trypanosome RNA pol I requires an RPA43 ortholog (Kuhn et al., 2007) representing an RPB7 paralog, Navarro *et al.* should make sure that their antibody does not cross-react with an RNA pol I-specific protein, e.g. TbRPA31 (Nguyen et al., 2007), the protein we favor in retrospect to be the RNA pol I paralog of RPB7 (see conclusion section in (Park et al., 2011)).

Finally, we are convinced that our data, in contrast to the title-based conclusion of Penate *et al.* (Peñate et al., 2009), demonstrate that RPB7 has no essential role in *T. brucei* RNA pol I transcription; it does not replace the function of an RPA43 ortholog. As we state in our publication, we cannot rule out a post-transcriptional or RNA pol II-based role of RPB7 in the expression of RNA pol I-synthesized transcripts because this was not the aim of our analysis.

Chapter V

Mono-allelic VSG expression by RNA polymerase I in *Trypanosoma brucei*: expression site control from both ends? (Review)

Abstract

Trypanosoma brucei is a vector borne, lethal protistan parasite of humans and livestock in sub-Saharan Africa. Antigenic Variation of its cell surface coat enables the parasite to evade adaptive immune responses and to live freely in the blood of its mammalian hosts. The coat consists of ten million copies of variant surface glycoprotein (VSG) that is expressed from a single *VSG* gene, drawn from a large repertoire and located near the telomere at one of fifteen so-called bloodstream expression sites (BESs). Thus, antigenic variation is achieved by switching to the expression of a different *VSG* gene. A BES is a tandem array of expression site-associated genes and a terminal VSG gene. It is polycistronically transcribed by a multifunctional RNA polymerase I (RNAPI) from a short promoter that is located 45-60 kb upstream of the *VSG* gene. The mechanism(s) restricting *VSG* expression to a single BES are not well understood. There is convincing evidence that epigenetic silencing and transcription attenuation play important roles. Furthermore, recent data indicated that there is regulation at the level of transcription initiation and that, surprisingly, the *VSG* mRNA appears to have a role in restricting *VSG* expression to a single gene. Here, we review BES expression regulation and propose a model in which telomere-directed, epigenetic BES silencing is opposed by BES promoter-directed, activated RNAPI transcription.

V-1. Introduction

The tsetse borne, unicellular parasite *Trypanosoma brucei,* which belongs to the phylogenetic order Kinetoplastida, is the only known organism that has evolved a multifunctional RNA polymerase I (RNAPI) system. This system is used to transcribe ribosomal gene units (*RRNA*) in the nucleolus, as in all eukaryotes, yet also to transcribe gene units that encode the parasite's major cell surface antigens (Günzl et al., 2003; Kooter and Borst, 1984). Trypanosomes have a unique mode of protein coding gene expression that allows them to utilize other RNA polymerases than RNAPII for the production of functional mRNA. In their genome, protein coding genes are arranged in long tandem arrays which are polycistronically transcribed. The precursor RNA is processed by spliced leader (SL) *trans* splicing and polyadenylation, resulting in mature, monocistronic mRNAs (Günzl, 2010; Michaeli, 2011; Preußer et al., 2012). Since in *trans* splicing the same capped leader sequence, derived from the SL RNA, is spliced onto the 5′ end of each mRNA, this process represents a post-transcriptional mode of capping that is decoupled from RNAPII transcription. Consequently, trypanosomes, in contrast to mammals (Grummt and Skinner, 1985), are able to use RNAPI to effectively and specifically express endogenous gene units that encode their major cell surface antigens (Rudenko et al., 1991; Zomerdijk et al., 1991a). This antigen, in mammalian-infective metacyclic and bloodstream form (BF) trypanosomes, is known as the variant surface glycoprotein (VSG), while the major cell surface antigen in insect-stage procyclic form trypanosomes is procyclin.

T. brucei causes Human and Animal African Trypanosomiasis (also known as Sleeping Sickness and Nagana, respectively) throughout sub-Saharan Africa (Fevre et al., 2006). The parasite lives freely in the bloodstream of its mammalian host, evading the immune system by antigenic variation of its cell surface coat. The coat consists of ten million copies of the same VSG,

shielding invariant membrane proteins from immune recognition (Schwede et al., 2011). *T. brucei* possesses roughly 2500 different *VSG* genes and pseudogenes (Cross et al., 2014), and periodic switching to the expression of an alternative *VSG* gene leads to antigenic variation. *VSG* genes are located in subtelomeric regions of 11 megabase, 5 intermediate-sized and ~100 minichromosomes covering, in total, ~30% of the genome (Ersfeld, 2011; Horn, 2014). However, the active *VSG* gene is invariably located next to the telomere within an expression site, with the coding region ending ~200-1800 bp upstream of the telomeric repeats. Metacyclic trypanosomes express a single *VSG* monocistronically from one of five metacyclic expression sites in which the RNAPI promoter is located ~1-4 kb upstream of the coding region (Cross et al., 2014; Ginger et al., 2002; Kolev et al., 2012). Conversely, BFs express the active *VSG* from one of fifteen polycistronic "bloodstream expression sites" (BESs) which comprise a tandem array of typically 8-9 expression-site associated genes (ESAGs) and a terminal *VSG* gene (**Figure V-1**) (Hertz-Fowler et al., 2008). ESAGs appear to be important for the successful infection of the mammalian host since they encode a variant heterodimeric transferrin receptor (ESAG6 and ESAG7), whose varying affinity for transferrins of different host species is thought to expand the parasite's host range (Bitter et al., 1998). These also encode adenylate cyclases (ESAG4) that inhibit the innate immune system upon trypanosome lysis (Salmon et al., 2012).

The BES promoter resides 45-60 kb upstream of the telomere (Zomerdijk et al., 1990). It extends only 67 bp upstream of the transcription initiation site and comprises two short sequence elements (Pham et al., 1996; Vanhamme et al., 1995). Both elements are required for efficient binding of the multi-subunit class I transcription factor A (CITFA) which is essential for RNAPI transcription in the trypanosome (Brandenburg et al., 2007). The active BES is transcribed

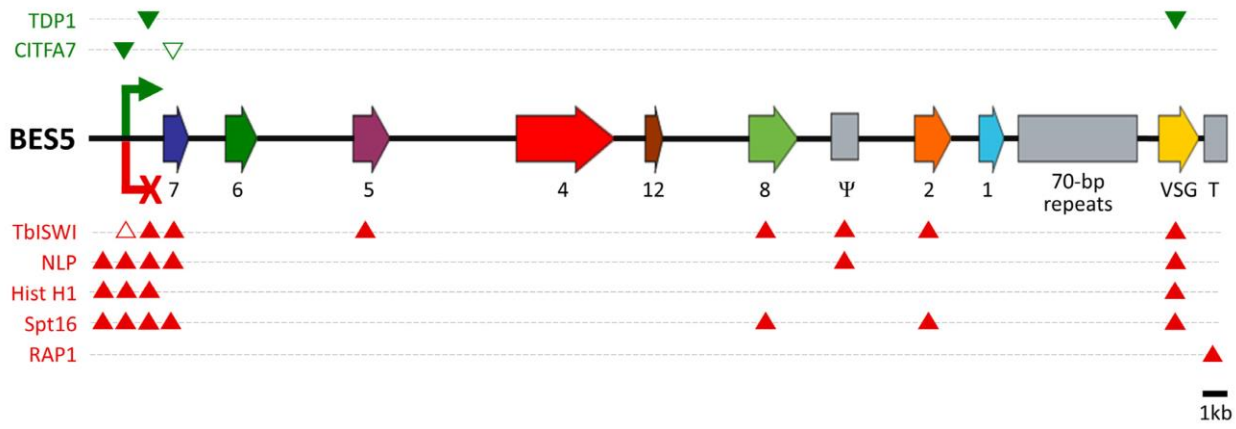


Figure V-1. Schematic outline of BES5 and interacting proteins

Depiction of BES5 (to scale) as a representative BES according to the published sequence (Hertz-Fowler et al., 2008). The diagram includes *ESAG*s (labeled 1, 2, 4-8 and 12), a *VSG* pseudogene (Ψ), 70 bp repeats preceding the terminal *VSG* gene, and the telomeric repeats (T). Note that some BESs have an additional promoter and an *ESAG10* gene ~14 kb upstream of the depicted promoter (not shown). The green arrow and red X represent the promoter when the BES is in the active state and silent state, respectively. Activating factors that are predominantly associated with the active BES are indicated above the diagram in green whereas factors which are implied in BES silencing are listed below the diagram in red. Filled and empty arrowheads indicate positive and negative ChIP results, respectively. Histone H1 and SPT16 associate predominately with silent sites whereas TbISWI and NLP were shown to interact equally with expression sites in both states.

outside the nucleolus (Chaves et al., 1998), apparently in a small compartment termed the expression site body (ESB) (Navarro and Gull, 2001). In BFs the switch to the expression of another *VSG* occurs by two principal ways: either the active BES is silenced while one of the silent BESs is activated, or a DNA recombination event replaces the *VSG* gene in the active BES with a *VSG* gene from the repertoire.

Antigenic variation and mono-allelic *VSG* expression in *T. brucei* have been a research focus for decades. Several factors involved in BES silencing have been identified (see below) and BES silencing has been linked to DNA replication/ORC1 (Benmerzouga et al., 2013; Tiengwe et al., 2012), chromosome maintenance (Kim et al., 2013b), and association of BESs with the nuclear lamina (DuBois et al., 2012). In addition, cohesin plays a critical role in maintaining the activated state of the BES during the cell cycle (Landeira et al., 2009). Recently, excellent and detailed reviews have addressed antigenic variation in trypanosomes and the biology of BES silencing (Alsford et al., 2012; Glover et al., 2013; Horn, 2014; Horn and McCulloch, 2010; Rudenko, 2010). Here we focus on the most recent findings of factors that appear to be directly involved in BES regulation, and propose a model in which BES-specific telomeric silencing is opposed by a mechanism that activates transcription initiation at the promoter of the active BES.

V-2. Telomeric Silencing

The active *VSG* gene, independent of whether it resides in metacyclic or bloodstream expression sites, is invariably located near the telomere, indicating that the telomere has an essential function in regulating *VSG* expression. Accordingly, repression of RNAPI-mediated transcription by the telomere was directly demonstrated by integrating a plasmid with seeds for *de novo* telomere formation either at BESs or, internally, at *RRNA* loci (Glover and Horn, 2006). At

the latter, tight repression extended only 2 kb upstream of telomeric repeats whereas, at inactive BESs, repression reached at least 5 kb in these experiments. The more extended repression of silent BESs was consistent with previous findings in which integration of RNAPI promoter-driven reporter cassettes at different positions of a silent BES were repressed, even when placed 14 kb upstream of the telomere (Horn and Cross, 1997). Several lines of evidence suggest that the pronounced silencing of BESs is dependent on the telomere. Depletion of the telomeric protein RAP1 led to de-repression of silent BESs, co-expression of multiple BES-encoded *VSG* genes, and the formation of additional extranucleolar RNAPI foci (Yang et al., 2009). Furthermore, depletion of the disruptor of telomeric silencing B (DOT1B), which methylates lysine 76 of trypanosome histone H3 (Janzen et al., 2006), similarly led to de-repression of silent BESs (Figueiredo et al., 2008). Direct evidence for repression of a BES from the telomere stems from a recent study in which induced expression of a *VSG* transgene, inserted into one of the *RRNA* loci, surprisingly led to a short-term, reversible attenuation of the active BES, indicating that *VSG* mRNA plays a direct role in the regulation of mono-allelic VSG expression (Batram et al., 2014). Interestingly, a time course experiment showed that this silencing of the active BES spread from the telomere towards the BES promoter in a DOT1B-dependent manner (Batram et al., 2014). Together, these data strongly indicated that BES silencing is directed by the telomere. Furthermore, it is likely that the *VSG* gene on silent BESs is protected from RNAPI transcription by more than one mechanism because *DOT1B* knockout cells could still shut down the active *VSG* gene upon ectopic *VSG* expression but were unable to attenuate expression of the remainder of the active BES (Batram et al., 2014).

V-3. BES transcription attenuation

Inactive BESs are completely silent only in regard to their telomere-proximal regions, including the terminal *VSG* gene. *VSG* mRNA from inactive BESs is 10^4 to 10^5-fold less abundant than that from the active BES (Figueiredo et al., 2008; Yang et al., 2009). Despite this strong difference, several observations have shown that transcription does initiate at silent BESs at a clearly detectable level. The first evidence came from a study in which insertion of a selectable marker gene 1 kb downstream of a "silent" BES promoter led to resistant parasites (Navarro and Cross, 1996). BES sequences are highly similar, especially at the promoter and in the proximal downstream region, differing from each other only by a few single nucleotide polymorphisms. However, the first genes within BESs are *ESAG7* and *ESAG6* which encode the heteromeric transferrin receptor and harbor short hypervariable regions that distinguish them from each other (Zomerdijk et al., 1991b). Analysis of *ESAG6* cDNA sequences, which on BESs are located ~5 kb downstream from the promoter, revealed that 20% of the *ESAG6* mRNA in BFs was derived from various silent BESs whereas 80% stemmed from the active BES, demonstrating that, even in the absence of selective pressure, productive transcription did occur in the promoter-proximal domain of inactive BESs (Ansorge et al., 1999). Subsequently, a vast cDNA clone analysis along whole BESs showed that silent BESs contributed much more to the promoter-proximal cDNA pool than to pools of promoter-distant cDNAs, revealing that transcription that initiated at silent BESs was attenuated along the BES (Vanhamme et al., 2000). Recently, this approach was repeated with single cells, confirming that silent BESs are transcribed in their promoter-proximal region and that transcription was attenuated further downstream within a single trypanosome (Kassem et al., 2014). Finally, the demonstration that BES silencing spreads gradually from telomere to promoter

and BES reactivation occurs gradually in the opposite direction (Batram et al., 2014) strongly supports the notion that transcription is attenuated at silent BESs.

V-4. Regulation of BES transcription Initiation

Although the promoters of inactive BESs are not "silent", there is now convincing evidence that there is substantial regulation at BES promoters. Consistently, promoter-proximal RNA levels were found to be much higher from the active versus silent BESs. Thus, when the neomycin phosphotransferase gene (*NEO*) was inserted 1 kb downstream of the promoter of an inactive BES, it conferred parasite resistance to a low concentration of the drug G418 (1 μg/ml) while the same gene, when inserted at the identical position of an active BES, boosted resistance at least 100-fold (Navarro and Cross, 1996). The finding that in BFs 80% of *ESAG6* mRNA stemmed from the active BES suggested that there is at least a 50-fold stronger *ESAG6* expression from the active BES than from the average silent BES. When Yang *et al.* (2009) introduced a luciferase gene immediately downstream of the active or a silent BES promoter, the active BES produced 1500-4000-fold more light units than the silent BES.

More direct evidence for BES regulation at the level of transcription initiation came from the analysis of CITFA. CITFA consists of seven subunits, CITFA1-7, which are conserved only among kinetoplastid organisms, and the dynein light chain DYNLL1 (also known as LC8). Silencing of *CITFA1*, *CITFA2* and *CITFA7* was lethal to BFs grown in culture and strongly and specifically reduced the abundance of rRNA and *VSG* mRNA (Brandenburg et al., 2007; Nguyen et al., 2012; Park et al., 2014). Accordingly, depletion of CITFA2 from extract virtually abolished RNAPI transcription *in vitro*, as assayed by ~100 bp-long primer extension products, and the purified CITFA complex produced a specific gel shift with the BES promoter (Brandenburg et al.,

2007). Moreover, a ChIP-seq analysis indicated that within a BES, CITFA7 occupancy was restricted to the promoter region (Nguyen et al., 2014). Together, these findings identified CITFA as a basal and general transcription initiation factor for RNAPI transcription in trypanosomes.

Interestingly, marking the active BES and a silent BES ~500 bp downstream of the transcription initiation site (Figueiredo et al., 2008) revealed that CITFA2 and CITFA7 predominantly occupied the promoter of the active BES relative to that of the marked silent BES, a phenotype that was maintained after consecutive *in situ* switches between the two marked sites (Nguyen et al., 2014). In accordance with CITFA's role as an RNAPI transcription initiation factor, higher CITFA occupancy at the active versus the silent BES promoter correlated with a ~70-fold higher abundance of promoter-proximal, unspliced RNA and a ~17-fold higher occupancy of the RNAPI-specific subunit RPB6z at the marker gene (Nguyen et al., 2014). Finally, *CITFA7* silencing led to a strong reduction of RNAPI occupancy and of promoter-proximal RNA levels, which directly demonstrated that CITFA binding to the promoter is required for high transcription rates *in vivo* (Nguyen et al., 2014). These data unequivocally showed that mono-allelic BES expression entails a mechanism that functions at the BES promoter, apparently limiting access of CITFA to silent BES promoters and/or ensuring maximal promoter occupancy of CITFA at the active BES.

It should be noted that this mechanism is not an "all or nothing"-mechanism because, in these experiments, the marked silent BES promoter was consistently occupied by CITFA above the level of negative control experiments. This finding is in accordance with promoter-proximal transcription occurring at silent BESs (see above) and it likely explains why hypersensitive DNase I sites in the promoter region, indicative of a bound transcription factor, were not restricted to the active BES but were also detected at a silent BES (Navarro and Cross, 1998). It appears that

trypanosomes cannot completely shut down transcription initiation from silent BESs. Alternatively, low level transcription of the promoter-proximal part of silent BESs might serve a biological function. For instance, co-expression of different forms of the heteromeric transferrin receptor, e.g. *ESAG6* and *ESAG7*, could ensure initial survival in different mammalian hosts.

V-5. Factors involved in BES transcription regulation

There is strong evidence that inactive BESs are silenced epigenetically. Thus, while silent BESs have a nucleosomal structure, the active BES is largely depleted of nucleosomes (Figueiredo and Cross, 2010; Stanne and Rudenko, 2010). Direct evidence that nucleosomes are important for BES promoter silencing stems from depleting histone H3, which rapidly led to a ~11-fold de-repression of a GFP gene introduced downstream of the promoter of a silent BES (Alsford and Horn, 2012). In addition, CAF-1b, a replication-dependent histone chaperone, and the replication-independent chaperone ASF1A, were shown to be important for the inheritance and maintenance of the silenced state of BESs (Alsford and Horn, 2012). Interestingly, silencing the gene of either chaperone led to apparent nucleosome depletion and a de-repression of the promoter-proximal BES region. However, it did not affect expression of the corresponding *VSG* gene suggesting that nucleosomal structure is particularly important for the regulation of BES promoter activity.

Several chromatin remodeling and modifying proteins have been implicated in BES repression so far. The first epigenetic factor found to play a role in BES regulation was the chromatin remodeler TbISWI (Hughes et al., 2007). Depletion of this factor increased the mRNA abundance of a reporter gene inserted promoter-proximally into a silent BES up to 60-fold, whereas only a fivefold increase of the corresponding silent *VSG* mRNA was observed. TbISWI was found to

occupy the entire length of both silent and active BESs, but was not enriched at BES promoters (Stanne et al., 2011). Although the specific function of TbISWI remains to be determined, these results suggest that TbISWI controls RNAPI transcription elongation rather than initiation.

Similar de-repression of a promoter-proximal reporter gene was observed when the histone deacetylase DAC3 (Wang et al., 2010), the linker histone H1 (Pena et al., 2014; Povelones et al., 2012), or the nucleoplasmin-like protein NLP (Narayanan et al., 2011) was depleted. The function of DAC3 appears to be promoter-specific since expression of the *VSG* gene in the marked BES was unaffected at both the mRNA and the protein level (Wang et al., 2010). However, direct association of DAC3 with BESs has not been demonstrated yet and it remains a possibility that DAC3's control of BES silencing is indirect.

The role of histone H1 in BES promoter repression has been more deeply investigated. Histone H1 is important for chromatin architecture and generally functions in chromatin condensation and transcription repression (Happel and Doenecke, 2009). Accordingly, co-silencing of the *T. brucei H1* multigene family opened up chromatin globally with the strongest effect on silent BES promoters (Pena et al., 2014). Metabolic labeling of nascent RNA then showed that histone H1 depletion resulted in an approximately six-fold higher promoter-proximal transcription rate at a silent BES, indicating that relaxation of the nucleosome structure in the promoter region led to an increase of the transcription initiation rate at the silent BES (Pena et al., 2014).

NLP is a ubiquitous nuclear protein and, accordingly, was found to be associated with all genomic loci analyzed, including the active and silent BESs (Narayanan et al., 2011). Despite this apparent general association with genomic DNA, NLP seems to be particularly important for BES promoter regulation. Depletion of NLP de-repressed a silent BES 45-65-fold, as measured by fluorescence derived from a promoter-proximal GFP gene. Moreover, *NLP* silencing also reduced

promoter-proximal gene expression from the active BES about threefold (Narayanan et al., 2011). While it was speculated that NLP may have a dual function in BES silencing and in promoting processive transcription at the active BES (Narayanan et al., 2011), it is equally possible that loss of NLP enabled competition between silent and the active BES for the RNAPI transcription machinery. However, the specific function of NLP in BES regulation remains to be determined.

SPT16 is a subunit of the trypanosome FACT ("facilitates chromatin transcription") complex (Patrick et al., 2008) and appears to have a direct role in BES promoter silencing because it was found highly enriched at a silent BES promoter (Denninger et al., 2010). Accordingly, *SPT16* silencing increased promoter-proximal GFP expression from a silent BES up to 25-fold, yet de-repression did not extend to the *VSG* genes of silent BESs. However, the de-repression effect was strongly correlated with an arrest in the G2/early M cell cycle phase, raising the possibility that SPT16 does not generally facilitate BES repression in the bloodstream trypanosome. Moreover, SPT16 depletion strongly reduced *VSG* expression from the active BES, suggesting a separate BES-related function of SPT16 in facilitating processive RNAPI transcription. Overall, the specific function of FACT in the multifunctional RNAPI system remains unclear. While SPT16 has been co-purified with RNAPII of the related organism *Leishmania major* (Martinez-Calvillo et al., 2007), its association with *T. brucei* RNAPI remains to be shown.

The epigenetic factors discussed so far, including RAP1 and DOT1B (see section 2, *Telomeric silencing*, above), function in BES silencing. The only such factor found to be important for efficient transcription of the active BES is the high mobility group protein TDP1, which belongs to a family of architectural chromatin proteins (Narayanan and Rudenko, 2013). Interestingly, TDP1 exhibited an inverse occupancy pattern to the core histone H3 at RNAPI-transcribed loci and was up to fivefold more abundant at the active BES promoter relative to a silent BES promoter.

Accordingly, TDP1 depletion decreased the abundance of pre-rRNA and *VSG* mRNA from the active BES. In addition, TDP1, a nuclear protein, exhibited predominant localization to the nucleolus and the ESB, and its DNA association was found throughout the active BES and *RRNA* gene units. Thus, it appears that TDP1 facilitates high rates of processive RNAPI transcription required for trypanosome survival (Narayanan and Rudenko, 2013).

V-6. A model of BES regulation

It is difficult to integrate the data from BES de-repression studies because, for most factors, specific functions in BES silencing have not been determined yet. Nevertheless, recent data strongly indicated that BES regulation occurs at both ends of expression sites. RAP1 and DOT1B depletion studies have clearly shown that BESs are silenced by a telomere-directed mechanism. Moreover, the demonstration that BES silencing spreads from the telomere towards the promoter (Batram et al., 2014) strongly supports a telomere-directed BES silencing mechanism. However, it is unlikely that this is the only mechanism regulating mono-allelic BES expression. If this was the case one would expect full activation of promoter-proximal transcription once telomeric silencing retreats beyond the promoter region, which should result in a leveling of the transcription rate between active and silent BESs (given the extremely high expression level of the active BES, it is unlikely that a trypanosome can support full activation of all fifteen BESs). However, in all cases reported, de-repression of silent BESs is, at best, moderate with the promoter-proximal expression level remaining manifold below that of the active BES. Furthermore, *RAP1* silencing did not strongly affect promoter-proximal transcription of de-repressed BESs and had only a minor influence on the high expression level of the active BES (Yang et al., 2009). Similarly, *DOT1B*

silencing did not affect expression of the active BES at all (Figueiredo et al., 2008). These results strongly argue for the presence of a separate mechanism involved in BES regulation.

Transcription attenuation has been proposed to be that mechanism and, as discussed, there is clear evidence that it does occur on silent BESs (Kassem et al., 2014; Vanhamme et al., 2000). Moreover, some data suggested that transcription attenuation is caused by inefficient transcript processing (Vanhamme et al., 2000) rather than by repressive chromatin. However, recent data do not support this scenario. The finding that *RAP1* silencing caused gradual BES de-repression with the greatest effect on telomere-proximal genes, strongly indicated that transcription elongation on silent BESs is "*antagonized*" by telomere-directed spreading of repressive chromatin (Yang et al., 2009). Furthermore, upon removal of the apparent transcription elongation barrier, e.g. pronounced telomeric silencing, by RAP1 (Yang et al., 2009) or DOT1B (Figueiredo et al., 2008) depletion, promoter-proximal expression remained magnitudes below that of the active site, making it unlikely that transcription attenuation accounts for the strong difference in promoter-proximal transcription observed between active and silent BESs. Hence, transcription attenuation appears to be a consequence of epigenetic silencing rather than a regulatory mechanism, and seems to be in place to prevent the low level of transcription that does initiate at silent BESs from reaching the distally located *VSG* gene.

Based on the *RAP1* silencing results on BES de-repression, Yang *et al.* (2009) suggested that there has to be a mechanism functioning on the BES promoter that could explain the striking difference in promoter-proximal expression levels between the active and de-repressed/silent BESs. The strongest support for this idea stems from the demonstration that CITFA, which is absolutely required for RNAPI transcription, is predominantly associated with the active BES promoter, versus a silent site, strongly indicating that there is a mechanism in place that allows

CITFA to preferentially interact with the active BES. In addition, the fact that depletion of several epigenetic factors increased promoter-proximal transcription with no or very little effect on the downstream *VSG* gene further supports the notion of a promoter-dependent regulatory mechanism.

Taking these data into account, we propose a model in which BESs are regulated by two opposing forces, namely telomere-directed epigenetic silencing acting on silent BESs and activated transcription initiation at the active BES (**Figure V-2**). In this model, the active BES promoter has unrestricted access to CITFA and RNAPI, allowing it to achieve the high transcription rate necessary for productive *VSG* expression. In addition, productive RNAPI transcription at the active BES is ensured by the presence of TDP1. At the same time, telomeric silencing at the active BES is impaired or pushed back so far that RNAPI transcription can extend productively past the *VSG* gene. In contrast, in this model, silent BES promoters are unable to recruit CITFA and RNAPI in sufficient amounts to allow for high transcription rates. In addition, a telomere-directed repressive epigenetic gradient spreading from the telomere into the BES causes transcription attenuation to prevent the low level of transcription, initiating at inactive BESs, from reaching the *VSG* gene.

How are BES promoters differentially regulated? An obvious mechanism that could prevent CITFA from interacting with silent BES promoters and from recruiting RNAPI is a repressive chromatin structure at the promoter. The epigenetic factors whose depletion led to promoter-proximal BES de-repression, e.g. DAC3, histone H1 and NLP, may be important to build up a repressive chromatin structure at the promoter. If this is correct, depletion of these factors should lead to higher CITFA and RNAPI occupancies at de-repressed promoters. An alternative idea for the low promoter activity at “silent BESs” has been put forward in studies of CITFA. CITFA was found to be concentrated in both the nucleolus and the ESB (Nguyen et al., 2014), and it retained

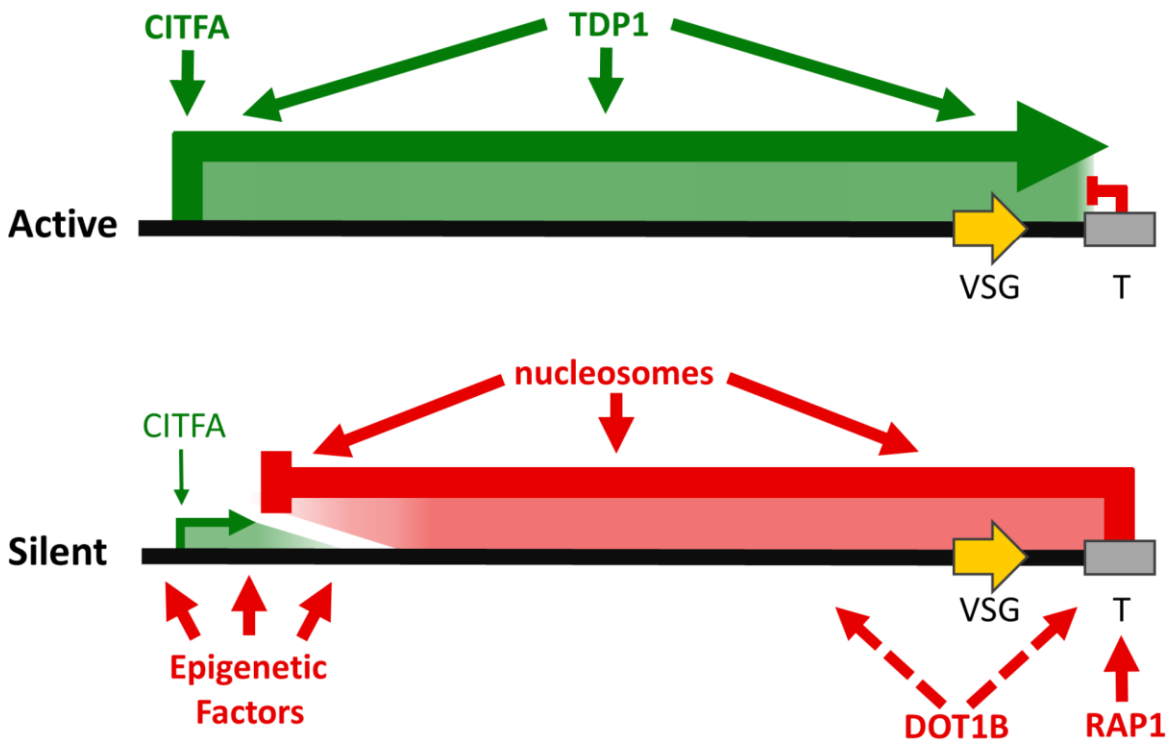


Figure V-2. Model of BES regulation in *T. brucei*

In the model of BES regulation, two opposing forces antagonize each other. The active BES is characterized by high transcription initiation rates and the lack of telomere-dependent epigenetic silencing, allowing unrestricted transcription elongation past the terminal VSG gene (green arrow). High processive transcription rates are facilitated by CITFA and TDP1. In silent BESs, low level RNAPI transcription initiation is opposed by BES-specific telomeric silencing that spreads towards the BES promoter causing transcription attenuation. This silencing depends on a nucleosomal structure, DOT1B and RAP1. RAP1 was shown to bind to telomeric repeats but the association of DOT1B with BESs remains to be determined (dotted line). Epigenetic factors may work together to build up a repressive chromatin structure at the promoter of silent BESs, preventing efficient binding of CITFA to its cognate DNA sequence elements. Alternatively, CITFA sequestration may limit the availability of the initiation factor for inactive BESs (not shown).

this localization even when its promoter-binding capability was impaired by depletion of the essential subunit CITFA1 (Park et al., 2014). It was therefore suggested that sequestration of CITFA into the nucleolus and the ESB could restrict maximal RNAPI transcription to these compartments. This idea is in line with a previous study in which BFs were forced to co-express two BESs simultaneously by antibiotic selection. The two marked BESs were consistently detected in close spatial proximity (Chaves et al., 1999), as if they were competing for an essential expression factor. Furthermore, it may explain why de-repressed BES promoters remain much less active than the promoter from the active BES.

Finally, this model of two opposing forces is supported by the monitoring of the shut-down/reactivation of the active BES upon ectopic expression of *VSG* mRNA (Batram et al., 2014). In these experiments the active BES was gradually inactivated from the telomere towards the promoter, most likely by an active, telomere-directed process of repressive chromatin spreading, whereas the reactivation of the same BES occurred in the reverse direction, possibly by removal of nucleosomes by the transcription machinery. An important question emanating from this study is how does the ectopically expressed *VSG* mRNA cross-talk to the active BES? Although this question is beyond the scope of this article, it is tempting to speculate that the *VSG* mRNA sequestered an important factor for RNAPI transcription, allowing repressive chromatin to spread onto, and silence, the active BES.

V-7. Conclusion

Mono-allelic *VSG* expression in *T. brucei* differs from other allelic exclusion systems, such as *var* gene expression in *Plasmodium falciparum* (Guizetti and Scherf, 2013; Kirkman and Deitsch, 2012) or olfactory receptor expression in mammals (Magklara and Lomvardas, 2013) by the fact

that the active *VSG* gene must be transcribed at an extremely high rate to enable rapidly proliferating trypanosomes to completely cover themselves with VSG. The careful measurement of RNA abundances and half lives indicated that the active *VSG* gene is transcribed at a 50-fold higher rate than a β-tubulin gene (Ehlers et al., 1987). At the same time, *T. brucei* must ensure that *VSG* genes on other BESs are not expressed. The parasite achieves this balancing act apparently by restricting full RNAPI transcription initiation to the active BES and by shielding *VSG* genes on silent BESs by a telomere-dependent silencing mechanism that causes attenuation of RNAPI transcription.

Acknowledgments

This work was supported by grants R01 AI059377 to A.G. and F30 AI110060 to J.K.K from the National Institute of Allergy and Infectious Diseases (NIAID) of the U.S. National Institutes of Health (NIH).

Chapter VI

Discussion and future directions

VI-1. Gene silencing by targeting heterologous sequences is a new and flexible tool for studying *T. brucei*

Previous work on CITFA revealed that it was a promoter-binding factor required for all transcription by RNA pol I in *T. brucei* (Brandenburg et al., 2007). CITFA was shown to consist of 8 subunits, CITFAs 1-7 (Brandenburg et al., 2007; Nguyen et al., 2012), conserved only among kinetoplastids, and a dynein light chain LC8 that is conserved among eukaryotes (Wickstead and Gull, 2007). CITFA was previously shown to co-localize to the sites where transcription by RNA pol I occurs, and was also shown to be present at *RRNA* and BES promoter (Brandenburg et al., 2007; Nguyen et al., 2012). For an overview of transcriptional regulators known to bind to active and silent BESs, including CITFA, please see **Figure V-2** on page 139. While CITFA2 and CITFA7 were previously shown to be essential for transcription by RNA pol I, the specific role of CITFA subunits in transcription was unclear. Furthermore, while it had been demonstrated that CITFA was localized to the nucleolus and the expression site body, or ESB, where *VSG* transcription takes place, its means of restriction to these compartments was unknown. While pursuing the above, we attempted to silence *CITFA1* by producing double stranded RNA (dsRNA) targeting *CITFA1*'s coding sequence and 3′ untranslated region (UTR), which failed, even though this approach had been successfully used in our lab with other genes.

As detailed in **Chapter II** and **Chapter III**, we were able to solve this problem, and investigate the role of CITFA1, by developing a new system for gene silencing that relied on targeting heterologous sequences. We demonstrated that it was possible to specifically silence *CITFA7* and *CITFA1* using two different fused heterologous sequences (Park et al., 2014), one being a tandem affinity tag while the other was a 3′ UTR from another member of the genus, *Trypanosoma cruzi*. The phenotype produced upon *CITFA7* silencing was in accordance with a previously published silencing of *CITFA7* that targeted its coding sequence (Nguyen et al., 2012), showing that our approach was valid. The lack of a phenotype upon replacing *CITFA7*'s endogenous 3′ UTR with one from *T. cruzi* suggests that the genomic elements required for RNA processing are shared between these two species. The main purpose of this work, however, was to develop a tool for specific gene silencing that was less reliant on targeting endogenous sequences, and there are numerous advantages to using this approach. Firstly, some genes, as was shown in **Chapter III**, were unaffected when dsRNA targeting their coding sequences or UTRs were expressed. Only though fusion of a targetable heterologous sequence could silencing be achieved. Secondly, gene silencing by production of dsRNA has traditionally required expression of a 500 bp long targeting sequence (Shi et al., 2000). This makes targeting smaller proteins or genes with shorter UTRs difficult, as they may not have the required sequence length for decisive silencing. Thirdly, due to the length of the dsRNA, the cells likely produce many different siRNAs, increasing the chance for a deleterious off target effect. However, as tested, induced expression of the two heterologous dsRNAs did not affect cell proliferation, eliminating such a possibility in this approach. Finally, our heterologous sequence targeting approach allows for genes with a high degree of similarity to be specifically and individually silenced. In the extreme case, it would even be possible to silence only one allele of a given gene, in order to study the effects of gene dosage.

Likewise, this method will allow for homologues of *T. brucei* proteins found in related species to be investigated for functional equivalency in *T. brucei*. Tagging an endogenous gene and targeting it for knockdown via a fused PTP tag or *T. cruzi* 3′ UTR, would allow for rescue studies to be performed with homologues from other kinetoplastids, or, as was demonstrated in **Chapter II**, mutants which vary only slightly from their wild-type counterparts. The importance of the phosphorylation sites in CITFA7 (Nguyen et al., 2012), for example, could be studied by attempting to rescue for the knockdown of the wild type protein with a mutant in which phosphorylated residues are mutated to comparable amino acids which cannot accept a phosphorylation. In short, the ability to specifically knock down target genes, regardless of their sequence length or homology, is a powerful tool that could enhance almost any genetic studies of protein function.

VI-2. CITFA1 is required for transcription but not CITFA complex assembly

Once we had established this new method for gene silencing, we were able to return to our original interest in determining the specific function of individual CITFA subunits. We were able to demonstrate that silencing *CITFA1* was lethal to bloodstream form (BF) trypanosomes, which emphasizes the essential nature of this complex, as *CITFA2* and *CITFA7* silencing were similarly lethal (Brandenburg et al., 2007; Nguyen et al., 2012). The loss of RNA pol I-synthesized transcripts upon *CITFA1* silencing confirmed CITFA's role as a RNA pol I transcription factor. *CITFA1* silencing also resulted in a loss of CITFA from the *RRNA* and BES promoters, indicating that CITFA1 has a role in DNA binding. Interestingly, however, a loss of CITFA1 protein was not accompanied by a loss of other CITFA subunits or a change in CITFA's localization. This reveals that the CITFA complex does not rely on CITFA1 for assembly, and suggest that CITFA is able

to localize to the ESB in the absence of DNA binding. Given that little is known about the nature of the ESB, beyond that it includes RNA pol I, CITFA, the active BES, and resulting *VSG* mRNA (Navarro and Gull, 2001), it may be possible to use DNA-disassociated CITFA, induced by *CITFA1* silencing, to investigate the proteins that are likely binding to and anchoring CITFA in the ESB.

Though CITFA1 was shown to be important for CITFA to bind RNA pol I promoters, and is likely a direct DNA contactor, this has yet to be conclusively shown. Tagging and purifying CITFA1 would allow for UV crosslinking studies to be performed side-by-side with CITFA purified by either a tagged CITFA2, CITFA4, or CITFA7. A shift in the lower of the two ~50 kDa bands observable when purified CITFA is subjected to UV crosslinking **(Figure II-5B)** would agree with other data, and would indicate that CITFA1, in addition to the already proven CITFA2 and CITFA4, is a direct DNA contactor.

VI-3. The CITFA2-LC8 interaction is essential for RNA pol I-mediated transcription in *T. brucei*

In addition to our general interest in determining the specific functions of individual CITFA subunits, we decided to investigate LC8 because of its role in transcription in other organisms, reviewed by Rapali et al., (2011b). *LC8* silencing caused a severe defect in mitotic progression, resulting in cells which were multinucleated yet had failed to divide. This agrees with work performed in HeLa cells, which also showed a mitotic block upon *LC8* silencing (Asthana et al., 2012). In addition to this phenotype, a drop in RNA pol I-synthesized transcripts was noted, with the drop in *VSG* mRNA being more pronounced than the effect upon *RRNA*. Given the severity of the apparent cytokinesis defects, pursuit of the role of LC8 in RNA pol I-mediated transcription

required identifying, and disrupting, LC8's interaction with the CITFA complex. We determined that LC8 interacted with a conserved 3-amino acid strength, TQV, at the N-terminus of CITFA2, and that this interaction promoted dimerization of CITFA2. Mutation or removal of this sequence eliminated the interaction between CITFA2 and LC8 both *in vitro* and *in vivo*. We were also able to show that this interaction is essential for trypanosome viability, and that disrupting it dramatically reduced *RRNA* and *VSG* mRNA levels. Furthermore, we showed that without this interaction, CITFA2 is unable to assemble into the CITFA complex or bind RNA pol I promoters. Lastly, we identified CITFAs 2, 4, and, likely, 1 as direct DNA contactors, and showed that without CITFA2, the remaining CITFA complex, though stable, was unable to bind promoter DNA or initiate transcription by RNA pol I. For a model summarizing these interactions between CITFA2 and LC8, and the assembly of the CITFA complex, please see **Figure II-7**, page 59. These data represent a massive step forward in understanding the roles of individual CITFA subunits, and is the first demonstration of LC8 being required for basal transcription. It also shows that LC8's function as a dimerizing 'hub' protein is of ancient evolutionary origin.

VI-4. Does LC8 have additional function in CITFA or in RNA pol I-mediated transcription?

Our data are consistent with the view that LC8, rather than functioning as a linker between two different proteins, instead promotes homodimerization, thereby imparting new function (Barbar, 2008; Barbar and Nyarko, 2014, 2015; King, 2008). It would be interesting to determine, however, if this was LC8's only role in the CITFA complex, as our work is far from conclusive in this regard. This could be tested by substituting the N-terminus of CITFA2 with sequences known to induce dimerization, such as coiled coil domains. If an N-terminal deletion CITFA2 mutant with the

addition of such domains was able to rescue for the knockdown of wild-type CITFA2, then it would be likely that LC8's only function in binding CITFA2 is to promote its dimerization. It would be interesting to test in these cells if LC8 was still present in CITFA that contained the mutant CITFA2. If LC8 were still present, it would indicate that LC8 may be interacting with other members of the complex. Though not as conserved as the LC8 site in CITFA2 (VGTQV), other CITFA subunits have sites which would be of almost comparable LC8 binding strength, according to binding strength assessments offered by Rapali et al. (2011a). CITFA5, both a and b homologues, contain a sequence, RVTQV, which is not conserved with other kinetoplastids. Given that CITFA5b amino acid sequence has 72% identity to CITFA5a, and that both have been identified by mass spectrometry in tandem affinity-purified CITFA eluates, it is tempting to speculate that these subunits are dimerized by LC8. Our lab has also performed experiments investigating a possible LC8 binding site in CITFA7, AGVQV, which is fairly well conserved in trypanosomatids, though not in *Leishmania* spp (data not shown). Unpublished data showed that CITFA7 and LC8 interacted in a yeast-two-hybrid assay and mutating this putative binding site in CITFA7 increased the expression from silent BESs in BF trypanosomes (Tu N Nguyen, Bao N Nguyen & A Günzl, unpublished), raising the possibility that a CITFA7-LC8 heterotetramer is responsible for localizing the CITFA complex to nucleolus and ESB. While it may seem unlikely for LC8 to be involved in more than one element of the same complex, it has been demonstrated with the estrogen receptor and KIBRA that LC8 is involved at multiple points in the activation of the estrogen response pathway (Rayala et al., 2005; Rayala et al., 2006).

It is possible that LC8 has additional effects upon the CITFA complex, or that it affords the cell an additional point of regulation to control the expression of either *RRNA* or *VSG* mRNA. The conservation of the TQV sequence in CITFA2 that binds LC8 with other, non-*VSG* expressing

kinetoplastids, supports the idea that LC8 is not playing a role in only *VSG* expression, though this may not be the case. The amino acids surrounding the TQV-motif are not entirely conserved, and are within close enough proximity to affect the CITFA2-LC8 interaction (Rapali et al., 2011a). It would be interesting, therefore, to express a *CITFA2* sequence from a related kinetoplastid in *T. brucei*, either as an entire unit, or as a chimera, with just the N-terminal portion fused to an N-terminally truncated *T. brucei CITFA2*, in order to determine if the CITFA2s present in related organisms were able to interact with LC8 or rescue the knockdown of endogenous *CITFA2*.

It has been shown in other organisms that phosphorylation of either LC8 or its binding partner can alter their interaction (Benison et al., 2009; Gallego et al., 2013; Lei and Davis, 2003; Song et al., 2008; Song et al., 2007). This regulatory phosphorylation occurs at S88 in human LC8, which is conserved as S89 in the genus *Trypanosoma* (**Figure II-S1B**), allowing for such a regulation to exist. Furthermore, phosphorylation of CITFA2 at amino acid T6, one of the conserved residues in the LC8 binding motif, could also be used as a means to block LC8 binding, as is the case in Nek9, a kinase involved in mitotic progression (Gallego et al., 2013). The phosphorylation state of CITFA2 and LC8 in *T. brucei* are, however, unknown. Before an investigation of the importance of phosphorylation of these proteins is performed, it will be necessary to determine whether these sites are phosphorylated in trypanosomes.

It should be noted that the above hypotheses regarding the function of LC8 in *T. brucei* are, in part, made possible by the work presented in **Chapter II.** Previous to this work, nothing was known regarding LC8 binding partners or binding site preferences in *T. brucei*. While the high degree of identity between human and trypanosome LC8 made it likely that characteristics were shared, the early divergence of kinetoplastids from other eukaryotic lineages made this uncertain. Our work, though only detailing one LC8 interaction, demonstrates that LC8 appears to bind

motifs similar to those detailed in other organisms, and that its role in promoting or stabilizing dimerization is the same. This means that the vast array of studies detailing LC8's binding preferences can now be tentatively used to predict potential LC8 binding partners in kinetoplastid organisms. Once additional studies on LC8 in these organisms have been performed, if they are also consistent with the non-kinetoplastid literature, it will allow for rapid progress in our understanding of LC8 in these organisms by allowing use of predictive bioinformatics.

VI-5. Does LC8 have functions in *T. brucei* outside of dynein and CITFA?

In addition to the above means of investigating LC8 function, it would likely be informative to tandem affinity-purify LC8 and characterize LC8-containing protein complexes by sucrose gradient sedimentation and mass spectrometry. However, it is currently unclear whether LC8 can be functionally tagged. This approach may even allow for the binding preferences of LC8 in *T. brucei* to be further established by sequence analysis *in silico*, if enough co-purifying proteins have recognizable LC8 binding motifs.

VI-6. CITFA and the ESB

The work contained in **Chapter II** demonstrated that, upon *CITFA2* silencing, CITFA3 is stable, but is lost from RNA pol I promoters. *LC8* silencing produced a similar effect, with CITFA6 also being shown to be stable in this case. CITFA2, therefore, represents the second CITFA subunit, with the first being CITFA1, to reveal that binding DNA is not required for CITFA assembly. This concept is intriguing, as it suggests a CITFA complex that assembles and localizes to sites of RNA pol I-mediated transcription independent of DNA and independent of RNA pol I, as an RNA pol I subunit has never co-precipitated or co-purified with CITFA. How then is CITFA

localizing to the nucleolus and ESB so effectively? One way to approach this question would be to tandem affinity purify the CITFA complex after *CITFA1* silencing, and subject the purified material to mass spectrometry. This may enhance the likelihood of identifying the proteins responsible for CITFA's localization, as it would ensure that CITFA is not interacting with DNA. While there is no evidence that CITFA interacts with RNA, it is possible that a CITFA-RNA interaction mediates its localization. While it was previously demonstrated that the accumulation of RNA pol I in the ESB is unaffected by DNase I treatment, implying that this compartment is not reliant on DNA binding or transcription (Navarro and Gull, 2001), which our data agrees with (Park et al., 2014), a similar experiment with RNase has not yet been performed. If CITFA's localization would be RNA-dependent, then preparation of total RNA from purified CITFA in conjunction with RNA-seq may reveal the identity of the interacting RNA, a method known as RNA immunoprecipitation sequencing, or RIP-seq. If RNase or nuclease treatment does not affect CITFA localization, it would reveal that the ESB had a protein-based architecture.

VI-7. The increasing therapeutic potential of LC8

Our results add *T. brucei* to the list of pathogens that rely on LC8 for viability. *LC8* silencing caused defects in both transcription and cytokinesis, and resulted in an extremely rapid death phenotype of cultured BF *T. brucei* that is rarely observed in this system. The importance of LC8 to infectious agents has also been shown for HIV (Jayappa et al., 2015), Ebola (Luthra et al., 2015), rabies (Tan et al., 2007), and *Toxoplasma gondii* (Qureshi et al., 2013), previously. Additionally, LC8 has been shown to have a role in promotion of cancerous phenotypes (Asthana et al., 2012; Vadlamudi et al., 2004). Given the nearly universal nature of LC8 (Wickstead and Gull, 2007), and the lack of available small molecule inhibitors, LC8 therapeutics remain strictly conceptual.

However, it may be feasible to target LC8's binding partners, kinases, or phosphatases, if such were able to be detailed. While not promising in the near term, the reliance of various pathogens on this molecule make it an exciting potential target.

VI-8. The role of RPB7 in RNA pol I transcription

We investigated whether RPB7, a known subunit of RNA pol II, was being utilized by RNA pol I in *T. brucei*, a claim published by the research group of Dr. Miguel Navarro (Spanish National Research Council, Granada, Spain) (Peñate et al., 2009). Immunoprecipitation and tandem affinity purification of RPB7 identified numerous RNA pol II subunits, yet failed to produce any subunits of RNA pol I, as shown in **Chapter IV**. Consistent with these results, ChIP experiments revealed a profile for RPB7 consistent with it functioning solely within RNA pol II. *In vitro* transcription assays revealed that depletion of RPB7 had no effect on either *VSG* or *RRNA* transcription, while it did affect an RNA pol II-transcribed gene. Lastly, immunofluorescence showed that RPB7 failed to co-localize with a nucleolar marker, the site in which RNA pol I-mediated transcription occurs. Taken together, these data clearly show that RPB7 is a *bona fide* RNA pol II subunit, while they fail to provide any evidence for an involvement of RPB7 in transcription by RNA pol I

Both we and the Navarro group published letters addressing the incongruent nature of the data from our two research teams (Günzl et al., 2011; Navarro et al., 2011). While specific arguments will not be fully covered here, a few key points are worth making. First, we feel that our data are clear and decisive on the topic of RPB7 involvement in RNA pol I. Secondly, it is possible, due to their using polyclonal antisera, that the Navarro group has correctly detailed the importance of a protein for RNA pol I-mediated transcription that is simply not RPB7. Since it is

very likely that trypanosome RNA pol I requires an RPA43 orthologue (Kuhn et al., 2007), the identity of which has not yet been determined in *T. brucei* and which would be an RPB7 paralog, it is possible that their antisera and experiments have identified the function of this protein, instead. We avoided this potential problem by tagging RPB7, which allowed us to use specific antibodies for RPB7 detection and purification. In addition, the data by Penate et al. (2009) overwhelmingly depended on RPB7 silencing. Since this likely affected expression of nearly all protein coding genes, it is possible that they were observing secondary effects on RNA pol I transcription. Therefore, we avoided this approach and used biochemical assays that directly addressed RPB7's role in RNA pol I-mediated transcription.

We do not claim, however, that RPB7 has no effect upon transcription by RNA pol I or the transcripts it produces. RPB7 is an interesting member of RNA pol II, as it forms a heterodimer with RPB4 and appears to participate in cellular processes which are separate from the immediate transcription process by RNA pol II, as reviewed in Sharma (2013). These include DNA repair, mRNA export and decay, and translation. Thus, it is possible that some of the phenotypes that Dr. Navarro's group attributes to RPB7's role in RNA pol I are, in fact, the result of RPB7's role outside of the context of RNA polymerases entirely. It will be interesting to see if future investigations of RPB7 (an RPB4 homolog has not been identified in trypanosomes yet) implicate these proteins in the same diverse functions in *T. brucei* as they have in other organisms.

References

Alibu, V.P., Storm, L., Haile, S., Clayton, C., and Horn, D. (2005). A doubly inducible system for RNA interference and rapid RNAi plasmid construction in Trypanosoma brucei. Molecular and biochemical parasitology *139*, 75-82.

Alsford, S., duBois, K., Horn, D., and Field, M.C. (2012). Epigenetic mechanisms, nuclear architecture and the control of gene expression in trypanosomes. Expert reviews in molecular medicine *14*, e13.

Alsford, S., and 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 Res. *40*, 10150-10160.

Alsford, S., Kawahara, T., Glover, L., and 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.

Ansorge, I., Steverding, D., Melville, S., Hartmann, C., and Clayton, C. (1999). Transcription of 'inactive' expression sites in African trypanosomes leads to expression of multiple transferrin receptor RNAs in bloodstream forms. Mol Biochem Parasitol. *101*, 81-94.

Arhin, G.K., Shen, S., Ullu, E., and Tschudi, C. (2004). A PCR-based method for gene deletion and protein tagging in Trypanosoma brucei. Methods in molecular biology *270*, 277-286.

Asante, D., Stevenson, N.L., and Stephens, D.J. (2014). Subunit composition of the human cytoplasmic dynein-2 complex. Journal of cell science *127*, 4774-4787.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B.P., Carrington, M., Depledge, D.P., Fischer, S., Gajria, B., Gao, X., *et al.* (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res. *38*, D457-D462.

Asthana, J., Kuchibhatla, A., Jana, S.C., Ray, K., and Panda, D. (2012). Dynein light chain 1 (LC8) association enhances microtubule stability and promotes microtubule bundling. The Journal of biological chemistry *287*, 40793-40805.

Babokhov, P., Sanyaolu, A.O., Oyibo, W.A., Fagbenro-Beyioku, A.F., and Iriemenam, N.C. (2013). A current analysis of chemotherapy strategies for the treatment of human African trypanosomiasis. Pathogens and global health *107*, 242-252.

Badjatia, N., Ambrosio, D.L., Lee, J.H., and Günzl, A. (2013). Trypanosome cdc2-related kinase 9 controls spliced leader RNA cap4 methylation and phosphorylation of RNA polymerase II subunit RPB1. Molecular and cellular biology *33*, 1965-1975.

Barbar, E. (2008). Dynein light chain LC8 is a dimerization hub essential in diverse protein networks. Biochemistry *47*, 503-508.

Barbar, E., Kleinman, B., Imhoff, D., Li, M., Hays, T.S., and Hare, M. (2001). Dimerization and folding of LC8, a highly conserved light chain of cytoplasmic dynein. Biochemistry *40*, 1596-1605.

Barbar, E., and Nyarko, A. (2014). NMR Characterization of Self-Association Domains Promoted by Interactions with LC8 Hub Protein. Computational and structural biotechnology journal *9*, e201402003.

Barbar, E., and Nyarko, A. (2015). Polybivalency and disordered proteins in ordering macromolecular assemblies. Seminars in cell & developmental biology *37*, 20-25.

Bastin, P., Ellis, K., Kohl, L., and Gull, K. (2000). Flagellum ontogeny in trypanosomes studied via an inherited and regulated RNA interference system. Journal of cell science *113 (Pt 18)*, 3321-3328.

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

Benashski, S.E., Harrison, A., Patel-King, R.S., and King, S.M. (1997). Dimerization of the highly conserved light chain shared by dynein and myosin V. The Journal of biological chemistry *272*, 20929-20935.

Benison, G., Chiodo, M., Karplus, P.A., and Barbar, E. (2009). Structural, thermodynamic, and kinetic effects of a phosphomimetic mutation in dynein light chain LC8. Biochemistry *48*, 11381-11389.

Benmerzouga, I., Concepcion-Acevedo, J., Kim, H.S., Vandoros, A.V., Cross, G.A., Klingbeil, M.M., and 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.

Bentley, D.L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Current Opinion in Cell Biology *17*, 251-256.

Berberof, M., Perez-Morga, D., and Pays, E. (2001). A receptor-like flagellar pocket glycoprotein specific to Trypanosoma brucei gambiense. Molecular and biochemical parasitology *113*, 127-138.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., *et al.* (2005). The genome of the African trypanosome Trypanosoma brucei. Science *309*, 416-422.

Biebinger, S., Rettenmaier, S., Flaspohler, J., Hartmann, C., Pena-Diaz, J., Wirtz, L.E., Hotz, H.R., Barry, J.D., and Clayton, C. (1996). The PARP promoter of Trypanosoma brucei is developmentally regulated in a chromosomal context. Nucleic acids research *24*, 1202-1211.

Bilbe, G. (2015). Infectious diseases. Overcoming neglect of kinetoplastid diseases. Science *348*, 974-976.

Bitter, W., Gerrits, H., Kieft, R., and Borst, P. (1998). The role of transferrin-receptor variation in the host range of Trypanosoma brucei. Nature *391*, 499-502.

Blum, J., Schmid, C., and Burri, C. (2006). Clinical aspects of 2541 patients with second stage human African trypanosomiasis. Acta tropica *97*, 55-64.

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

Boisvert, F.-M., van Koningsbruggen, S., Navascués, J., and Lamond, A.I. (2007). The multifunctional nucleolus. Nature Reviews Molecular Cell Biology *8*, 574-585.

Brandao, A., and Jiang, T. (2009). The composition of untranslated regions in Trypanosoma cruzi genes. Parasitology international *58*, 215-219.

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

Broadhead, R., Dawe, H.R., Farr, H., Griffiths, S., Hart, S.R., Portman, N., Shaw, M.K., Ginger, M.L., Gaskell, S.J., McKean, P.G., *et al.* (2006). Flagellar motility is required for the viability of the bloodstream trypanosome. Nature *440*, 224-227.

Brown, S.D., Huang, J., and Van der Ploeg, L.H. (1992). The promoter for the procyclic acidic repetitive protein (PARP) genes of Trypanosoma brucei shares features with RNA polymerase I promoters. Molecular and cellular biology *12*, 2644-2652.

Brun, R., and Blum, J. (2012). Human African trypanosomiasis. Infectious disease clinics of North America *26*, 261-273.

Brun, R., Blum, J., Chappuis, F., and Burri, C. (2010). Human African trypanosomiasis. Lancet *375*, 148-159.

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

Chaves, I., Zomerdijk, J., Dirks, M.A., Dirks, R.W., Raap, A.K., and Borst, P. (1998). Subnuclear localization of the active variant surface glycoprotein gene expression site in

Trypanosoma brucei. Proceedings of the National Academy of Sciences of the United States of America *95*, 12328-12333.

Checchi, F., Filipe, J.A., Haydon, D.T., Chandramohan, D., and Chappuis, F. (2008). Estimates of the duration of the early and late stage of gambiense sleeping sickness. BMC infectious diseases *8*, 16.

Criswell, P.S., Ostrowski, L.E., and Asai, D.J. (1996). A novel cytoplasmic dynein heavy chain: expression of DHC1b in mammalian ciliated epithelial cells. Journal of cell science *109 (Pt 7)*, 1891-1898.

Cross, G.A., Kim, H.S., and Wickstead, B. (2014). Capturing the variant surface glycoprotein repertoire (the VSGnome) of Trypanosoma brucei Lister 427. Molecular and biochemical parasitology *195*, 59-73.

Cross, M., Wieland, B., Palfi, Z., Günzl, A., Rothlisberger, U., Lahm, H.W., and Bindereif, A. (1993). The trans-spliceosomal U2 snRNP protein 40K of Trypanosoma brucei: cloning and analysis of functional domains reveals homology to a mammalian snRNP protein. The EMBO journal *12*, 1239-1248.

Daniels, J.P., Gull, K., and Wickstead, B. (2012). The trypanosomatid-specific N terminus of RPA2 is required for RNA polymerase I assembly, localization, and function. Eukaryotic cell *11*, 662-672.

Das, A., Li, H., Liu, T., and Bellofatto, V. (2006). Biochemical characterization of Trypanosoma brucei RNA polymerase II. Molecular and biochemical parasitology *150*, 201-210.

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

Devaux, S., Lecordier, L., Uzureau, P., Walgraffe, D., Dierick, J.-F., Poelvoorde, P., Pays, E., and Vanhamme, L. (2006). Characterization of RNA polymerase II subunits of Trypanosoma brucei. Molecular and biochemical parasitology *148*, 60-68.

Dick, T., Ray, K., Salz, H.K., and Chia, W. (1996). Cytoplasmic dynein (ddlc1) mutations cause morphogenetic defects and apoptotic cell death in Drosophila melanogaster. Molecular and cellular biology *16*, 1966-1977.

Doyle, J.J., Hirumi, H., Hirumi, K., Lupton, E.N., and Cross, G.A. (1980). Antigenic variation in clones of animal-infective Trypanosoma brucei derived and maintained in vitro. Parasitology *80*, 359-369.

DuBois, K.N., Alsford, S., Holden, J.M., Buisson, J., Swiderski, M., Bart, J.M., Ratushny, A.V., Wan, Y., Bastin, P., Barry, J.D., *et al.* (2012). NUP-1 Is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. PLoS biology *10*, e1001287.

Edwards, A.M., Kane, C.M., Young, R.A., and Kornberg, R.D. (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. The Journal of biological chemistry *266*, 71-75.

Ehlers, B., Czichos, J., and Overath, P. (1987). RNA turnover in *Trypanosoma brucei*. Mol.Cell Biol. *7*, 1242-1249.

Engstler, M., Pfohl, T., Herminghaus, S., Boshart, M., Wiegertjes, G., Heddergott, N., and Overath, P. (2007). Hydrodynamic flow-mediated protein sorting on the cell surface of trypanosomes. Cell *131*, 505-515.

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

Ferrante, A., and Allison, A.C. (1983). Alternative pathway activation of complement by African trypanosomes lacking a glycoprotein coat. Parasite immunology *5*, 491-498.

Fevre, E.M., Picozzi, K., Jannin, J., Welburn, S.C., and Maudlin, I. (2006). Human African trypanosomiasis: Epidemiology and control. Advances in parasitology *61*, 167-221.

Figueiredo, L.M., and Cross, G.A. (2010). Nucleosomes are depleted at the VSG expression site transcribed by RNA polymerase I in African trypanosomes. Eukaryot.Cell *9*, 148-154.

Figueiredo, L.M., Janzen, C.J., and Cross, G.A. (2008). A histone methyltransferase modulates antigenic variation in African trypanosomes. PLoS.Biol. *6*, e161.

Franco, J.R., Simarro, P.P., Diarra, A., and Jannin, J.G. (2014). Epidemiology of human African trypanosomiasis. Clinical epidemiology *6*, 257-275.

Fridolfsson, H.N., Ly, N., Meyerzon, M., and Starr, D.A. (2010). UNC-83 coordinates kinesin-1 and dynein activities at the nuclear envelope during nuclear migration. Developmental biology *338*, 237-250.

Gallego, P., Velazquez-Campoy, A., Regue, L., Roig, J., and Reverter, D. (2013). Structural analysis of the regulation of the DYNLL/LC8 binding to Nek9 by phosphorylation. The Journal of biological chemistry *288*, 12283-12294.

Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Molecular biology and evolution *14*, 685-695.

Gibbons, B.H., Asai, D.J., Tang, W.J., Hays, T.S., and Gibbons, I.R. (1994). Phylogeny and expression of axonemal and cytoplasmic dynein genes in sea urchins. Molecular biology of the cell *5*, 57-70.

Ginger, M.L., Blundell, P.A., Lewis, A.M., Browitt, A., Günzl, A., and 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. Eukaryot.Cell *1*, 1000-1009.

Glover, L., and Horn, D. (2006). Repression of polymerase I-mediated gene expression at *Trypanosoma brucei* telomeres. EMBO reports *7*, 93-99.

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

Goggolidou, P., Stevens, J.L., Agueci, F., Keynton, J., Wheway, G., Grimes, D.T., Patel, S.H., Hilton, H., Morthorst, S.K., DiPaolo, A., *et al.* (2014). ATMIN is a transcriptional regulator of both lung morphogenesis and ciliogenesis. Development *141*, 3966-3977.

Gouy, M., Guindon, S., and Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular biology and evolution *27*, 221-224.

Grummt, I., and Skinner, J.A. (1985). Efficient transcription of a protein-coding gene from the RNA polymerase I promoter in transfected cells. Proc.Natl.Acad.Sci.U.S.A *82*, 722-726.

Guizetti, J., and Scherf, A. (2013). Silence, activate, poise and switch! Mechanisms of antigenic variation in Plasmodium falciparum. Cell Microbiol. *15*, 718-726.

Günzl, A. (2012). RNA Metabolism in Trypanosomes. In RNA Metabolism in Trypanosomes, A. Bindereif, ed., pp. 1-27.

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

Günzl, A., Bindereif, A., Ullu, E., and Tschudi, C. (2000). Determinants for cap trimethylation of the U2 small nuclear RNA are not conserved between Trypanosoma brucei and higher eukaryotic organisms. Nucleic acids research *28*, 3702-3709.

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

Günzl, A., Cross, M., and Bindereif, A. (1992). Domain structure of U2 and U4/U6 small nuclear ribonucleoprotein particles from Trypanosoma brucei: identification of trans-spliceosomal specific RNA-protein interactions. Molecular and cellular biology *12*, 468-479.

Günzl, A., Kirkham, J.K., Nguyen, T.N., Badjatia, N., and Park, S.H. (2015). Mono-allelic VSG expression by RNA polymerase I in Trypanosoma brucei: expression site control from both ends? Gene *556*, 68-73.

Günzl, A., Park, S.H., Nguyen, T.N., Kirkham, J.K., and Lee, J.H. (2011). Response to "Role of RPB7 in RNA pol I transcription in Trypanosoma brucei". Molecular and biochemical parasitology *180*, 45-46.

Hall, J.P., Wang, H., and Barry, J.D. (2013). Mosaic VSGs and the scale of Trypanosoma brucei antigenic variation. PLoS pathogens *9*, e1003502.

Happel, N., and Doenecke, D. (2009). Histone H1 and its isoforms: contribution to chromatin structure and function. Gene *431*, 1-12.

Harel-Sharvit, L., Eldad, N., Haimovich, G., Barkai, O., Duek, L., and Choder, M. (2010). RNA Polymerase II Subunits Link Transcription and mRNA Decay to Translation. Cell *143*, 552-563.

Hertz-Fowler, C., Figueiredo, L.M., Quail, M.A., Becker, M., Jackson, A., Bason, N., Brooks, K., Churcher, C., Fahkro, S., Goodhead, I., *et al.* (2008). Telomeric expression sites are highly conserved in *Trypanosoma brucei*. PLoS.One. *3*, e3527.

Hirumi, H., and 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.

Hodi, Z., Nemeth, A.L., Radnai, L., Hetenyi, C., Schlett, K., Bodor, A., Perczel, A., and Nyitray, L. (2006). Alternatively spliced exon B of myosin Va is essential for binding the tail-associated light chain shared by dynein. Biochemistry *45*, 12582-12595.

Horn, D. (2014). Antigenic variation in African trypanosomes. Molecular and biochemical parasitology *195*, 123-129.

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

Horn, D., and McCulloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. Curr.Opin.Microbiol. *13*, 700-705.

Hou, Y., and Witman, G.B. (2015). Dynein and intraflagellar transport. Experimental cell research *334*, 26-34.

Hughes, K., Wand, M., Foulston, L., Young, R., Harley, K., Terry, S., Ersfeld, K., and Rudenko, G. (2007). A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. EMBO J. *26*, 2400-2410.

Inoue, M., Nakamura, Y., Yasuda, K., Yasaka, N., Hara, T., Schnaufer, A., Stuart, K., and Fukuma, T. (2005). The 14-3-3 proteins of Trypanosoma brucei function in motility, cytokinesis, and cell cycle. The Journal of biological chemistry *280*, 14085-14096.

Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., *et al.* (2005). The genome of the kinetoplastid parasite, Leishmania major. Science *309*, 436-442.

Janz, L., and Clayton, C. (1994). The PARP and rRNA promoters of *Trypanosoma brucei* are composed of dissimilar sequence elements that are functionally interchangeable. Mol.Cell.Biol. *14*, 5804-5811.

Janzen, C.J., Hake, S.B., Lowell, J.E., and Cross, G.A. (2006). Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in Trypanosoma brucei. Molecular cell *23*, 497-507.

Jasiak, A.J., Hartmann, H., Karakasili, E., Kalocsay, M., Flatley, A., Kremmer, E., Strasser, K., Martin, D.E., Soding, J., and Cramer, P. (2008). Genome-associated RNA Polymerase II Includes the Dissociable Rpb4/7 Subcomplex. Journal of Biological Chemistry *283*, 26423-26427.

Jayappa, K.D., Ao, Z., Wang, X., Mouland, A.J., Shekhar, S., Yang, X., and Yao, X. (2015). Human immunodeficiency virus type 1 employs the cellular dynein light chain 1 protein for reverse transcription through interaction with its integrase protein. Journal of virology *89*, 3497-3511.

Jess, W., Hammer, A., and Cornelissen, A.W. (1989). Complete sequence of the gene encoding the largest subunit of RNA polymerase I of Trypanosoma brucei. FEBS Lett *249*, 123-128.

Jung, Y., Kim, H., Min, S.H., Rhee, S.G., and Jeong, W. (2008). Dynein light chain LC8 negatively regulates NF-kappaB through the redox-dependent interaction with IkappaBalpha. The Journal of biological chemistry *283*, 23863-23871.

Jurado, S., Gleeson, K., O'Donnell, K., Izon, D.J., Walkley, C.R., Strasser, A., Tarlinton, D.M., and Heierhorst, J. (2012). The Zinc-finger protein ASCIZ regulates B cell development via DYNLL1 and Bim. The Journal of experimental medicine *209*, 1629-1639.

Kaiser, F.J., Tavassoli, K., Van den Bemd, G.J., Chang, G.T., Horsthemke, B., Moroy, T., and Ludecke, H.J. (2003). Nuclear interaction of the dynein light chain LC8a with the TRPS1 transcription factor suppresses the transcriptional repression activity of TRPS1. Human molecular genetics *12*, 1349-1358.

Kalidas, S., Li, Q., and Phillips, M.A. (2011). A Gateway(R) compatible vector for gene silencing in bloodstream form Trypanosoma brucei. Molecular and biochemical parasitology *178*, 51-55.

Kassem, A., Pays, E., and 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.

Kato, C.D., Nanteza, A., Mugasa, C., Edyelu, A., Matovu, E., and Alibu, V.P. (2015). Clinical profiles, disease outcome and co-morbidities among T. b. rhodesiense sleeping sickness patients in Uganda. PloS one *10*, e0118370.

Kelly, S., Wickstead, B., and Gull, K. (2005). An in silico analysis of trypanosomatid RNA polymerases: insights into their unusual transcription. Biochem Soc Trans *33*, 1435-1437.

Kidane, A.I., Song, Y., Nyarko, A., Hall, J., Hare, M., Lohr, F., and Barbar, E. (2013). Structural features of LC8-induced self-association of swallow. Biochemistry *52*, 6011-6020.

Kim, H.S., Li, Z., Boothroyd, C., and Cross, G.A. (2013a). Strategies to construct null and conditional null Trypanosoma brucei mutants using Cre-recombinase and loxP. Molecular and biochemical parasitology *191*, 16-19.

Kim, H.S., Park, S.H., Günzl, A., and Cross, G.A. (2013b). 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*, e57001.

King, S.M. (2008). Dynein-independent functions of DYNLL1/LC8: redox state sensing and transcriptional control. Science signaling *1*, pe51.

King, S.M., Barbarese, E., Dillman, J.F., 3rd, Patel-King, R.S., Carson, J.H., and Pfister, K.K. (1996). Brain cytoplasmic and flagellar outer arm dyneins share a highly conserved Mr 8,000 light chain. The Journal of biological chemistry *271*, 19358-19366.

King, S.M., and Patel-King, R.S. (1995). The M(r) = 8,000 and 11,000 outer arm dynein light chains from Chlamydomonas flagella have cytoplasmic homologues. The Journal of biological chemistry *270*, 11445-11452.

Kirkman, L.A., and Deitsch, K.W. (2012). Antigenic variation and the generation of diversity in malaria parasites. Current opinion in microbiology *15*, 456-462.

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

Kolev, N.G., Tschudi, C., and Ullu, E. (2011). RNA interference in protozoan parasites: achievements and challenges. Eukaryotic cell *10*, 1156-1163.

Kooter, J.M., and Borst, P. (1984). Alpha-amanitin-insensitive transcription of variant surface glycoprotein genes provides further evidence for discontinuous transcription in trypanosomes. Nucleic Acids Res. *12*, 9457-9472.

Kuhn, C.D., Geiger, S.R., Baumli, S., Gartmann, M., Gerber, J., Jennebach, S., Mielke, T., Tschochner, H., Beckmann, R., and Cramer, P. (2007). Functional architecture of RNA polymerase I. Cell *131*, 1260-1272.

Lacomble, S., Portman, N., and Gull, K. (2009). A protein-protein interaction map of the Trypanosoma brucei paraflagellar rod. PloS one *4*, e7685.

LaCount, D.J., Bruse, S., Hill, K.L., and Donelson, J.E. (2000). Double-stranded RNA interference in Trypanosoma brucei using head-to-head promoters. Molecular and biochemical parasitology *111*, 67-76.

Landeira, D., Bart, J.M., Van Tyne, D., and Navarro, M. (2009). Cohesin regulates VSG monoallelic expression in trypanosomes. The Journal of cell biology *186*, 243-254.

Landeira, D., and Navarro, M. (2007). Nuclear repositioning of the VSG promoter during developmental silencing in Trypanosoma brucei. The Journal of cell biology *176*, 133-139.

Lasda, E.L., and Blumenthal, T. (2011). Trans-splicing. Wiley interdisciplinary reviews. RNA *2*, 417-434.

Laufer, G., and Günzl, A. (2001). In-vitro competition analysis of procyclin gene and variant surface glycoprotein gene expression site transcription in Trypanosoma brucei. Molecular and biochemical parasitology *113*, 55-65.

Laufer, G., Schaaf, G., Bollgonn, S., and Günzl, A. (1999). In vitro analysis of alpha-amanitin-resistant transcription from the rRNA, procyclic acidic repetitive protein, and variant surface glycoprotein gene promoters in Trypanosoma brucei. Molecular and cellular biology *19*, 5466-5473.

Lee, J.H., Cai, G., Panigrahi, A.K., Dunham-Ems, S., Nguyen, T.N., Radolf, J.D., Asturias, F.J., and Günzl, A. (2010). A TFIIH-associated mediator head is a basal factor of small nuclear spliced leader RNA gene transcription in early-diverged trypanosomes. Molecular and cellular biology *30*, 5502-5513.

Lee, J.H., Jung, H.S., and Günzl, A. (2009). Transcriptionally active TFIIH of the early-diverged eukaryote Trypanosoma brucei harbors two novel core subunits but not a cyclin-activating kinase complex. Nucleic acids research *37*, 3811-3820.

Lee, M.G. (1996). An RNA polymerase II promoter in the hsp70 locus of Trypanosoma brucei. Molecular and cellular biology *16*, 1220-1230.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nature Protocols *1*, 729-748.

Lei, K., and Davis, R.J. (2003). JNK phosphorylation of Bim-related members of the Bcl2 family induces Bax-dependent apoptosis. Proceedings of the National Academy of Sciences of the United States of America *100*, 2432-2437.

Li, W., Yi, P., and Ou, G. (2015). Somatic CRISPR-Cas9-induced mutations reveal roles of embryonically essential dynein chains in Caenorhabditis elegans cilia. The Journal of cell biology *208*, 683-692.

Li, Z., Lee, J.H., Chu, F., Burlingame, A.L., Günzl, A., and Wang, C.C. (2008). Identification of a novel chromosomal passenger complex and its unique localization during cytokinesis in Trypanosoma brucei. PloS one *3*, e2354.

Li, Z., Umeyama, T., and Wang, C.C. (2009). The Aurora Kinase in Trypanosoma brucei plays distinctive roles in metaphase-anaphase transition and cytokinetic initiation. PLoS pathogens *5*, e1000575.

Lightcap, C.M., Kari, G., Arias-Romero, L.E., Chernoff, J., Rodeck, U., and Williams, J.C. (2009). Interaction with LC8 is required for Pak1 nuclear import and is indispensable for zebrafish development. PloS one *4*, e6025.

Liniger, M., Bodenmuller, K., Pays, E., Gallati, S., and Roditi, I. (2001). Overlapping sense and antisense transcription units in Trypanosoma brucei. Molecular microbiology *40*, 869-878.

Liu, B., Xiang, X., and Lee, Y.R. (2003). The requirement of the LC8 dynein light chain for nuclear migration and septum positioning is temperature dependent in Aspergillus nidulans. Molecular microbiology *47*, 291-301.

Lo, K.W., Naisbitt, S., Fan, J.S., Sheng, M., and Zhang, M. (2001). The 8-kDa dynein light chain binds to its targets via a conserved (K/R)XTQT motif. The Journal of biological chemistry *276*, 14059-14066.

Logan-Klumpler, F.J., De, S.N., Boehme, U., Rogers, M.B., Velarde, G., McQuillan, J.A., Carver, T., Aslett, M., Olsen, C., Subramanian, S., *et al.* (2012). GeneDB--an annotation database for pathogens. Nucleic Acids Res. *40*, D98-108.

Lotan, R., Goler-Baron, V., Duek, L., Haimovich, G., and Choder, M. (2007). The Rpb7p subunit of yeast RNA polymerase II plays roles in the two major cytoplasmic mRNA decay mechanisms. The Journal of cell biology *178*, 1133-1143.

Luthra, P., Jordan, D.S., Leung, D.W., Amarasinghe, G.K., and Basler, C.F. (2015). Ebola virus VP35 interaction with dynein LC8 regulates viral RNA synthesis. Journal of virology *89*, 5148-5153.

Lutje, V., Seixas, J., and Kennedy, A. (2010). Chemotherapy for second-stage Human African trypanosomiasis. The Cochrane database of systematic reviews, CD006201.

Luz Ambrosio, D., Lee, J.H., Panigrahi, A.K., Nguyen, T.N., Cicarelli, R.M., and Günzl, A. (2009). Spliceosomal proteomics in Trypanosoma brucei reveal new RNA splicing factors. Eukaryotic cell *8*, 990-1000.

Magklara, A., and Lomvardas, S. (2013). Stochastic gene expression in mammals: lessons from olfaction. Trends Cell Biol.

Makokha, M., Hare, M., Li, M., Hays, T., and Barbar, E. (2002). Interactions of cytoplasmic dynein light chains Tctex-1 and LC8 with the intermediate chain IC74. Biochemistry *41*, 4302-4311.

Manna, P.T., Boehm, C., Leung, K.F., Natesan, S.K., and Field, M.C. (2014). Life and times: synthesis, trafficking, and evolution of VSG. Trends in parasitology *30*, 251-258.

Martinez-Calvillo, S., Saxena, A., Green, A., Leland, A., and Myler, P.J. (2007). Characterization of the RNA polymerase II and III complexes in *Leishmania major*. Int.J.Parasitol. *37*, 491-502.

Matthews, K.R. (2015). 25 years of African trypanosome research: From description to molecular dissection and new drug discovery. Molecular and biochemical parasitology *200*, 30-40.

Maudlin, I., and Welburn, S.C. (1989). A single trypanosome is sufficient to infect a tsetse fly. Annals of tropical medicine and parasitology *83*, 431-433.

McConville, M.J., and Ferguson, M.A. (1993). The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. The Biochemical journal *294 (Pt 2)*, 305-324.

Michaeli, S. (2011). Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome. Future.Microbiol. *6*, 459-474.

Milkereit, P., and Tschochner, H. (1998). A specialized form of RNA polymerase I, essential for initiation and growth-dependent regulation of rRNA synthesis, is disrupted during transcription. The EMBO journal *17*, 3692-3703.

Mitsuzawa, H., Kanda, E., and Ishihama, A. (2003). Rpb7 subunit of RNA polymerase II interacts with an RNA-binding protein involved in processing of transcripts. Nucleic acids research *31*, 4696-4701.

Morris, J.C., Wang, Z., Drew, M.E., and Englund, P.T. (2002). Glycolysis modulates trypanosome glycoprotein expression as revealed by an RNAi library. The EMBO journal *21*, 4429-4438.

Mugnier, M.R., Cross, G.A., and Papavasiliou, F.N. (2015). The in vivo dynamics of antigenic variation in Trypanosoma brucei. Science *347*, 1470-1473.

Naisbitt, S., Valtschanoff, J., Allison, D.W., Sala, C., Kim, E., Craig, A.M., Weinberg, R.J., and Sheng, M. (2000). Interaction of the postsynaptic density-95/guanylate kinase domain-associated

protein complex with a light chain of myosin-V and dynein. The Journal of neuroscience : the official journal of the Society for Neuroscience *20*, 4524-4534.

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

Narayanan, M.S., and Rudenko, G. (2013). TDP1 is an HMG chromatin protein facilitating RNA polymerase I transcription in African trypanosomes. Nucleic Acids Res. *41*, 2981-2992.

Navarro, M., and Cross, G.A. (1996). DNA rearrangements associated with multiple consecutive directed antigenic switches in Trypanosoma brucei. Mol.Cell Biol. *16*, 3615-3625.

Navarro, M., and Cross, G.A. (1998). In situ analysis of a variant surface glycoprotein expression-site promoter region in Trypanosoma brucei. Mol Biochem Parasitol. *94*, 53-66.

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

Navarro, M., Penate, X., Landeira, D., and Lopez-Farfan, D. (2011). Role of RPB7 in RNA pol I transcription in Trypanosoma brucei. Molecular and biochemical parasitology *180*, 43-44.

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in Trypanosoma brucei. Proceedings of the National Academy of Sciences of the United States of America *95*, 14687-14692.

Nguyen, T.N., Muller, L.S., Park, S.H., Siegel, T.N., and 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., and 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*. Eukaryot. Cell *11*, 1573-1581.

Nguyen, T.N., Schimanski, B., and Günzl, A. (2007). Active RNA polymerase I of Trypanosoma brucei harbors a novel subunit essential for transcription. Molecular and cellular biology *27*, 6254-6263.

Nguyen, T.N., Schimanski, B., Zahn, A., Klumpp, B., and Günzl, A. (2006). Purification of an eight subunit RNA polymerase I complex in Trypanosoma brucei. Molecular and biochemical parasitology *149*, 27-37.

Nyarko, A., Hare, M., Hays, T.S., and Barbar, E. (2004). The intermediate chain of cytoplasmic dynein is partially disordered and gains structure upon binding to light-chain LC8. Biochemistry *43*, 15595-15603.

Nyarko, A., Song, Y., Novacek, J., Zidek, L., and Barbar, E. (2013). Multiple recognition motifs in nucleoporin Nup159 provide a stable and rigid Nup159-Dyn2 assembly. The Journal of biological chemistry *288*, 2614-2622.

Oberholzer, M., Langousis, G., Nguyen, H.T., Saada, E.A., Shimogawa, M.M., Jonsson, Z.O., Nguyen, S.M., Wohlschlegel, J.A., and Hill, K.L. (2011). Independent analysis of the flagellum surface and matrix proteomes provides insight into flagellum signaling in mammalian-infectious Trypanosoma brucei. Molecular & cellular proteomics : MCP *10*, M111 010538.

Pace, C.N., and Scholtz, J.M. (1998). A helix propensity scale based on experimental studies of peptides and proteins. Biophysical journal *75*, 422-427.

Pal, A., Hall, B.S., Jeffries, T.R., and Field, M.C. (2003). Rab5 and Rab11 mediate transferrin and anti-variant surface glycoprotein antibody recycling in Trypanosoma brucei. The Biochemical journal *374*, 443-451.

Palenchar, J.B., Liu, W., Palenchar, P.M., and Bellofatto, V. (2006). A Divergent Transcription Factor TFIIB in Trypanosomes Is Required for RNA Polymerase II-Dependent Spliced Leader RNA Transcription and Cell Viability. Eukaryotic cell *5*, 293-300.

Park, S.H., Nguyen, B.N., Kirkham, J.K., Nguyen, T.N., and 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*, 785-795.

Park, S.H., Nguyen, T.N., Kirkham, J.K., Lee, J.H., and Günzl, A. (2011). Transcription by the multifunctional RNA polymerase I in Trypanosoma brucei functions independently of RPB7. Molecular and biochemical parasitology *180*, 35-42.

Paschal, B.M., and Vallee, R.B. (1987). Retrograde transport by the microtubule-associated protein MAP 1C. Nature *330*, 181-183.

Patrick, K.L., Luz, P.M., Ruan, J.P., Shi, H., Ullu, E., and Tschudi, C. (2008). Genomic rearrangements and transcriptional analysis of the spliced leader-associated retrotransposon in RNA interference-deficient Trypanosoma brucei. Mol.Microbiol. *67*, 435-447.

Pays, E., Vanhollebeke, B., Vanhamme, L., Paturiaux-Hanocq, F., Nolan, D.P., and Perez-Morga, D. (2006). The trypanolytic factor of human serum. Nature reviews. Microbiology *4*, 477-486.

Pazour, G.J., Wilkerson, C.G., and Witman, G.B. (1998). A dynein light chain is essential for the retrograde particle movement of intraflagellar transport (IFT). The Journal of cell biology *141*, 979-992.

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., *et al.* (2014). Trypanosoma

brucei histone H1 inhibits RNA polymerase I transcription and is important for parasite fitness in vivo. Molecular microbiology *93*, 645-663.

Peñate, X., López-Farfán, D., Landeira, D., Wentland, A., Vidal, I., and Navarro, M. (2009). RNA pol II subunit RPB7 is required for RNA pol I-mediated transcription in Trypanosoma brucei. EMBO reports *10*, 252-257.

Peyroche, G., Levillain, E., Siaut, M., Callebaut, I., Schultz, P., Sentenac, A., Riva, M., and Carles, C. (2002). The A14-A43 heterodimer subunit in yeast RNA pol I and their relationship to Rpb4-Rpb7 pol II subunits. Proceedings of the National Academy of Sciences *99*, 14670-14675.

Peyroche, G., Milkereit, P., Bischler, N., Tschochner, H., Schultz, P., Sentenac, A., Carles, C., and Riva, M. (2000). The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3. The EMBO journal *19*, 5473-5482.

Pfister, K.K., Fay, R.B., and Witman, G.B. (1982). Purification and polypeptide composition of dynein ATPases from Chlamydomonas flagella. Cell motility *2*, 525-547.

Pfister, K.K., Shah, P.R., Hummerich, H., Russ, A., Cotton, J., Annuar, A.A., King, S.M., and Fisher, E.M. (2006). Genetic analysis of the cytoplasmic dynein subunit families. PLoS genetics *2*, e1.

Pham, V.P., Qi, C.C., and Gottesdiener, K.M. (1996b). A detailed mutational analysis of the VSG gene expression site promoter. Mol.Biochem.Parasitol. *75*, 241-254.

Pham, V.P., Rothman, P.B., and Gottesdiener, K.M. (1997). Binding of trans-acting factors to the double-stranded variant surface glycoprotein (VSG) expression site promoter of Trypanosoma brucei. Molecular and biochemical parasitology *89*, 11-23.

Povelones, M.L., Gluenz, E., Dembek, M., Gull, K., and Rudenko, G. (2012). Histone H1 plays a role in heterochromatin formation and VSG expression site silencing in Trypanosoma brucei. PLoS.Pathog. *8*, e1003010.

Preußer, C., Jae, N., Günzl, A., and Bindereif, A. (2012). Pre-mRNA splicing in *Trypanosoma brucei*: factors, Mechanisms, and Regulation. In RNA Metabolism in Trypanosomes, A. Bindereif, ed. (Springer Press), pp. 49-76.

Qureshi, B.M., Hofmann, N.E., Arroyo-Olarte, R.D., Nickl, B., Hoehne, W., Jungblut, P.R., Lucius, R., Scheerer, P., and Gupta, N. (2013). Dynein light chain 8a of Toxoplasma gondii, a unique conoid-localized beta-strand-swapped homodimer, is required for an efficient parasite growth. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *27*, 1034-1047.

Radnai, L., Rapali, P., Hodi, Z., Suveges, D., Molnar, T., Kiss, B., Becsi, B., Erdodi, F., Buday, L., Kardos, J., *et al.* (2010). Affinity, avidity, and kinetics of target sequence binding to LC8 dynein light chain isoforms. The Journal of biological chemistry *285*, 38649-38657.

Rapali, P., Radnai, L., Suveges, D., Harmat, V., Tolgyesi, F., Wahlgren, W.Y., Katona, G., Nyitray, L., and Pal, G. (2011a). 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*, e18818.

Rapali, P., Szenes, A., Radnai, L., Bakos, A., Pal, G., and Nyitray, L. (2011b). DYNLL/LC8: a light chain subunit of the dynein motor complex and beyond. The FEBS journal *278*, 2980-2996.

Rayala, S.K., den Hollander, P., Balasenthil, S., Yang, Z., Broaddus, R.R., and Kumar, R. (2005). Functional regulation of oestrogen receptor pathway by the dynein light chain 1. EMBO reports *6*, 538-544.

Rayala, S.K., den Hollander, P., Manavathi, B., Talukder, A.H., Song, C., Peng, S., Barnekow, A., Kremerskothen, J., and Kumar, R. (2006). Essential role of KIBRA in co-activator function of dynein light chain 1 in mammalian cells. The Journal of biological chemistry *281*, 19092-19099.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. Nature biotechnology *17*, 1030-1032.

Robays, J., Nyamowala, G., Sese, C., Betu Ku Mesu Kande, V., Lutumba, P., Van der Veken, W., and Boelaert, M. (2008). High failure rates of melarsoprol for sleeping sickness, Democratic Republic of Congo. Emerging infectious diseases *14*, 966-967.

Ross, R., and Thomson, D. (1910). A Case of Sleeping Sickness showing Regular Periodical Increase of the Parasites Disclosed. British medical journal *1*, 1544-1545.

Rost, B., Yachdav, G., and Liu, J. (2004). The PredictProtein server. Nucleic acids research *32*, W321-326.

Ruan, J.-p., Ullu, E., and Tschudi, C. (2007). Characterization of the Trypanosoma brucei cap hypermethylase Tgs1. Molecular and biochemical parasitology *155*, 66-69.

Rudenko, G. (2010). Epigenetics and transcriptional control in African trypanosomes. Essays Biochem. *48*, 201-219.

Rudenko, G., Chung, H.M., Pham, V.P., and Van der Ploeg, L.H. (1991a). RNA polymerase I can mediate expression of CAT and neo protein- coding genes in Trypanosoma brucei. EMBO J. *10*, 3387-3397.

Ruepp, S., Furger, A., Kurath, U., Renggli, C.K., Hemphill, A., Brun, R., and Roditi, I. (1997). Survival of Trypanosoma brucei in the tsetse fly is enhanced by the expression of specific forms of procyclin. The Journal of cell biology *137*, 1369-1379.

Russell, J., and Zomerdijk, J.C.B.M. (2005). RNA-polymerase-I-directed rDNA transcription, life and works. Trends in Biochemical Sciences *30*, 87-96.

Salmon, D., Vanwalleghem, G., Morias, Y., Denoeud, J., Krumbholz, C., Lhomme, F., Bachmaier, S., Kador, M., Gossmann, J., Dias, F.B., *et al.* (2012). Adenylate cyclases of Trypanosoma brucei inhibit the innate immune response of the host. Science *337*, 463-466.

Schimanski, B., Brandenburg, J., Nguyen, T.N., Caimano, M.J., and Günzl, A. (2006). A TFIIB-like protein is indispensable for spliced leader RNA gene transcription in Trypanosoma brucei. Nucleic acids research *34*, 1676-1684.

Schimanski, B., Klumpp, B., Laufer, G., Marhöfer, R.J., Selzer, P.M., and Günzl, A. (2003). The second largest subunit of Trypanosoma brucei's multifunctional RNA polymerase I has a unique N-terminal extension domain. Molecular and biochemical parasitology *126*, 193-200.

Schimanski, B., Laufer, G., Gontcharova, L., and Günzl, A. (2004). The *Trypanosoma brucei* spliced leader RNA and rRNA gene promoters have interchangeable TbSNAP50-binding elements. Nucleic Acids Res. *32*, 700-709.

Schimanski, B., Nguyen, T.N., and Günzl, A. (2005a). Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in Trypanosoma brucei. Molecular and cellular biology *25*, 7303-7313.

Schimanski, B., Nguyen, T.N., and Günzl, A. (2005b). Highly Efficient Tandem Affinity Purification of Trypanosome Protein Complexes Based on a Novel Epitope Combination. Eukaryotic cell *4*, 1942-1950.

Schmid, C., Richer, M., Bilenge, C.M., Josenando, T., Chappuis, F., Manthelot, C.R., Nangouma, A., Doua, F., Asumu, P.N., Simarro, P.P., *et al.* (2005). Effectiveness of a 10-day melarsoprol schedule for the treatment of late-stage human African trypanosomiasis: confirmation from a multinational study (IMPAMEL II). The Journal of infectious diseases *191*, 1922-1931.

Schmitz, K.M., Mayer, C., Postepska, A., and Grummt, I. (2010). Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. Genes & Development *24*, 2264-2269.

Schwede, A., and Carrington, M. (2010). Bloodstream form Trypanosome plasma membrane proteins: antigenic variation and invariant antigens. Parasitology *137*, 2029-2039.

Schwede, A., Jones, N., Engstler, M., and Carrington, M. (2011). The VSG C-terminal domain is inaccessible to antibodies on live trypanosomes. Mol.Biochem.Parasitol. *175*, 201-204.

Sheader, K., Vaughan, S., Minchin, J., Hughes, K., Gull, K., and Rudenko, G. (2005). Variant surface glycoprotein RNA interference triggers a precytokinesis cell cycle arrest in African

trypanosomes. Proceedings of the National Academy of Sciences of the United States of America *102*, 8716-8721.

Sherman, D.R., Janz, L., Hug, M., and Clayton, C. (1991). Anatomy of the parp gene promoter of Trypanosoma brucei. The EMBO journal *10*, 3379-3386.

Shi, H., Djikeng, A., Mark, T., Wirtz, E., Tschudi, C., and Ullu, E. (2000). Genetic interference in Trypanosoma brucei by heritable and inducible double-stranded RNA. Rna *6*, 1069-1076.

Shiflett, A.M., Faulkner, S.D., Cotlin, L.F., Widener, J., Stephens, N., and Hajduk, S.L. (2007). African trypanosomes: intracellular trafficking of host defense molecules. The Journal of eukaryotic microbiology *54*, 18-21.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology *7*, 539.

Smith, J.L., Levin, J.R., and Agabian, N. (1989). Molecular characterization of the Trypanosoma brucei RNA polymerase I and III largest subunit genes. The Journal of biological chemistry *264*, 18091-18099.

Song, C., Wen, W., Rayala, S.K., Chen, M., Ma, J., Zhang, M., and Kumar, R. (2008). Serine 88 phosphorylation of the 8-kDa dynein light chain 1 is a molecular switch for its dimerization status and functions. The Journal of biological chemistry *283*, 4004-4013.

Song, Y., Benison, G., Nyarko, A., Hays, T.S., and Barbar, E. (2007). Potential role for phosphorylation in differential regulation of the assembly of dynein light chains. The Journal of biological chemistry *282*, 17272-17279.

Stanne, T.M., Kushwaha, M., Wand, M., Taylor, J.E., and Rudenko, G. (2011). TbISWI regulates multiple polymerase I (Pol I)-transcribed loci and is present at Pol II transcription boundaries in Trypanosoma brucei. Eukaryotic cell *10*, 964-976.

Stanne, T.M., and Rudenko, G. (2010). Active VSG expression sites in Trypanosoma brucei are depleted of nucleosomes. Eukaryot.Cell *9*, 136-147.

Stearns, D.J., Kurosawa, S., Sims, P.J., Esmon, N.L., and Esmon, C.T. (1988). The interaction of a Ca2+-dependent monoclonal antibody with the protein C activation peptide region. Evidence for obligatory Ca2+ binding to both antigen and antibody. The Journal of biological chemistry *263*, 826-832.

Stelter, P., Kunze, R., Flemming, D., Hopfner, D., Diepholz, M., Philippsen, P., Bottcher, B., and Hurt, E. (2007). Molecular basis for the functional interaction of dynein light chain with the nuclear-pore complex. Nature cell biology *9*, 788-796.

Steverding, D. (2008). The history of African trypanosomiasis. Parasites & vectors *1*, 3.

Stewart, M., Haile, S., Jha, B.A., Cristodero, M., Li, C.H., and Clayton, C. (2010). Processing of a phosphoglycerate kinase reporter mRNA in *Trypanosoma brucei* is not coupled to transcription by RNA polymerase II. Mol.Biochem.Parasitol. *172*, 99-106.

Stuart, K., Brun, R., Croft, S., Fairlamb, A., Gurtler, R.E., McKerrow, J., Reed, S., and Tarleton, R. (2008). Kinetoplastids: related protozoan pathogens, different diseases. The Journal of clinical investigation *118*, 1301-1310.

Stuchell-Brereton, M.D., Siglin, A., Li, J., Moore, J.K., Ahmed, S., Williams, J.C., and Cooper, J.A. (2011). Functional interaction between dynein light chain and intermediate chain is required for mitotic spindle positioning. Molecular biology of the cell *22*, 2690-2701.

Sunter, J., Wickstead, B., Gull, K., and Carrington, M. (2012). A new generation of T7 RNA polymerase-independent inducible expression plasmids for Trypanosoma brucei. PloS one *7*, e35167.

Tan, G.S., Preuss, M.A., Williams, J.C., and Schnell, M.J. (2007). The dynein light chain 8 binding motif of rabies virus phosphoprotein promotes efficient viral transcription. Proceedings of the National Academy of Sciences of the United States of America *104*, 7229-7234.

Thuita, J.K., Kagira, J.M., Mwangangi, D., Matovu, E., Turner, C.M., and Masiga, D. (2008). Trypanosoma brucei rhodesiense transmitted by a single tsetse fly bite in vervet monkeys as a model of human African trypanosomiasis. PLoS neglected tropical diseases *2*, e238.

Tiengwe, C., Marcello, L., Farr, H., Dickens, N., Kelly, S., Swiderski, M., Vaughan, D., Gull, K., Barry, J.D., BELL, S.D., *et al.* (2012). Genome-wide analysis reveals extensive functional interaction between DNA replication initiation and transcription in the genome of Trypanosoma brucei. Cell Rep. *2*, 185-197.

Tschudi, C., Djikeng, A., Shi, H., and Ullu, E. (2003). In vivo analysis of the RNA interference mechanism in Trypanosoma brucei. Methods *30*, 304-312.

Turner, C.M. (1997). The rate of antigenic variation in fly-transmitted and syringe-passaged infections of Trypanosoma brucei. FEMS microbiology letters *153*, 227-231.

Turner, C.M., and Barry, J.D. (1989). High frequency of antigenic variation in Trypanosoma brucei rhodesiense infections. Parasitology *99 Pt 1*, 67-75.

Újvári, A., and Luse, D.S. (2005). RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. Nat Struct Mol Biol *13*, 49-54.

Uzureau, P., Uzureau, S., Lecordier, L., Fontaine, F., Tebabi, P., Homble, F., Grelard, A., Zhendre, V., Nolan, D.P., Lins, L., *et al.* (2013). Mechanism of Trypanosoma brucei gambiense resistance to human serum. Nature *501*, 430-434.

Vadlamudi, R.K., Bagheri-Yarmand, R., Yang, Z., Balasenthil, S., Nguyen, D., Sahin, A.A., den Hollander, P., and Kumar, R. (2004). Dynein light chain 1, a p21-activated kinase 1-interacting substrate, promotes cancerous phenotypes. Cancer cell *5*, 575-585.

Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D.P., Lins, L., Van Den Abbeele, J., Pays, A., Tebabi, P., Van Xong, H., Jacquet, A., *et al.* (2003). Apolipoprotein L-I is the trypanosome lytic factor of human serum. Nature *422*, 83-87.

Vanhamme, L., Pays, A., Tebabi, P., Alexandre, S., and Pays, E. (1995). Specific binding of proteins to the noncoding strand of a crucial element of the variant surface glycoprotein, procyclin, and ribosomal promoters of *Trypanosoma brucei*. Mol.Cell.Biol. *15*, 5598-5606.

Vanhamme, L., Poelvoorde, P., Pays, A., Tebabi, P., Xong, H.V., and Pays, E. (2000). Differential RNA elongation controls the variant surface glycoprotein gene expression sites of Trypanosoma brucei. Molecular microbiology *36*, 328-340.

Vazquez, M., Atorrasagasti, C., Bercovich, N., Volcovich, R., and Levin, M.J. (2003). Unique features of the Trypanosoma cruzi U2AF35 splicing factor. Molecular and biochemical parasitology *128*, 77-81.

Vazquez, M.P., Mualem, D., Bercovich, N., Stern, M.Z., Nyambega, B., Barda, O., Nasiga, D., Gupta, S.K., Michaeli, S., and Levin, M.J. (2009). Functional characterization and protein-protein interactions of trypanosome splicing factors U2AF35, U2AF65 and SF1. Molecular and biochemical parasitology *164*, 137-146.

Wagner, W., Fodor, E., Ginsburg, A., and Hammer, J.A., 3rd (2006). The binding of DYNLL2 to myosin Va requires alternatively spliced exon B and stabilizes a portion of the myosin's coiled-coil domain. Biochemistry *45*, 11564-11577.

Walgraffe, D., Devaux, S., Lecordier, L., Dierick, J.-F., Dieu, M., Van den Abbeele, J., Pays, E., and Vanhamme, L. (2005). Characterization of subunits of the RNA polymerase I complex in Trypanosoma brucei. Molecular and biochemical parasitology *139*, 249-260.

Wang, L., Hare, M., Hays, T.S., and Barbar, E. (2004). Dynein light chain LC8 promotes assembly of the coiled-coil domain of swallow protein. Biochemistry *43*, 4611-4620.

Wang, Q.P., Kawahara, T., and Horn, D. (2010). Histone deacetylases play distinct roles in telomeric VSG expression site silencing in African trypanosomes. Mol.Microbiol. *77*, 1237-1245.

Wang, S.C., Lin, K.M., Chien, S.J., Huang, L.T., Hsu, C.N., and Tain, Y.L. (2014). RNA silencing targeting PIN (protein inhibitor of neuronal nitric oxide synthase) attenuates the development of hypertension in young spontaneously hypertensive rats. Journal of the American Society of Hypertension : JASH *8*, 5-13.

Wang, W., Lo, K.W., Kan, H.M., Fan, J.S., and Zhang, M. (2003). Structure of the monomeric 8-kDa dynein light chain and mechanism of the domain-swapped dimer assembly. The Journal of biological chemistry *278*, 41491-41499.

Wang, Z., Morris, J.C., Drew, M.E., and Englund, P.T. (2000). Inhibition of Trypanosoma brucei gene expression by RNA interference using an integratable vector with opposing T7 promoters. The Journal of biological chemistry *275*, 40174-40179.

Wheeler, R.J. (2010). The trypanolytic factor-mechanism, impacts and applications. Trends in parasitology *26*, 457-464.

White, R.J. (2008). RNA polymerases I and III, non-coding RNAs and cancer. Trends in Genetics *24*, 622-629.

Wickstead, B., Ersfeld, K., and Gull, K. (2002). Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of Trypanosoma brucei. Molecular and biochemical parasitology *125*, 211-216.

Wickstead, B., and Gull, K. (2007). Dyneins across eukaryotes: a comparative genomic analysis. Traffic *8*, 1708-1721.

Williams, J.C., Roulhac, P.L., Roy, A.G., Vallee, R.B., Fitzgerald, M.C., and Hendrickson, W.A. (2007). Structural and thermodynamic characterization of a cytoplasmic dynein light chain-intermediate chain complex. Proceedings of the National Academy of Sciences of the United States of America *104*, 10028-10033.

Wilson, M.J., Salata, M.W., Susalka, S.J., and Pfister, K.K. (2001). Light chains of mammalian cytoplasmic dynein: identification and characterization of a family of LC8 light chains. Cell motility and the cytoskeleton *49*, 229-240.

Wirtz, E., and Clayton, C. (1995). Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. Science *268*, 1179-1183.

Wirtz, E., Hartmann, C., and Clayton, C. (1994). Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes. Nucleic acids research *22*, 3887-3894.

Wirtz, E., Leal, S., Ochatt, C., and 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*, 89-101.

Wu, H., and King, S.M. (2003). Backbone dynamics of dynein light chains. Cell motility and the cytoskeleton *54*, 267-273.

Yang, X., Figueiredo, L.M., Espinal, A., Okubo, E., and Li, B. (2009). RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. Cell *137*, 99-109.

Zaros, C., and Thuriaux, P. (2004). Rpc25, a conserved RNA polymerase III subunit, is critical for transcription initiation. Molecular microbiology *55*, 104-114.

Zomerdijk, J.C., Kieft, R., and Borst, P. (1991a). Efficient production of functional mRNA mediated by RNA polymerase I in Trypanosoma brucei. Nature *353*, 772-775.

Zomerdijk, J.C., Kieft, R., Duyndam, M., Shiels, P.G., and Borst, P. (1991b). Antigenic variation in Trypanosoma brucei: a telomeric expression site for variant-specific surface glycoprotein genes with novel features. Nucleic acids research *19*, 1359-1368.

Zomerdijk, J.C., Ouellette, M., ten Asbroek, A.L., Kieft, R., Bommer, A.M., Clayton, C.E., and Borst, P. (1990). The promoter for a variant surface glycoprotein gene expression site in Trypanosoma brucei. EMBO J. *9*, 2791-2801.

Copyright Permissions

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jul 03, 2015

This is a License Agreement between Justin K Kirkham ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Justin K Kirkham
Customer address	713 Jacobs Ln NEWINGTON, CT 06111
License number	3661530312157
License date	Jul 03, 2015
Licensed content publisher	Elsevier
Licensed content publication	Molecular and Biochemical Parasitology
Licensed content title	Response to "Role of RPB7 in RNA pol I transcription in Trypanosoma brucei"
Licensed content author	Arthur Günzl,Sung Hee Park,Tu N. Nguyen,Justin K. Kirkham,Ju Huck Lee
Licensed content date	November 2011
Licensed content volume number	180
Licensed content issue number	1
Number of pages	2
Start Page	45
End Page	46
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes

Will you be translating?	No
Title of your thesis/dissertation	Investigating the role of LC8, CITFA, and RPB7 in RNA polymerase I-mediated transcription in Trypanosoma brucei
Expected completion date	Jan 2017
Estimated size (number of pages)	200
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jul 03, 2015

This is a License Agreement between Justin K Kirkham ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Justin K Kirkham
Customer address	713 Jacobs Ln NEWINGTON, CT 06111
License number	3661530158624
License date	Jul 03, 2015
Licensed content publisher	Elsevier
Licensed content publication	Gene
Licensed content title	Mono-allelic VSG expression by RNA polymerase I in Trypanosoma brucei: Expression site control from both ends?
Licensed content author	Astrid Engel,Justin K. Kirkham,Tu N. Nguyen,Nitika Badjatia,Sung Hee Park
Licensed content date	1 February 2015
Licensed content volume number	556
Licensed content issue number	1
Number of pages	6
Start Page	68
End Page	73
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes

Will you be translating?	No
Title of your thesis/dissertation	Investigating the role of LC8, CITFA, and RPB7 in RNA polymerase I-mediated transcription in Trypanosoma brucei
Expected completion date	Jan 2017
Estimated size (number of pages)	200
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jul 03, 2015

This is a License Agreement between Justin K Kirkham ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Justin K Kirkham
Customer address	713 Jacobs Ln NEWINGTON, CT 06111
License number	3661511278372
License date	Jul 03, 2015
Licensed content publisher	Elsevier
Licensed content publication	Molecular and Biochemical Parasitology
Licensed content title	Transcription by the multifunctional RNA polymerase I in Trypanosoma brucei functions independently of RPB7
Licensed content author	Sung Hee Park,Tu N. Nguyen,Justin K. Kirkham,Ju Huck Lee,Arthur Günzl
Licensed content date	November 2011
Licensed content volume number	180
Licensed content issue number	1
Number of pages	8
Start Page	35
End Page	42
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Title of your	Investigating the role of LC8, CITFA, and RPB7 in RNA polymerase

thesis/dissertation	I-mediated transcription in Trypanosoma brucei
Expected completion date	Jan 2017
Estimated size (number of pages)	200
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this

licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
  - via their non-commercial person homepage or blog
  - by updating a preprint in arXiv or RePEc with the accepted manuscript
  - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
  - directly by providing copies to their students or to research collaborators for their personal use
  - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- after the embargo period
  - via non-commercial hosting platforms such as their institutional repository
  - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs

and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant

DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.7

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.