Universidade de Lisboa

Faculdade de Medicina de Lisboa

UNDERSTANDING CHROMATIN DYNAMICS DURING ANTIGENIC VARIATION IN *TRYPANOSOMA BRUCEI*

Francisco Maria dos Santos e Silva Aresta Branco

Orientadora: Prof. Doutora Luísa Miranda Figueiredo

Tese especialmente elaborada para a obtenção do grau de Doutor em Ciências Biomédicas, Especialidade de Microbiologia e Parasitologia

2016

Universidade de Lisboa

Faculdade de Medicina de Lisboa

UNDERSTANDING CHROMATIN DYNAMICS DURING ANTIGENIC VARIATION IN *TRYPANOSOMA BRUCEI*

Francisco Maria dos Santos e Silva Aresta Branco

Orientadora: Doutora Luísa Miranda Figueiredo

Tese especialmente elaborada para a obtenção do grau de Doutor em Ciências Biomédicas, Especialidade de Microbiologia e Parasitologia

Júri:

Presidente: Doutor José Augusto Gamito Melo Cristino, Professor Catedrático e Presidente do Conselho Científico da Faculdade de Medicina da Universidade de Lisboa.

Vogais:

- Professor Markus Engstler, *Full Professor* of Cell and Developmental Biology, University of Wuerzburg, Germany;
- Doutor Lars Erwin Theodoor Domingos Jansen, Investigador Principal do Instituto Gulbenkian de Ciência;
- Doutora Elsa Margarida Teixeira Rodrigues, Professora Auxiliar da Faculdade de Farmácia da Universidade de Lisboa;
- Doutor Rui Miguel Prudêncio Cunha Pignatelli, Especialista de Reconhecido Mérito e Competência, Investigador do Instituto de Medicina Molecular da Faculdade de Medicina da Universidade de Lisboa;
- Doutora Vanessa Alexandre Zuzarte Luís, Especialista de Reconhecido Mérito e Competência, Investigadora do Instituto de Medicina Molecular da Faculdade de Medicina da Universidade de Lisboa;
- Doutora Luísa Miranda Figueiredo, Especialista de Reconhecido Mérito e Competência, Investigador do Instituto de Medicina Molecular da Faculdade de Medicina da Universidade de Lisboa, (Orientadora).

Fundação para a Ciência e Tecnologia (FCT) – SFRH/BD/80718/2011

2016

SUMMARY

African trypanosomiasis is a disease restricted to sub-Saharian Africa and it assumes two forms: Human African trypanosomiasis (HAT or sleeping sickness) or Animal African trypanosomiasis (AAT or n'gana), both being fatal if untreated. While HAT causes an estimated number of 15000 new cases each year and 70 million people are at risk, AAT causes an enormous economic burden due to livestock elimination [1].

HAT is caused by the unicellular *Trypanosoma brucei* parasite, which is transmitted by the tsetse transmitting vector. Upon a tsetse blood meal, parasites are delivered into a mammalian host where they proliferate in the bloodstream, lymphatic system and inside preferred tissues, such as the adipose tissue [2]. The parasites can progressively migrate to the central nervous system, leading to neurological disorders, coma and death [3].

Each *Trypanosoma brucei* parasite expresses a single type of Variant Surface Glycoprotein (VSG) at its cell surface. To avoid being eliminated by the host immune system, parasites use antigenic variation, which consists in periodically changing to a new VSG coat, obligating the immune system to reinitiate a new adaptive response and generate new antibodies. VSG genes are transcribed by RNA polymerase I (Pol I) from specialized subtelomeric loci termed Bloodstream Expression Sites (BESs) [4]. Although there are around 15 BESs in the genome, only one is active at any given time, thus ensuring VSG monoallelic expression. Replacement of the VSG coat starts in the nucleus by changing the actively transcribed VSG. This process takes place through: i) switching by recombination, which involves DNA recombination between the active and a silent BES encompassing the VSG gene or by copying a VSG sequence from archival copies present in the genome into the active BES [5, 6]; ii) *in situ* switching, where the active BES becomes silenced with the concomitant activation of a silent BES [7, 8].

VSG *in situ* switching is very poorly understood. We do not know the key players or the sequence of events that regulate it. As the actively transcribed BES possesses a very open chromatin conformation relative to transcriptionally silent BESs [9, 10], we proposed to understand how chromatin and transcription interweave during an *in situ* switching.

Another event in which the active BES becomes silenced is during differentiation from bloodstream to procyclic (present in the tsetse) forms [11]. In this situation, we observed that transcriptional silencing precedes chromatin changes. Thus, we postulated that during an *in situ* switching event transcription is probably also halted prior to closing chromatin structure. We chose to use a tetracycline-inducible BES transcriptional silencing system from the Horn lab that induces VSG silencing [12], resulting in an increase of the switching frequency to 8%. Blockage of transcription was confirmed by the absence of Pol I in the active BES and led to a rapid decrease in transcript levels of this locus within the first 8 hr after BES silencing. Despite such significant changes at the transcriptional level, chromatin of the previously active BES remained in an open conformation. We hypothesized that the cell-cycle was necessary to induce chromatin remodeling as in Pol I-transcribed ribosomal DNA (rDNA) genes in yeast, but this revealed unfruitful.

We observed that TDP1 (Trypanosome DNA-binding protein 1), an essential high mobility group box protein [13, 14], remained in the active BES even after BES silencing had been triggered. Depletion of TDP1 in these conditions led to closure of chromatin indicating that TDP1 is a key factor in stabilizing open chromatin conformation under transcriptional switching. We showed that during *in situ* switching, cells probe more than one silent BES, inclusively expressing other VSGs. We determined that after 24 hr of BES silencing, roughly half of the cells were committed to switch while the remaining could revert to transcribe the initial BES. Interestingly, cells that at 24 hr of BES silencing were not probing a specific BES (and VSG) did not switch to that BES. Overall, these results led us to propose a model in which under an *in situ* switching, transcription of the active BES is halted but its chromatin is maintained in an open conformation by TDP1. This allows parasites to probe silent BES to make a decision of switching or returning to the initial BES.

The importance of TDP1 in Pol I transcription and in the switching process led us to ask if TDP1 overexpression could interfere with VSG monoallelic expression. Using a TDP1-overexpressing cell-line, we observed that genome-wide chromatin becomes more open (at different extent) except in the active BES and rDNA genes, suggesting that these two loci already have the chromatin fully open. Furthermore, the mRNA transcript levels

only increased in silent BES, MES (Metacyclic Expression Sites) and procyclin loci confirming the role of TDP1 as a Pol I transcription facilitator [14]. Importantly, we observed the expression of a silent VSG upon TDP1 overexpression confirming the disruption of monoallelic expression.

Overall, this dissertation enlightens the relevance of chromatin during an *in situ* switching and elucidates the roles of TDP1 as key factor for antigenic variation.

Keywords: *Trypanosoma brucei*; monoallelic expression; VSG *in situ* switching; chromatin; TDP1

RESUMO

A tripanossomíase Africana é uma doença restrita à África subsariana e que assume duas formas: tripanossomíase Humana Africana (THA ou doença do sono) ou tripanossomíase Animal Africana (TAA ou *n'gana*), sendo ambas fatais quando não tratadas. Enquanto que a THA causa, anualmente, um número estimado de 15000 novos casos, encontrando-se 70 milhões de pessoas em risco de contrair a doença, a AAT provoca um enorme fardo económico devido à eliminação de gado [1].

A THA é causada pelo parasita *Trypanosoma brucei* que é transmitido pelo vetor de transmissão denominado tsé-tsé, uma mosca do género *Glossina*. Durante a refeição sanguínea da tsé-tsé, os parasitas são injetados no hospedeiro mamífero onde proliferam na corrente sanguínea, sistema linfático e no interior de tecidos preferenciais, tal como o tecido adiposo [2]. Os parasitas podem progressivamente migrar para o sistema nervoso central, causando problemas neurológicos, coma e morte [3].

Cada parasita expressa cerca de 10 milhões de cópias de uma proteína denominada *Variant Surface Glycoprotein* (VSG) à superfície da célula. Apesar do genoma do parasita conter cerca de 2000 genes que codificam para VSGs, apenas um gene, e consequentemente um tipo de proteína, é expresso a cada momento. Para evitar a eliminação pelo sistema imune do hospedeiro, os parasitas usam o mecanismo de variação antigénica, que consiste na mudança periódica para uma nova cobertura de VSG, obrigando o sistema imune a re-iniciar uma nova resposta adaptativa e a gerar novos anticorpos. As VSGs são transcritas pela RNA polimerase I (Pol I) a partir de locais subteloméricos especializados denominados *Bloodstream Expression Sites* (BESs) [4]. Apesar de existirem cerca de 15 BESs no genoma, apenas um se encontra ativo a cada momento, estando os restantes silenciados, assegurando a expressão monoalélica de VSGs. A substituição da cobertura de VSGs começa no núcleo através da mudança da VSG ativamente transcrita. Este processo pode ocorrer através de: i) recombinação de DNA, no qual ocorre troca de sequência genómica de VSG entre o BES ativo e um BES silenciado ou através da cópia e inserção no BES ativo de uma sequência genómica de VSG que se encontre em regiões genómicas destinadas ao arquivo de genes VSG [5, 6]; ii) mudança *in situ*, onde o BES ativo é silenciado com a concomitante ativação de um BES silenciado [7, 8].

A mudança *in situ* de VSG é pouco compreendida. Não se conhecem os fatores chave, nem sequer a sequência de eventos que estão envolvidos nesta mudança. O BES ativo é altamente transcrito e possui uma conformação de cromatina aberta, composta de muito poucos nucleossomas. Os BESs silenciados são pouco transcritos e apenas na zona do promotor. A sua cromatina é fechada, compacta e repleta de nucleossomas [9, 10]. Dada a correlação entre taxa de transcrição e conformação de cromatina, neste trabalho propomo-nos a compreender como é que a cromatina e transcrição interagem e são modificadas durante uma mudança *in situ* de VSG.

Um outro momento em que o BES ativo é silenciado ocorre durante a diferenciação do parasita da forma sanguínea (no mamífero) para a forma procíclica (na tsé-tsé). Neste processo, todos os VSGs são silenciados, para dar lugar a prociclínas nas superfície da célula. Nesta situação, observámos que o silenciamento da transcrição precede alterações na estrutura da cromatina. Desta forma, postulámos que durante uma mudança *in situ*, a transcrição é, também, provavelmente suspensa antes da estrutura da cromatina ser fechada. Para podermos estudar a relação transcrição/cromatina durante uma mudança de VSG *in situ,* modificámos geneticamente uma linha repórter inicialmente desenvolvida no laboratório Horn [12]. Nesta linha celular, a transcrição do BES ativo é bloqueada através de um sistema indutível de tetraciclína, sendo que, para sobreviverem, os parasitas necessitam de mudar a transcrição para outro BES e expressar outro VSG à superfície da célula. Começámos por verificar que todas as células alteram o VSG que expressam mas apenas 8% das células sobrevivem, sendo esta frequência 100 a 10000 vezes superior ao previamente observado. O bloqueio da transcrição foi confirmado pela ausência de Pol I no BES ativo e pelo rápido decréscimo nos níveis de transcritos desta região nas primeiras 8 horas após o silenciamento do BES. Apesar de alterações significativas a nível transcricional, a cromatina do BES previamente ativo manteve-se numa conformação aberta. De seguida, testámos se o ciclo celular era importante para induzir uma reestruturação da cromatina, tal como o DNA ribossomal (rDNA) transcrito pela Pol I em leveduras, mas esta hipótese revelou-se infrutífera.

Antes do início desta dissertação, já tinha sido demonstrado que a TDP1 (*Trypanosome DNA-binding protein 1*) é uma proteína de grupo com grande mobilidade, essencial ao parasita e que está presente em regiões genómicas transcritas pela Pol I [13,

14]. Primeiramente, mostrámos que a TDP1 se mantém no BES ativo mesmo após o silenciamento do BES ter sido despoletado. Ao induzirmos o silenciamento do BES ativo após a depleção da TDP1, verificámos que a cromatina do BES ativo ficou mais compactada indicando que a TDP1 é um factor chave na estabilização da conformação de cromatina aberta aquando de alterações do *status* transcricional do BES ativo. Mostrámos ainda que durante uma mudança *in situ*, os parasitas sondam mais que um BES silenciado ao mesmo tempo, inclusivamente expressando outros VSGs. Determinámos que, após 24 horas de silenciamento do BES, aproximadamente metade das células já estavam molecularmente comprometidas para mudar a transcrição para outro BES, enquanto que as restantes células podiam reverter e re-transcrever o BES inicial. Curiosamente, as células que, às 24 horas após o silenciamento do BES ativo, não estivessem especificamente a sondar um determinado BES (e o seu VSG) não mudavam a transcrição para esse BES e consequentemente para a expressão do VSG desse BES. Globalmente, estes resultados levaram-nos a propôr um modelo em que durante a indução de uma mudança *in situ*, a transcrição do BES ativo é suspensa mas a cromatina é mantida numa conformação aberta pela TDP1. Isto permite aos parasitas sondar BES silenciados e avaliar a funcionalidade das proteínas expressas nesses BES para decidir se muda de BES e de VSG expresso à superfície da célula ou se voltam a transcrever o BES inicial mantendo a expressão do mesmo VSG. Considerando esta linha celular, a re-transcrição do BES inicial que está bloqueado com o sistema indutível de tetraciclina leva à morte dos parasitas.

A importância da TDP1 na transcrição da Pol I e no processo de mudança de VSG levou-nos a perguntar se o ganho de função da TDP1, através do aumento do conteúdo celular desta proteína, poderia interferir com a expressão monoalélica. Utilizando uma linha celular que permite sobre-expressar a TDP1, observámos que a cromatina do genoma fica globalmente mais aberta (a diferentes níveis consoante as regiões genómicas) exceptuando o BES ativo e os genes ribossomais, sugerindo que estas duas regiões possuem a cromatina completamente aberta. Adicionalmente, os nivéis de transcritos de RNA mensageiro apenas aumentaram para os BES silenciados, MES (*Metacyclic Expression Sites*, que contém genes de VSGs que apenas são expressos quando os parasitas se encontram nas glândulas salivares da tsé-tse prontos a serem injectados no hospedeiro mamífero) e regiões de prociclínas confirmando o papel da

TDP1 como um facilitador de transcrição de Pol I [14]. Também observámos a expressão de um VSG silenciado para além do VSG ativo após a sobre-expressão da TDP1, confirmando a disrupção da expressão monoalélica.

Globalmente, esta dissertação esclarece a relevância da cromatina durante uma mudança *in situ* e elucida os papéis da TDP1 como um fator chave para a variação antigénica.

Palavras-chave: *Trypanosoma brucei*; expressão monoalélica; mudança *in situ* de VSG; cromatina; TDP1.

ACKNOWLEDGMENTS

First of all, I cannot be more grateful to my supervisor, Luísa Figueiredo. Thank you for giving me the opportunity six years ago to join and spend this amazing time in your lab. All the support and advice you gave me, all the times I had personal problems and you worried and manage to have time and sit to calm me, all the times your listened to my ideas and gave me the freedom to pursue them, all the times you told me to focus and pressured me to correct my weaknesses… Thank you for bringing me to the awesome parasitology world and inspiring me to continue an academic career. Truly truly thank you! And I will try not call you again at 4 a.m. =D

To my lab mates, past and present, thank you for all these years of spectacular scientific and social environment, you really are the best. To Margarida, you really were a huge support and such an amazing friend so many times. To Sandra, you were always available to help me and calming me when I needed. To Idálio, I really could not have a better student than you and I learned so much with you too. To Filipa and Ana, who received me so well in the lab and whose help when I arrived to the lab was extremely important. To Daniel, Mafalda, Fábio, Pena, Leonor, Fabien, Nina, Miguel Ferreira, Eleonora and Helena for the friendship and support you gave me. To João Rodrigues for the input and help you've been giving me. To everybody from Miguel Prudêncio's and Maria Mota's labs for all the input in the parasitology sessions and the good times.

To my thesis committee members Maria Mota, João Barata and Diogo Castro for all the great input you provided in our meeting. To Sérgio Almeida for the several conversations and very good advice you always gave me.

To Marco Antunes and Bruno Vidal, I cannot express how your friendship was and is so much important through all these years! You are truly amazing friends. In the innumerous good times you were there but you never failed me in the bad times. A very special thank you to both of you. To David, Tiago, Paulo and Nuno, your 16 year-long great friendship is part of me and I'll always remember how you manage to make me forget the problems and realize that we need to have fun in life. To Damas, Marquito, Marta and Henri, my daily Biochemistry friends with whom I shared so many college projects making me love science and are part of the reason why I decided to pursuit a

PhD. To Miguel Dinis, your great friendship and support in the last months have been extremely important for me to finish this thesis. To Júlia, an amazing person and great friend who is very important to me and always gave me indispensable advice.

To Carolina for being by my side even now that you are far away, for all your love, for making me feel happy and special, for handling my bad humor, and for giving me such a strong force and confidence boost in these final months. You are really special to me!

And finally, but most importantly, THANK YOU to my family. You are the foundation of the person I am. Thank you for all the unconditional support, for all the values you taught me, all the strength that you always gave me and for making me pursuit all my dreams. I dedicate this dissertation to you!

TABLE OF CONTENTS

FIGURES AND TABLES INDEX

CHAPTER 1 – GENERAL INTRODUCTION

CHAPTER 2 – A TRANSCRIPTION-INDEPENDENT EPIGENETIC MECHANISM IS ASSOCIATED WITH ANTIGENIC SWITCHING IN *TRYPANOSOMA BRUCEI*

CHAPTER 3 – OVEREXPRESSION OF A HIGH MOBILITY GROUP BOX PROTEIN COMPROMISES VSG MONOALLELIC EXPRESSION IN *TRYPANOSOMA BRUCEI*

ABBREVIATIONS

ºC	Celsius degree
3'UTR	3' untranslated region
5'UTR	5' untranslated region
AAT	Animal African trypanosomiasis
ala	Alanine
APOL1	Apolipoprotein L1
AT	African trypanosomiasis
ATF	Adipose tissue form
ATP	Adenosine triphosphate
Base J	β-D-glucopyranosyloxymethyluracil
BARP	brucei alanine-rich protein
BDF	Bromodomain factor
BES	Bloodstream expression site
BIR	Break-induced replication
BLE	Bleomycin resistance
bp	base pair
BSF	Bloodstream form
BSR	Blasticidin-S resistance
cAMP	Cyclic adenosine monophosphate
CSB	Cockayne Syndrome protein B
CDK	Cyclin-dependent kinase
cDNA	complementary DNA
ChIP	Chromatin immunoprecipitation
CITFA	Class I transcription factor A
CSF	Cerebrospinal fluid
C-terminal	Carboxy-terminal
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DSB	Double strand break
dsRNA	Double-stranded RNA
DTM	Differentiating trypanosome medium
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethyleneglycoltetraacetic acid
EMP1	Erythrocyte membrane protein 1
eNoSC	Energy-dependent nucleolar silencing complex
ESAG	Expression site-associated gene
ESB	Expression site body
FACT	Facilitator of chromatin transactions
FAIRE	Formaldehyde-assisted isolation of regulatory elements

G1	Gap 1
G2	Gap 2
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
gDNA	Genomic DNA
GFP	Green fluorescent protein
GLB1	GFP, luciferase, blasticidin-S resistance in BES1
GLB1-R3	GLB1 with RFP::NPT in BES3
GLB1-TDP1::TY1	GLB1 with a TY1 tag fused to endogenous TDP1 allele
GLB1-TDP1::3xcMyc	GLB1 with a triple cMyc tag fused to endogenous TDP1 allele
Glu	Glutamic acid
Gly	Glycine
GPI	Glycosylphosphatidylinositol
GRESAG2	Gene related to ESAG2
HA	Hemagglutinin
HAT (disease)	Human African trypanosomiasis
HAT (protein)	Histone acetyltransferase
HCl	Hydrogen chloride
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMGB	High mobility group box
HP1	Heterochromatin protein 1
HR	Homologous recombination
hr	Hours
kbp	Kilobase pair
kDa	Kilodalton
KCl	Potassium chloride
KOH	Potassium hydroxide
KD	Knock-down
ISWI	Imitation of switch
LiCl	Lithium chloride
LSF	Long slender form
M (cell-cycle)	Mitosis
M (concentration)	Molar
Mbp	Megabase pair
MES	Metacyclic expression site
MgSO4	Magnesium sulfate
μg	Microgram
μl	Microliter
μm	Micrometer
ml	Mililiter
mRNA	Messenger ribonucleic acid
MSP-B	Major surface protease B
MTOC	Microtubule organizing center
mVSG	Metacyclic VSG

NA	Neuraminidase
NaCl	Sodium chloride
Na2HPO4	Sodium hydrogen phosphate
NaH2PO4	Sodium dihydrogen phosphate
Nek	NimA-related kinase
NLP	Nucleoplasmin-like protein
NLS	Nuclear localization signal
NOR	Nucleolus organizer region
NoRC	Nucleolar remodeling complex
NP-40	Nonyl phenoxypolyethoxylethanol
NPT	Neomycin phosphotransferase
NR-elements	NlaIII repeat elements
N-terminal	Amino-terminal
NuRD	Nucleosome remodeling and deacetylation
OR	Olfactory receptor
ORC	Origin recognition complex
ORF	Open reading frame
OSN	Olfactory sensory neuron
PAD1	Protein associated with differentiation 1
PAG	Procyclin associated gene
PARP1	Poly(ADP-ribose)-polymerase 1
PBS	Phosphate-buffered saline
pCPTcAMP	8-(4-chlorophenylthio)-cAMP
PCR	Polymerase chain reaction
PF	Procyclic form
P. falciparum	*Plasmodium falciparum*
PfEMP1	*P. falciparum* erythrocyte membrane protein 1
PI	Propidium iodide
PIC	Pre-initiation complex
PL1A	puromycin N-acetyl-transferase, luciferase in BES1 active
PL1S	puromycin N-acetyl-transferase, luciferase in BES1 silent
PIP5K	1-phosphatidylinositol-4-phosphate-5 kinase
PIP5Pase	1-phosphatidylinositol-4-phosphate-5 phosphatase
PMSF	phenylmethylsulfonyl fluoride
Pol I	RNA polymerase I
Pol II	RNA polymerase II
Pol III	RNA polymerase III
Poly(Y)	Polypyrimidine
POT	PL1S overexpressing TDP1
pRNA	Promoter-associated RNA
Pro	Proline
PTU	Polycistronic transcription unit
qPCR	quantitative PCR

rDNA	Ribosomal DNA
RENT	Regulator of nucleolar silencing and telophase
RFP	Red fluorescent protein
RFP::NPT	RFP gene fused to a NPT gene
RNA	Ribonucleic acid
RNAi	RNA interference
RNase A	Ribonuclease A
RT	Reverse transcriptase
S	Synthesis
SDS	Sodium dodecyl sulfate
SIF	Stumpy induction factor
Sir2	Silencing information regulator 2
siRNA	Small interfering RNA
SL1	Selectivity factor 1
SRA	Serum resistance-associated gene
SS	Short stumpy form
SUMO	Small ubiquitin-like modifier
SWI/SNF	Switch / sucrose non-fermentable
TBP	TATA-binding protein
T. brucei	*Trypanosoma brucei*
T. brucei brucei	*Trypanosoma brucei brucei*
T. b. gambiense	*Trypanosoma brucei gambiense*
T. b. rhodesiense	*Trypanosoma brucei rhodesiense*
T. congolense	*Trypanosoma congolense*
T. cruzi	*Trypanosoma cruzi*
T. equiperdum	*Trypanosoma equiperdum*
T. vivax	*Trypanosoma vivax*
TDB	Trypanosome dilution buffer
TERT	Telomerase reverse transcriptase
Tet	Tetracycline
TetO	Tetracycline operator
TetR	Tetracycline repressor
TDP1	Trypanosoma DNA-binding protein 1
Thr	Threonine
TIP5	TTF-I interacting protein 5
TOR	Target of rapamycin
TRF4	TATA-binding protein related factor 4
tRNA	Transfer RNA
TSS	Transcription start site
TTF-I	Transcription termination factor 1
TTS	Transcription termination site
UAF	Upstream activating factor
UBF	Upstream binding factor

UPR	Unfolded protein response
UTR	Untranslated region
VEX1	VSG exclusion 1
Vmp	Variant major protein
VSG	Variant surface glycoprotein
WHO	World Health Organization

CHAPTER 1

GENERAL INTRODUCTION

AFRICAN TRYPANOSOMIASIS

African trypanosomiasis (AT) is an endemic disease in the sub-Saharan Africa [15]. This disease is caused by a protozoan parasite from the *Trypanosoma* genus, which is transmitted by the tsetse, a fly from the genus *Glossina*. Thus, the distribution of AT is correlated with the distribution of the tsetse in Africa. Depending on the mammalian host, there are two different forms of the disease: the Animal and the Human African trypanosomiasis (AAT and HAT, respectively).

AAT, also known as n'gana, is specifically caused by *T. congolense*, *T.vivax* and *T. b. brucei*. AAT affects both wild and domestic animals although in the former the infection symptoms are mild whereas in the latter, the infection often causes death [16]. As a consequence, AAT causes a severe economic burden. According to the Food and Agriculture Organization of the United Nations, losses are estimated to be in the range of US$ 1-1.2 billion in cattle alone.

HAT, or sleeping sickness, is caused by two subspecies of *T. brucei*, *T. b. gambiense* and *T. b. rhodesiense* [3, 15]. Rarely, some human infections were reported to be caused by animal infective trypanosome species [17]. *T. b. gambiense* infections occur in the Central and Western Africa, giving rise to a chronic form of HAT. It accounts for ~98% of all HAT reported cases leading to death in 3 to 4 years if untreated. On the other hand, *T. b. rhodesiense* infections occur in the Eastern Africa, giving rise to an acute and rapidly progressing form of HAT. Thus, it accounts for ~2% of HAT cases causing death in 6 months if left untreated. In 2014, according to the WHO (World Health Organization), the number of HAT new reported cases was 3796, the estimated number of cases below 15000 but the estimated population at risk is still around 70 million people [1].

HAT progresses in two stages: an early stage or haemolymphatic phase, characterized by the presence of the parasites in the bloodstream and lymphatic system and a late stage or neurological phase where the parasites are able to cross the blood-brain barrier invading the central nervous system. Both forms of HAT present similar signs and symptoms of the disease, although with different severity and frequency, being the *rhodesiense* form of the disease more acute and with shorter time-lapse of the early stage [3, 16]. For the early stage, the symptoms account as headache, physical weakness,

intermittent fever, dermatologic problems, anemia, musculoskeletal pains, lymphadenopathies, cardiac disorders, endocrine disturbances and hepatosplenomegaly. The late stage infected patients present a disturbed sleep-wake cycle and neuropsychiatric symptoms such as confusion, sensory disturbances, extreme lethargy, poor condition and coma. Importantly, the 24 hour circadian rhythm is disrupted and alterations in the circadian fluctuations of cortisol, prolactin, and growth hormone secretion, as well as of plasma renin activity were already detected in patients [18, 19].

Nowadays the diagnosis of HAT is performed according to a decision tree [20]. Initially, a serological test for antibody detection is performed proceeding afterwards to lymph node examination and further analytical techniques. If the exams turn to be positive, a stage diagnosis is performed by a cerebrospinal fluid (CSF) lumbar punction to further decide the treatment. In the absence of trypanosomes in the CSF, it is considered that the patient is in the early stage of the disease and pentamidine (for *T. b. gambiense*) and suramin (for *T. b. rhodesiense*) are the drugs given to the patients. Although with slight undesirable effects, these drugs are well-tolerated by the patients. On the contrary, if trypanosomes are detected in the CSF, it means that the patients are already in the late stage of HAT. In this case, the drugs administered to the patients are: melarsoprol, which is used for both *T. b. gambiense* and *rhodesiense* infections but elicits several side-effects that can be fatal; eflornithine, which is less toxic but only effective against *T. b. gambiense*; and a combination eflornithine/nifurtimox in which nifurtimox increases the efficacy of eflornithine as a first line defense for *T. b. gambiense* but not tested for *T. b. rhodesiense* yet.

TRYPANOSOMA BRUCEI BIOLOGY ASPECTS

Trypanosoma brucei is a unicellular eukaryote that taxonomically is placed in the supergroup Excavata, phylum Euglenozoa, class Kinetoplastida. Kinetoplastids diverged very early in the eukaryotic lineage of evolution [21], such that Euglenozoa has been considered to belong to the root of the eukaryotic tree of evolution [22]. Although trypanosomes branched from higher eukaryotes possessing unique features, *T. brucei* is often used as a model organism since it is easy to cultivate *in vitro* and its genome is sequenced [23], making it easy to manipulate. It is also striking how this model organism became a platform for the discovery of several important aspects and pathways that are common within eukaryotic cells. Amongst them, the presence of glycosylphosphatidylinositol (GPI) anchors [24], a post-translational modification in the C-terminus of proteins that allows attachment of proteins on the extracellular moiety of the plasma membrane; RNA editing [25], a process by which an RNA transcript can be subjected to sequence nucleotide changes, either by insertion, deletion or replacement; *trans*-splicing [26-28], a process by which two or more different RNA transcripts are joined or spliced into a single messenger RNA (mRNA). Moreover, and specifically concerning kinetoplastids, it was on *T. brucei* that the DNA base J (β-D-glucopyranosyloxymethyluracil) [29, 30], and the glycolytic specialized organelle glycosome [31, 32] were first found.

Life cycle of *T. brucei*

The life cycle of *T. brucei* oscillates between two hosts: the mammal and the transmitting vector tsetse (Figure 1). When the tsetse bites the mammalian host, it uptakes blood as a meal but also releases parasites in the growth-arrested metacyclic form [33] through its salivary glands. These parasites, when in the bloodstream of the host, re-acquire cell division through binary fission and adopt a morphological long slender form (LSF), being commonly termed bloodstream forms (BSFs). BSFs are able to establish an infection by evading the host immune system through antigenic variation of the cell surface variant surface glycoproteins (VSG) (see section 1.5). The dense VSG coat also serves as a physical protection since it avoids lysis from the alternative pathway of

complement activation [34]. The changes between metacyclics and BSFs are also accompanied by a change of the VSG coat from the metacyclic subset (mVSGs) to the bloodstream subset of VSGs, which takes around 5-6 days [35]. Throughout the infection, trypanosomes also invade other body fluids (lymphatic system and spinal fluid) and tissues with a high focus in the adipose tissue [2] and dermis [36]. A recent study from our group also showed that parasites adapt their metabolism when living in the adipose tissue, including activation of the β-oxidation pathway in *T. brucei* [2]. Ultimately, the parasites progress to the central nervous system, causing the neurological disorders.

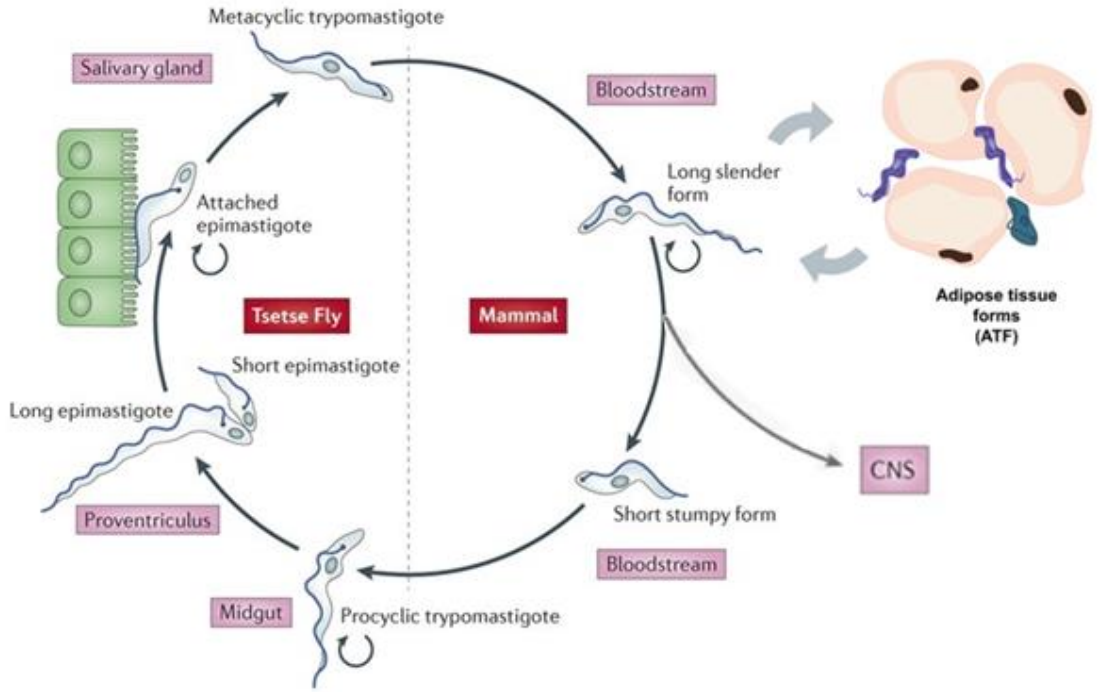


Figure 1 – Representation of the life cycle of *Trypanosoma brucei* and its different forms on both mammal and tsetse hosts. Injection of metacyclic forms during a tsetse meal is followed by proliferation of long slender or bloodstream forms (BSFs) in the bloodstream and lymphatic system and accumulate at high density in the adipose tissue (ATFs) later in infection. At high density, BSFs differentiate to short stumpy forms (SSs) which are uptaken by the tsetse. In the adipose tissue both BSFs (in purple) and SSs (in dark grey) co-exist. SSs differentiate in procyclic forms (PFs) in the tsetse midgut, which will migrate to the proventriculus and assymetrically divide in a long and a short epimastigote. The short epimastigote migrates and attaches to the epithelia of the salivary glands to finally differentiate into metacyclic forms which are competent to infect a new mammalian host upon a new tsetse meal (image adapted from [2, 37]).

Proliferation of the parasites in the bloodstream leads to the accumulation of a mysterious stumpy induction factor (SIF) that drives the differentiation of BSFs to another life cycle stage in the bloodstream, the short stumpy forms (SSs), at high parasiteamia levels [38, 39]. The response to SIF is density-dependent similar to quorum sensing mechanism and involves an extensive cascade of factors that ultimately control cell-cycle and gene expression [40]. SSs are mainly characterized by a small sized morphology, expression of PAD1 (Protein Associated with Differentiation 1, a carboxylate transporter) and a G_0/G_1 cell-cycle arrest [39, 41, 42], which prepares them to be uptaken by the tsetse during a blood meal. The transcriptome [43-45] and the proteome [46] of SS becomes altered in relation to BSFs, namely its metabolism is prepared for living in the tsetse midgut environment. Furthermore, these are the only forms that can survive after transmission back to the tsetse [47].

Upon entry in the tsetse midgut, the parasites start differentiating to procyclic forms (PFs) and start a process of development in the tsetse that lasts 20-30 days [3]. The drop of temperature and the sensing of citrate in the fly midgut are the cues for this differentiation [48], which is initiated by a Nek-related kinase defining a commitment step [49]. One study performed a transcript analysis throughout the process of differentiation from pure SSs to PFs [44], whereas another study [50] conducted an experiment starting differentiation with pleomorphic BSFs. Some similarities were obtained but the latter study suffered a 12-hour delay as a transition to “stumpy-like” forms before committing for procyclics while the former study started with pure SS. Other studies observed the steady-state differences of the transcriptome [43, 45] and proteome [46] in SSs and PFs, but the majority of the works has focused on observing changes between BSF and PF (for transcriptome [43-45, 50-53], for proteome [46, 54, 55]). Another major change that takes place during differentiation to procyclics is the replacement of the cell surface coat from VSG to procyclins, which is a consequence of adaptation to a new host. As such, the C-terminal domain of highly acidic procyclins allows protection against digestive enzymes from the tsetse midgut [56]. In a single cell, the 10^7 VSG molecules are shed from the surface by MSP-B (Major Surface Protease B, a zinc metalloprotease) and GPI-PLC (Glycophosphatidylinositol Phospholipase C) [57, 58] in the first 18 hours [44]. At the same time, two classes of procyclins, EP (for Glu-Pro repeats) and GPEET (for Gly-Pro-Glu-

Glu-Thr repeats), become expressed within 6 hours [44, 59]. After 7 days, only EP procyclins are expressed at the cell surface of procyclics [60].

Upon this stage, procyclics migrate to the tsetse proventriculus, where they give rise to a long epimastigote and ultimately divide assimetrically into a long and a short epimastigote [61]. At this stage, epimastigotes express another surface protein termed BARP (for *brucei* alanine-rich protein) [62]. Unlike the long epimastigotes that eventually die, the short epimastigotes migrate to the salivary glands attaching to the epithelial cells. There, sexual exchange may undergo between epimastigotes [63, 64], before differentiating to metacyclic forms, where binary fission stops and acquisition of a new mVSG coat occurs [65]. In the last step, metacyclics detach from the epithelial cells and are ready to infect a new mammalian host.

Several transitions of *T. brucei* life cycle can be mimicked *in vitro*. Differentiation from BSFs to PFs is accomplishable by reducing the temperature from 37ºC to 27ºC and addition of citrate/*cis*-aconitate to the medium [48, 66]. However, performing this in monomorphic cell-lines bypasses the formation of stumpy forms probably due to stable maintenance in laboratory conditions that leads to a failure in the signaling pathway [67]. Nonetheless, addition of a cAMP derivative termed 8-(4-chlorophenylthio)-cAMP (pCPTcAMP) allows slender-to-"stumpy-like" differentiation in monomorphs [67]. The inability to naturally differentiate to stumpy forms is the reason why monomorphic strains are more virulent in mice than pleomorphic strains [68].

Until recently, only these two transitions could be performed *in vitro*. Kolev and collegues recapitulated the transitions procyclic-to-epimastigote-to-metacyclic by overexpressing the RNA-binding protein RBP6 in PFs [69], where individual metacyclics express a single mVSG at the cell surface [70].

Genome organization of *T. brucei*

The nucleus of *T. brucei* has a diameter size of ~2.5 μm [71] and possesses a very unusual set of chromosomes divided into three different classes: 11 diploid megabase chromosomes that vary in size (0.9 to 5.7 Mbp) 1-5 intermediate chromosomes (200 to

900 kbp) and ~100 minichromosomes (30 to 150 kbp) [23, 72, 73]. Field isolates, but not laboratory strains, contain extrachromosomal DNA sequences termed NlaIII repeat (NR)-elements, which may derive from subtelomeric recombinational events, that can last for several generations [74].

The megabase chromosomes encompass the housekeeping genes, including all RNA polymerase II (Pol II) and RNA polymerase III (Pol III) transcribed genes, which are organized in several polycistronic transcription units (PTU) [75], similar to operons in bacteria. The subtelomeric regions of the megabase chromosomes are also the home of the Bloodstream and Metacyclic Expression Sites (BES and MES, respectively) (these loci will be discussed ahead). Intermediate chromosomes are composed of non-repetitive DNA, repetitive sequences of 177-bp satellite sequences and functional BES together with telomeric repeats [76], and it is possible that these chromosomes have derived from subtelomeric breakage of megabase chromosomes [77]. As for minichromosomes, they comprise around 10% of nuclear DNA [73], and are composed of long 177-bp satellite sequences which have an inversion point in the center of the minichromosome, plus 70-bp repeats followed by VSG genes (the only genes in minichromosomes) and the telomeric repeats [78, 79].

Trypanosoma brucei also possess another type of DNA which is not present in the nucleus: the kinetoplast. The kinetoplast is composed of a huge network mesh of maxicircles (analogous to mitochondrial DNA) and minicircles [80]. There are between 50-100 maxicircles with a size of around 23 kb and 5.000-10.000 minicircles of around 1 kb in a network [81, 82]. The maxicircles encode for 9S and 12S mitochondrial ribosomal RNAs and mitochondrial membrane proteins [83-85], which are subjected to RNA editing, while the minicircles give rise to guide RNAs for RNA editing [86, 87].

The special organization of VSG genes and Expression Sites

The genome of the best sequenced *T. brucei* strain (TREU927) contains around 2000 VSG genes, distributed amongst VSG arrays localized in the internal parts of megabase chromosomes, in BESs and MESs of megabase and intermediate (in this case,

only BES are present) chromosomes and as single genes in minichromosomes (Figure 2) [23, 88, 89].

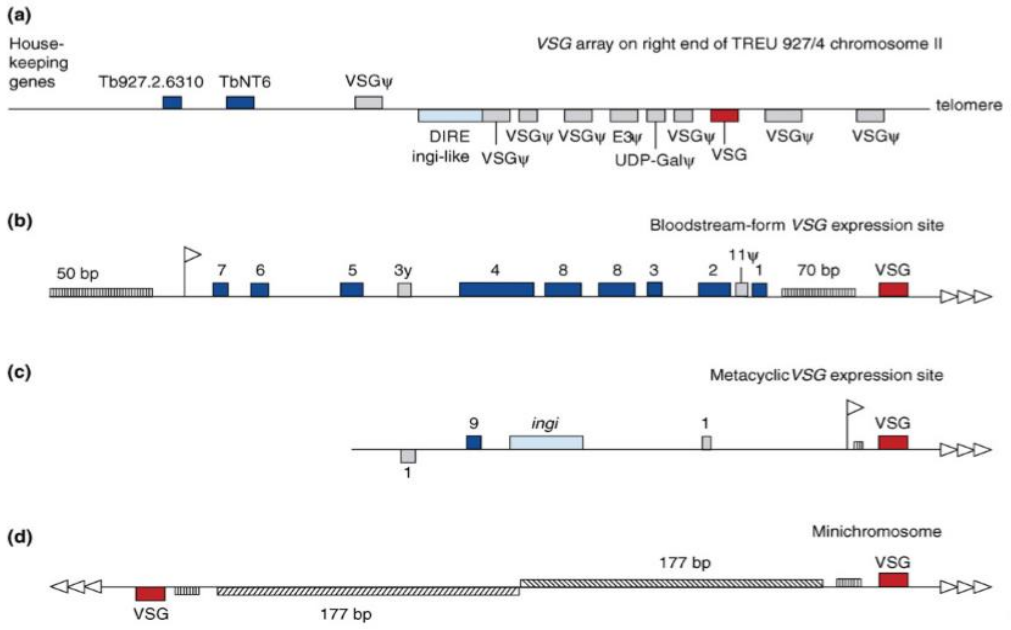


Figure 2 - Organization of VSG genes in the genome of *Trypanosoma brucei*. (a) Example of an internal array nearby housekeeping genes (in blue) containing functional VSG genes (in red) and pseudogenes (in grey) and an *ingi*-like retrotransposon element (in light blue). **(b)** Typical representation of a bloodstream expression site (BES). Upstream of the promoter (white flag) are located the 50 bp-repeats, while downstream of the promoter are located the expression-site associated genes (ESAGs, in blue), possible pseudo-ESAGs (in grey), 70 bp-repeats and the VSG gene (in red) before the telomeric repeats. **(c)** Typical representation of a metacyclic expression site (MES). This monocistronic locus only contains the promoter (white flag) and the VSG gene (in red) and occasionally degenerate 70 bp-repeats in between (small stripped box). Upstream the promoter ESAGs (in blue), pseudo-ESAGs (in grey) or retrotransposons (light blue) can also be present. **(d)** Representation of a minichromosome. This genomic class is composed of 177 bp-repeats, 70 bp-repeats (vertical stripped boxes) and VSGs (in red) upstream the telomeric repeats. (Imagen taken from [90]).

The number of annotated VSGs has increased in the last years from 806 [23] (although these count only the VSG arrays genes) to 2563 genes [91], which may be explained by different analyzed strains, better technology for assembly and analysis and the type of chromosomes that were targeted in the studies. In the latter study, the VSGnome of the common laboratory Lister 427 strain was analyzed and showed interesting features about the VSG repertoire: of the 2563 VSG genes, only 325 (13%) are

fully functional while the remaining exist as incomplete genes or pseudogenes; 360 genes exist with two or more copies; nucleotide and protein sequences of Lister 427 VSGs are very dissimilar from those in TREU927 strain (less than 50%); although the nearly 100 minichromosomes have potential for 200 VSG genes, only 65 minichromosome associated VSG were identified meaning that not all minichromosomes contain VSGs. Also, this study identified six mVSG genes although in *T. b. rhodesiense* a set of 29 mVSG proteins has previously been detected [92].

VSG genes can only be transcribed from specialized loci, the BES in bloodstream forms or MES in metacyclic forms (Figure 2). In the genome of Lister 427 there are 15 BESs that are distributed in the subtelomeric regions of megabase and intermediate chromosomes [93]. BESs have ~45-60 kb in length and its structure is the following: a 50 bp-repeat sequence, a Pol I promoter, a variable number of expression site-associated genes (ESAGs) intercalated in some cases with pseudo-ESAGs, a 70 bp-repeat sequence, a VSG gene and the telomeric repeats (Figure 2) [93]. There are 12 ESAGs termed ESAG1-12 (Table 1), although each BES possesses a different combination of ESAGs, which may even include either absence or duplication of an ESAG according to the BES [93]. The expression of a specific BES with a specific combination of ESAGs may be related to the adaptation to different mammalian hosts [94]. It is noteworthy that BES8 is the only BES that could not be activated *in vitro*, having the particularity that contains a single ESAG7 besides its VSG [93].

T. b. rhodesiense has a special ESAG, the serum resistance-associated gene (SRA), which confers resistance to the human serum [95]. SRA is derived from a VSG [96] and can bind with high affinity to the trypanolytic factor apolipoprotein L1 (APOL1) inside the cell. SRA avoids the insertion of APOL1 in the plasma membrane for pore formation and cell death, and instead leads APOL1 for degradation as a mechanism of resistance in humans [97].

On the other hand, MES are also subtelomeric regions in the megabase chromosome and, structurally, are one of two examples, together with splice leader (SL) genes, of monocistronic transcription in *T. brucei* [98]. MES possess a small size between ~1-5 kb and the structure of a MES involves a promoter, highly degenerate 70-bp repeats

and a mVSG (Figure 2) [98, 99]. In some cases, a partial ESAG or a VSG pseudogene may be present upstream of the promoter [91, 100].

Table 1 - Predicted functions of ESAGs.

Name	Functions	Reference
ESAG1	Predicted integral membrane protein.	[94]
ESAG2	GPI-anchored protein localized in the flagellar pocket.	[94]
ESAG3	May be secreted or partners with another protein at the cell surface.	[94]
ESAG4	Adenylate cyclase.	[101, 102]
ESAG5	Predicted integral membrane protein with transmembrane domains. Similar to a human lipid binding/transfer family of proteins.	[94, 103]
ESAG6	Subunit of a heterodimeric receptor for transferrin.	[104]
ESAG7	Subunit of a heterodimeric receptor for transferrin.	[104]
ESAG8	Putative regulatory protein that accumulates in the nucleolus and may regulate RNA stability.	[105, 106]
ESAG9	GPI-anchored protein enriched in SS. Are highly shed and may interact with external environment.	[94, 107]
ESAG10	Encodes a homologue of *Leishmania* biopterin receptor. It is present in around half the detected BES but always upstream of the known BES promotes having itself a secondary promoter.	[93, 108]
ESAG11	GPI-anchored protein.	[94]
ESAG12	Unknown function.	[93]
SRA	Human serum resistance. Derived from a VSG. Specific of *T. b. rhodesiense*.	[95, 96]

GENERAL TRANSCRIPTION IN *TRYPANOSOMA BRUCEI*

Transcription in *T. brucei* is highly unusual as almost all genes are organized in polycistronic rather than monocistronic transcription units [23, 75]. There are around 150 PTUs bearing the housekeeping genes [109, 110], but with no apparent functional relation between the genes present in each PTU [111]. The regions between each PTU are named strand switch regions (SSR) which can transcriptionally be convergent or divergent. However, promoters for Pol II transcription have not been identified in *T. brucei* with the sole exception of the promoter that transcribed the splice leader genes [112, 113]. Strikingly, the great majority of the genes in *T. brucei* do not contain introns except for two genes: Poly(A) polymerase and an ATP-dependent DEAD/H RNA helicase [23, 114]. Unlike Pol I transcription (which will be approached in section 1.4), Pol II and Pol III transcription is constitutive. Thus, regulation of mRNA abundance is dependent of post-transcriptional mechanisms such as *trans*-splicing, polyadenylation and mRNA stability (albeit these processes are common to Pol I-transcribed PTUs), which will be further discussed. Other processes like mRNA export and translation control are also crucial for gene expression in trypanosomes [115-117]. However, these regulatory steps will not be further described as they fall beyond the scope of this dissertation.

RNA Polymerase II and III transcription

Pol II in *T. brucei* is composed of twelve subunits, termed RPB1-12 [118, 119]. Interestingly, these parasites contain two paralogues of RPB5 (RPB5 and RPB5z), RPB6 (RPB6 and RPB6z) and RPB10 (RPB10 and RPB10z), but only the non-z isoforms are associated to Pol II [120]. Discovery of Pol II promoters has remained elusive (except for SL genes) although the structural and spatial organization of putative bidirectional promoter has already been proposed [121]. Rather, probable transcription start sites (TSS) have been associated to a number of epigenetic marks: acetylated H4K10 (H4K10ac), H2A.Z, H2B.V, tri-methylated H3K4 (H3K4me3) histones and the bromodomain factor 3 (BDF3) are enriched at probable Pol II TSS [109, 122]. It is proposed that the presence of these histones in the nucleosomes of TSS make them unstable to facilitate nucleosome eviction prior to transcription [109]. At the end of the

PTUs, the transcription termination sites (TTS) are also enriched in specific epigenetic marks: H3V and H4V histone variants [109]. Loss of H3V leads to the production of small interfering RNAs (siRNAs) in regions of overlapping transcription at convergent strand switch regions [123]. Additionally, flanking regions of Pol II PTUs are enriched in base J [124]. Base J or β-D-glucopyranosyloxymethyluracil is a hypermodified DNA base which is only present in kinetoplastids which is generated in a two steps by thymidine hydroxylases JBP1 and JBP2 (first step) and a glucosyltransferase (second step) [125-127]. In *Leishmania*, loss of base J leads to transcriptional readthrough at TTSs with a dramatic effect in cell survival [128]. On the other hand, base J depletion in *T. brucei* does not have a readthrough effect at the TTSs but rather allows transcription of downstream genes within some specific PTUs, i.e. base J attenuates Pol II transcription in the middle of specific PTUs [129]. This effect was also observed upon depletion of H3V [123].

Being the only Pol II known promoter, SL gene promoter has been used to unveil the transcription factors behind Pol II transcription: TFIIB is essential and required for the association of Pol II to the preinitiation complex [130]; TATA-binding protein related factor 4 (TRF4) is also essential, being recruited to the gene promoter [131] and tightly associated with the tSNAP complex [132, 133]; and TFIIH is a basal transcription factor that works in transcription initiation and elongation [134, 135].

Pol III transcribes the 5S ribosomal DNA (rDNA), transfer RNAs (tRNAs), and a subset of U-rich small nuclear RNAs, which are dispersed throughout the megabase chromosomes [136-140]. Pol III is typically composed of seventeen subunits although several subunits are common to either Pol I or Pol II [141]. Only one transcription factor for Pol III transcription has been identified, the TFIIIB subunit BRF1 [133], although the structure of the tRNA gene promoters contains elements which allow the binding of TFIIIB and TFIIIC [142].

Determinants of mRNA levels

The generation of mature mRNAs from the polycistronic mRNA occurs through coupled *trans*-splicing and polyadenylation [143]. This coupled mechanism happens co-transcriptionally a few minutes after synthesis of the splicing acceptor site [144]. Within a

polycistronic mRNA, an upstream gene is polyadenylated at the 3' untranslated region (3'UTR) while a downstream gene is *trans*-spliced at the 5' untranslated region (5'UTR). *Trans*-splicing is a two-step transesterification reaction in which a 39-nucleotide sequence from the ~140-nucleotide SL RNA is capped, highly modified, and added to the 5'UTR of the mRNA [145, 146]. Contrary to *cis*-splicing, this process involves the formation of an intermediate Y structure instead of a lariat structure [27]. Both types of splicing possess the same motif requirements, a polypyrimidine [poly(Y)] tract, a GT dinucleotide at the 5' splice site, an AG dinucleotide at the 3' splice acceptor site and, eventually, exonic enhancer motifs [147-149]. Polyadenylation occurs by the addition of a chain of adenosine monophosphates to the 3'UTR by a complex that contains the Poly(A) polymerase [150]. The poly(Y) tract is also important for the polyadenylation signaling [151]. More than 32000 unique splice-acceptor sites have been identified for ~8900 genes and more than 50000 polyadenylation sites were identified for ~8000 genes [152], showing that some genes contain alternative splice-acceptor and polyadenylation sites.

The stability and degradation of the mRNAs (together with gene copy number) are key features to establish mRNA abundance and variability. As previously mentioned, the *T. brucei* transcriptome changes between the different life-cycle stages [43-45, 50-53]. Moreover, each BSF cell is calculated to contain 20000 mRNA molecules while each PF cell contains twice as much mRNA molecules [153]. In *T. brucei*, stability of mRNA depends on the presence of regulatory sequences in the 3'UTR and the binding of specific RNA-binding proteins. The importance of the 3'UTR in gene expression control was initially established when the 3'UTR of the procyclin, THT-1 transporter, aldolase, phosphoglycerate B and C mRNAs were found to be developmentally regulated [154, 155]. In procyclin mRNAs, a 16-mer stem-loop and a 26-mer poly(Y) sequences in the 3'UTR were unveiled as elements that are important for mRNA stability in PFs and degradation in BSFs [156]. RNA-binding proteins are also important for the stabilization of specific mRNAs and life-cycle progression [157]. As examples, knock-down of RBP10 leads to down-regulation of several enriched mRNAs in BSFs and its overexpression in PFs leads to accumulation of those mRNAs [158], knock-down of ALBA3/4 in procyclic forms leads to cell-cycle arrest together with nucleus repositioning [159] and overexpression of RBP6 in procyclics promotes developmental transitions until metacyclic forms [69].

The steady-state levels of an mRNA also depend on its rate of degradation. Usually, the degradation starts with the deadenylation by the CAF1-NOT and PAN2/PAN3 complexes [116, 160], followed by decapping and, ultimately, 5’ -> 3’ degradation by XRNA and 3’-> 5’ degradation by the exosome [161, 162]. For mRNAs with very short half-life, a direct decapping and rapid degradation by XRNA can occur, bypassing the deadenylation step [161].

These post-transcriptional mechanisms also occur in Pol I transcripts, although, regulation at the transcriptional level is probably more important as it will be explained below.

RNA POLYMERASE I TRANSCRIPTION

In almost all eukaryotic cells, Pol I solely transcribes rDNA genes which accounts for ~60% of all synthesized RNA [163], the rate of ribosome biosynthesis being proportional to the growth rate and proliferation [164]. rDNA transcription occurs in the nucleolus where 200 repeats (in yeast) or 300-400 (in humans) rDNA transcription units encode for three major rRNA species: 18S, 5.8S and 25S (28S in humans). A full 35S or 47S pre-rRNA precursor is synthesized, processed into the 3 rRNA species and, together with Pol III-synthesized 5S rRNA and Pol II-transcribed ribosomal proteins, are assembled in ribosomes and exported from the nucleus [165]. Besides the rDNA genes, the rDNA repeats also contain an intergenic spacer, possessing different lengths according to the species, that can harbor some genetic elements such as enhancers or termination sites [166].

In *T. brucei*, Pol I also synthesizes a 45S pre-rRNA which is processed into the 18S and 5.8S but the large subunit is processed into two large (LSUa and LSUb) and four small (sr1, 2, 4 and 6) RNA molecules [167]. However, this parasite possesses a highly reduced number of rRNA repeats: only 9 repeats which are divided amongst 4 different chromosomes [140]. Strikingly, African trypanosomes are the only known organisms in which Pol I machinery transcribes protein coding genes, specifically the genes present in procyclin, MES and BES loci [4]. Although not scientifically proven, this surprising fact may be a consequence of two factors: Pol I is known to be the fast-acting polymerase amongst the canonical polymerases allowing high transcription rate of surface protein mRNAs and the lack of Pol II transcriptional regulation may require another polymerase machinery to circumvent the life-cycle dependent necessities for different surface proteins.

RNA polymerase I machinery and transcription cycle in yeast and humans

Saccharomyces cerevisiae Pol I machinery is composed of fourteen subunits (Table 2), of which five are shared with Pol II and Pol III, two are shared only with Pol III and the remaining seven are specific of this complex [168].

Table 2 - RNA polymerase I subunits in *Saccharomyces cerevisiae* (yeast), *Homo sapiens* (human) and *Trypanosoma brucei* (adapted from [169]).

S. cerevisiae	*H. sapiens*	*T. brucei*
ScRPA190	HsRPA190	TbRPA1
ScRPA135	HsRPA127	TbRPA2
ScRPA49	PAF53	-
ScRPA43	HsRPA43	TbRPA31? [a]
ScRPA40	HsRPA40	TbRPC40
ScRPA34.5	CAST	-
ScRPB5	HsRPB5	TbRPB5z
ScRPB6	HsRPB6	TbRPB6z
ScRPA19	HsRPA19	TbRPC19
ScRPB8	HsRPB8	TbRPB8
ScRPA14	-	-
ScRPA12.2	HsRPA12.2	TbRPA12
ScRPB10	HsRPB10	TbRPB10z
ScRPB12	HsRPB12	? [b]

[a] It has been proposed that the trypanosomatid-specific TbRPA31 may be the ortholog of RPA43.

[b] Although *in silico* approaches have identified a putative TbRPB12, no experimental evidence has been identified for this factor in RNA Pol I composition of *T. brucei*.

The transcription cycle is the sequence of events that start with the assembly of the pre-initiation complex (PIC) at the promoter until the termination step that releases the transcript and the polymerase. Transcription initiation with the formation of PIC in yeast is dependent of four factors: Rrn3, core factor, TATA-binding protein (TBP) and upstream activating factor (UAF). Rrn3 binds to RPA43 subunit to allow bridging between the polymerase and the core factor in the rDNA promoter [170]. The core factor is a three subunit complex essential for transcription that enhances polymerase binding to the promoter and positioning in the transcription start site [171, 172]. UAF is a six subunit complex in which Uaf30 subunit provides efficient binding to the rDNA promoter allowing activation of the rDNA genes [173, 174]. TBP, although not required for basal transcription [172], promotes full activation of transcription by binding to UAF and

helping to recruit the core factor [175]. Transcription elongation requires subcomplex RPA49 and RPA34.5 to promoter elongation [176] as well as topoisomerases I and II to remove the accumulation of negative (in the wake of polymerase) and positive (ahead of the polymerase) supercoiling [177, 178]. Moreover, Spt4/5 heterodimer is considered to be a negative regulator of elongation [179], while Paf1 complex increases the rate of elongation of Pol I [180]. To terminate the transcription, 10-15 bp T-rich sequence together with proteins Reb1 and Nsf1 are important to pause Pol I transcription and release the transcript [181, 182].

The human Pol I complex is very similar to the yeast complex as homologues of all subunits except of RPA14 have been identified [169] (Table 2). Formation of the PIC in humans is dependent of Rrn3, selectivity factor 1 (SL1, homologue of the core factor) and upstream binding factor (UBF). Human Rrn3 fulfills the same function as in yeast bridging the polymerase to SL1 [183]. SL1 is composed of TBP and at least three TBP-associated proteins and also mediates recruitment of Pol I which has no sequence-specific DNA binding activity [184]. UBF is not functionally related to UAF since it is a high mobility group box (HMGB) protein and is required to activate transcription through the interaction of its acidic C terminus with SL1 [184, 185]. The role of UBF as a HMGB and its importance for Pol I transcription will be discussed ahead. Human Pol I transcription activation is promoted by hPAF53/CAST heterodimer which is the ortholog of RPA49 and RPA34.5 subcomplex [186] and elongation is regulated by the phosphorylation of HMG boxes of UBF [187]. Transcription termination also depends of a T-rich stretch upstream of a 11 bp Sal box [188]. This Sal box allows the binding of TTF-I to promote transcription pausing. In mice, the factor PTRF binds to Pol I and TTF-I to allow release of the polymerase and transcript from the DNA template [189].

Epigenetic control of rDNA genes in yeast and humans

Although yeast and mammalian nucleoli contain hundreds of rDNA repeats, only ~50% are transcriptionally active [190, 191]. Ribosomal RNA genes are organized into Nucleolus Organizer Regions (NORs) which can exist between amongst active and inactive states, which show very distinct chromatin structures. Initial studies using psoralen-

crosslinking showed that two chromatin conformations co-exist in the nucleoli [192]: one is nucleosome-enriched, i.e. closed chromatin, and corresponds to inactive repeats whereas the other contain much less nucleosome content in the gene body, i.e. open chromatin, and corresponds to the active repeats. However, it has been reported that even rDNA active repeats contain histones in a dynamic chromatin structure subjected to remodeling [193]. More recently, a third chromatin structure in rDNA genes was identified in mammalian cells. It consists in a poised chromatin structure prone to transcriptional activation and it is characterized by an unmethylated promoter associated with PIC components but not Pol I, together with a bivalent histone modification in nucleosome specifically localized in a transcriptional repressive position [194]. Several epigenetic regulators control the establishment and maintenance of the different chromatin states through cell-cycle progression, which range from DNA methylation proteins, histone modifying enzymes, chromatin remodelers and non-coding RNAs.

In mammalian cells, active rDNA repeats are enriched in UBF, nucleolin, DNA hypomethylation and also contain histone modifications such as H3K4me3 and acetylation of histones H3 and H4 [195] (Figure 3). Activation of rDNA repeats is highly dependent of the transcription termination factor TTF-I. This factor is able to bind the termination sequence T_0 localized just upstream the TSS and also the termination sequence $T_{1\text{-}10}$ localized downstream the transcription unit. TTF-I can oligomerize and bridge a DNA loop between T_0 and $T_{1\text{-}10}$ to spatially create a chromatin hub of high transcriptional activity [196]. Also, TTF-I recruits the Cockayne Syndrome protein B (CSB) to promote Pol I transcription elongation [197]. CSB is also responsible for the recruitment of the histone H4 and H3K9 acetyltransferase PCAF to poised rDNA genes to promote association of Pol I to the promoters and, consequently, initiation of rDNA transcription [198].

Mammalian inactive or silent rDNA repeats contain several epigenetic marks such as trimethylation of H3K9, H4K20 and H3K27, histone H4 deacetylation or CpG methylation in the promoter or enhancer [199-202] (Figure 3). Here, TTF-I is also considered a key regulator to repress rDNA genes. TTF-I can interact with the large subunit of the Nucleolar Remodeling Complex (NoRC) TIP5 (for TTF-I interacting protein 5), leading to the recruitment of DNA methyltransferases, histone methyltransferase and

deacetylation enzymes and repositioning of the promoter-proximal nucleosome further downstream to the TSS blocking pre-RNA synthesis [200, 203, 204]. In this context, non-coding RNAs also play important roles in rDNA chromatin establishment. A promoter-associated RNA (pRNA) derived from a Pol I promoter ~2kb upstream of the rDNA TSS is crucial to bind TIP5 after NoRC repositioned the promoter-proximal nucleosome [205, 206]. The pRNA is able to recruit DNMT3b and methylate CpG residue at position -133 inhibiting UBF binding and PIC formation, together with histone deacetylases and methyltransferases [204, 207, 208]. CpG methylation is also important for rDNA silencing as inactivation of DNA methyltransferases DNMT1 and DNMT3b or inhibition by 5-aza-2'-deoxycytidine leads to the opening of silent repeats and even cryptic transcription by RNA Pol II [209, 210].

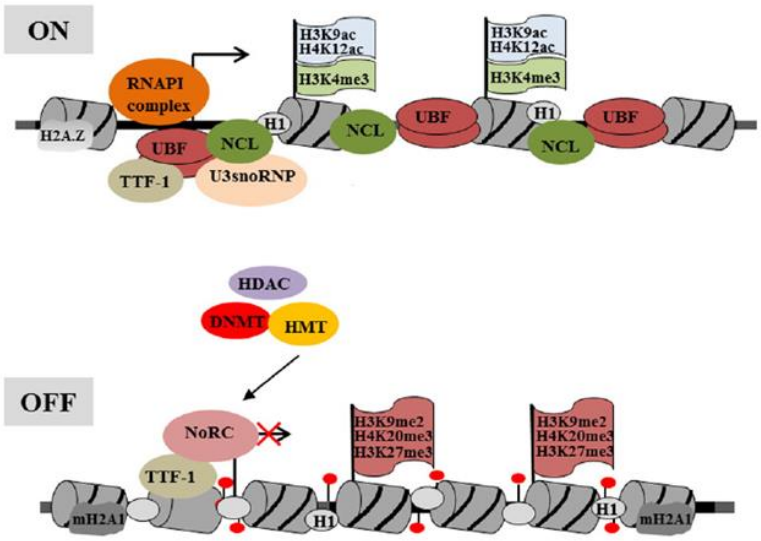


Figure 3 – Structural organization of rDNA repeats in mammalian cells. RNA polymerase I actively transcribed repeats (ON) possess a chromatin enriched in UBF and nucleolin (NCL). Histone modifications occur in histones along the gene body. Silenced repeats (OFF) have a very compact chromatin with several histone methylation marks. The nucleolar remodeling complex (NoRC) recruits DNA methyltransferases (DNMT), histone methyltransferase (HMT) and deacetylation (HDAC) enzymes to establish a heterochromatin status and block transcription. In both status, TTF-I is important to recruit factors that promote both transcription and heterochromatin establishment. (Image adapted from [211])

In conditions of cell growth arrest, a poised chromatin establishment occurs and is dependent of the NuRD (nucleosome remodeling and deacetylation) complex [212]. This

complex is recruited by TTF-I and CSB to unmethylated promoters in order to form a repressive nucleosome in the core promoters which as the bivalent histone mark H3K4me3 and H3K27me3 [194]. Moreover, NuRD can negatively regulate TIP5 avoiding the CpG methylation at the promoter [213]. When mammalian cells enter a starvation state, the rDNA chromatin is silenced by eNoSC (for energy-dependent nucleolar silencing complex) in an unstable and reversible manner to conserve energy until suitable energetic conditions occur again [214]. eNoSC binds to rDNA promoters to deacetylate and methylate histone H3 causing the establishment of heterochromatin through the rDNA loci [215].

In yeast, it was shown that a strain with only 42 rDNA repeats in the genome did not significantly change the rate of rRNA biosynthesis since the loading of Pol I in the promoter and transcription initiation becomes increased [216]. Thus, it is not the number of active rDNA repeats that determines transcription rate. Ribosomal RNA silencing in yeast occurs through two different pathways (Figure 4): Sir2 (for silent information regulator 2) dependent or independent pathways.

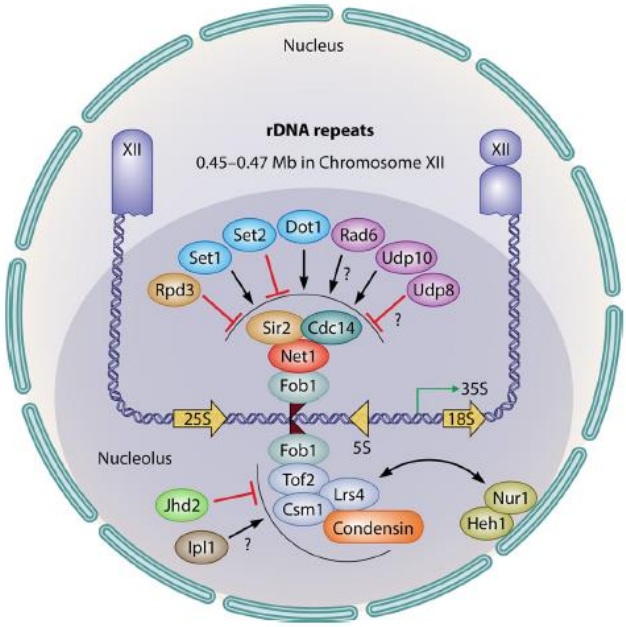


Figure 4 – Pathways for yeast rDNA silencing. Sir2-dependent pathway (above the DNA duplex) requires the RENT complex (Sir2, Cdc14 and Net1) to deacetylate histones H3 and H4. Sir-independent pathway (below the DNA duplex) is based on the binding of Tof2, Csm1 and Lrs4 to induce rDNA silencing. Both pathways require the binding of Fob1 to the nontranscribed spacer 1 region to recruit both complexes. Several factors exert positive (black arrow) or negative (red blocking signal) effects on the rDNA silencing process. (Image taken from [211])

The former is dependent of the RENT complex (regulator of nucleolar silencing and telophase) which is composed of Sir2, Net1 and Cdc14 [217, 218] in which Sir2 exerts its deacetylase activity in histones H3 and H4 to start inducing heterochromatin formation [219]. As for the Sir2-independent pathway, a hierarchical binding of the proteins Tof2, Csm1 and Lrs4 is also necessary to rDNA silencing [220]. Moreover, in both these pathways the binding of the complexes to the rDNA occurs in the nontranscribed spacer 1 region, specifically through the interaction with Fob1 protein and both act synergistically to ensure rDNA silencing at the loci [220]. Several histone modifying enzymes and inner nuclear membrane associated proteins are able to modulate the action of these two pathways (reviewed in [211], Figure 4).

One of the major differences between mammalian and yeast rDNA gene control is that, although also having a proportion of rDNA genes in silent state, DNA methylation does not occur in yeast cells since its genome does not contain DNA methylases. However, remodeling of rDNA chromatin in yeast is related with the cell-cycle stage. Thus, when cells enter the S stage of the cell-cycle a replication-dependent deposition of nucleosomes occurs in almost all the active rDNA genes which will adopt a closed conformation [221]. A similar effect has also been previously observed in a human cell-line [222]. Re-entrance in G_2 stage leads to Pol I transcription-dependent opening of a random subset of rDNA genes [221, 223]. However, when the replication stage is blocked, closed rDNA genes start adopting an open conformation stabilized by the HMGB protein Hmo1. This protein also stabilizes active rDNA genes when transcription is halted under G_1 arrest [221]. Other physiological processes such as DNA damage or stationary growth have also been associated to chromatin assembly into the active rDNA repeats, which are reversible when normal condition are re-established [224, 225]. Importantly, when yeast cells enter the stationary phase of growth, the transcription was highly altered but the proportion of active/silent genes was very small, suggesting that in these conditions the transcription is primarily regulated rather than the chromatin states of rDNA [225].

RNA polymerase I machinery and transcription cycle in *T. brucei*

In silico studies have predicted twelve homologs of the yeast Pol I subunits in *T. brucei* genome, of which ten have already been characterized [226-228] (Table 2). Another subunit, the Pol II RPB7 which was suggested to be yeast RPA43 ortholog, was also shown to be involved in rDNA and BES transcription in BSFs [229], but later studies proved otherwise [230]. On the other hand, it was proposed that RPA31, a subunit only conserved amongst trypanosomatids, may be the yeast RPA43 counterpart [227, 230]. Also, only the z isoforms of the existing paralogues of RPB5 (RPB5 and RPB5z), RPB6 (RPB6 and RPB6z) and RPB10 (RPB10 and RPB10z) are associated to Pol I [120].

Structure of Pol I promoters in *T. brucei* is very diversified [231]. Procyclin and rDNA promoters are similar between them, while BESs promoters are shorter but conserved amongst them. MES promoters are apparently shorter but not conserved amongst them. However, even though their structure is different, rDNA promoter is interchangeable with a BES promoter [232]. Pol I recruitment to its promoters is dependent of class I transcription factor A (CITFA) complex, composed of eight subunits [233, 234]. Dynein light chain LC8 subunit binds CITFA2 subunit promoting CITFA complex assembly and promoter binding [235], but interaction of CITFA with Pol I remains unknown. Thus, CITFA is highly enriched in rDNA and active BES promoters in BSFs, although small levels are also detected in silent BESs promoters [236]. In terms of transcription elongation, ELP3b, an ortholog of yeast Pol II Elongator subunit Elp3, might have a role in this process as it is a negative regulator of rDNA, but surprisingly not BES, transcription [237]. Little is known about termination of Pol I transcription in *T. brucei*.

Epigenetic control of rDNA genes in *T. brucei*

The epigenetic regulation of ribosomal genes in *T. brucei* is still poorly understood. Of the 9 rDNA repeats, it is still unknown if all are active or only a fraction. Nonetheless, one may speculate that all or almost all possess an open chromatin conformation since depletion of histone H1 which is a linker between nucleosomes does not lead to conformational changes in rDNA promoters or 18S gene [238]. However, chromatin of genes downstream of 18S subunit has not been evaluated.

Table 3 - Histones and chromatin-associated proteins involved in RNA polymerase I gene expression regulation in *Trypanosoma brucei*.

Class	Name	Functions	Reference
Canonical histones	H1[a], H3[b]	Maintaining BESs silent[a,b], maintaining other Pol I loci silent[a], VSG switching[a], cell cycle[b]	[238-240]
Histone deacetylase	DAC3	Maintaining BESs silent	[12]
Histone acetyltransferases	HAT1*[a], HAT2[b], HAT3[c] ELP3b*[d]	Telomeric silencing[a], VSG switching[c], rDNA transcription[d], cell cycle[a,b], DNA repair[c], H4K10ac[b], H4K4ac[c]	[237, 241-243]
Histone methyltransferases	DOT1A[a], DOT1B[b]	VSG switching[b], active BES attenuation[b]; cell cycle[a,b] differentiation to PFs[b]; H3K76 me/me2[a] H3K76me3[b]	[244-246]
SUMOylation enzymes	SIZ1	Maintaining BES active, recruitment of Pol I	[247]
Chromatin remodelers	ISWI[a], NLP[b], RCCP[c], FYRP[d]	Maintaining BESs silent[a-d]/active[b], maintaining other Pol I loci silent[a]	[248-250]
Histone chaperones	ASF1A[a], CAF-1b[b], FACT[c], NLP[d]	Telomeric silencing[a,b], maintaining BESs silent[a-d]/active[c,d], maintaining other Pol I loci silent[d]	[240, 251-253]
Chromatin architectural proteins	TDP-1	Maintaining BES active, maintaining other Pol I loci active	[14]
Telomeric proteins	RAP1[a], TIF2[b], TRF[c], TERT[d]	Maintaining BESs silent[a], VSG switching[b-d], cell cycle[a,c], telomere and subtelomere integrity[b-d]	[254-262]
Nuclear architecture	NUP-1[a] PIP5Pase[b], PIP5K[c]	Maintaining BESs silent[a-c], maintaining other Pol I loci silent[a,b], VSG switching[a,c], Pol I localization[b,c], nucleus integrity[a], cell cycle[a], chromosome positioning[a-c]	[263, 264]
Replication proteins	Cohesin complex[a], MCM-BP[b] ORC1/CDC6[c],	Inheritance of active BES[a], maintaining BESs silent[b,c], maintaining other Pol I loci silent[b,c], VSG switching[a,c], widespread mRNA abundance[c], cell cycle[a-c]	[265-268]
Bromodomain proteins	BDF2[a], BDF3[b]	Maintaining BESs silent [a-b], maintain genomic VSGs silent[b], maintaining other Pol I loci silent [a-b], cell cycle[a-b]	[269]
Others	VEX1	Sequestration by active BES for monoallelic expression, Pol I sequence homology-dependent silencing	[270]

*No experimental evidence for the respective enzymatic/chromatin binding activity; a,b,c,d,e indicate the roles identified for the respective factor.

So far only three factors are known to influence the status of rDNA genes (Table 3): TDP1, a high mobility group box protein that binds rDNA active repeats and facilitates Pol I transcription (TDP1 will be extensively described in section 1.4.7) [14]; ELP3b, a specific rDNA negative regulator of elongation, which possesses a putative histone acetyltransferase activity although it is still not understood how it inhibits transcription [237]; and ISWI has a role as a transcriptional silencer in the rDNA spacer region between two repeats, possibly to reinforce this sequence function as an insulator [250].

Epigenetic control of procyclin loci in *T. brucei*

Even though procyclin loci are transcribed by Pol I, their location lies within the core of megabase chromosomes 6 and 10 [140]. These loci are subdivided between different clusters according to the genes each one shelters [271] (Figure 5). Two clusters contain EP1 and EP2 procyclin genes downstream the promoter but while one cluster also expresses procyclin associated gene 1 (PAG1), the other cluster expresses PAG2. Other PAG genes may be present downstream, whereas at least PAG1 was shown to serve as an insulator for downstream transcription of Pol I [272]. Another cluster is composed by GPEET procyclin, followed by EP3 procyclin, PAG3 and GRESAG2 (gene related to ESAG2) genes. EP genes may have different isoforms depending upon parasite strain [56, 273].

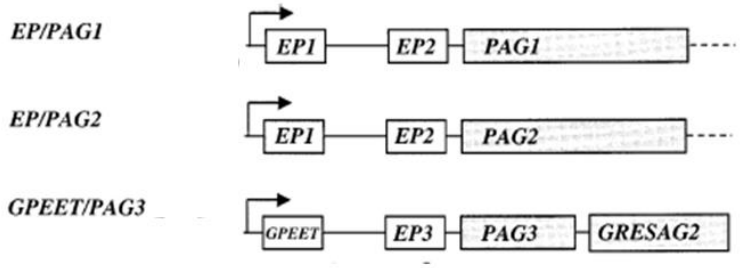


Figure 5 – Organization of the procyclin loci. Two different EP loci are distinguished by the presence of different PAG genes while GPEET locus has a different genomic constitution relative to EP loci. (Image adapted from [271])

Control of procyclin loci is dependent of transcriptional and post-transcriptional levels. When inserting a reporter gene with an actin 3'UTR (not developmentally

regulated) downstream a procyclin promoter, its mRNA levels were ~8.6-fold lower in BSFs compared to PFs, indicating that the procyclin promoter is down-regulated in BSFs [274]. However, procyclin mRNA levels are much lower than 8.6-fold in BSFs relative to PFs. This is due to the elements present in the 3'UTR, which make procyclin mRNA unstable in BSFs, prone to rapid turnover as well as affecting the translation [156]. Moreover, differentiation stimuli such as temperature, *cis*-aconitate and incubation medium are involved in transcription elongation of procyclin genes [275]. Transcriptional, post-transcriptional and translational regulation causes a 1000-fold downregulation of procyclin levels in BSFs [276]. Nonetheless, even within PFs a strict control of expression of EP and GPEET is undertaken. Whilst in the first 7 days of differentiation, both EP and GPEET procyclins are expressed, after day 7 only EP procyclin is expressed [277]. GPEET mRNA becomes destabilized through a glycerol-responsive element in the 3'UTR, a mechanism that is dependent on the activity of mitochondrial enzymes [278].

Chromatin associated factors also play an important role in the control of procyclin expression (Table 3). In PFs, the procyclin loci are located in the nucleolar periphery (Figure 6) [11]. In BSFs, it is not directly proven where these loci are located. However, upon knockdown (KD) of NUP-1 lamina-like protein in BSFs, which regulates nuclear periphery heterochromatin organization, procyclin genes become upregulated suggesting that procyclin loci may be positioned in the nuclear periphery [263].

Transcription from EP1 loci promoter in BSFs becomes ~4-fold and ~6-fold upregulated upon nucleoplasmin-like protein (NLP) and chromatin remodeler ISWI KD, respectively [250, 251]. A conditional knock-out of a component of a replication complex MCM-BP in BSFs leads to ~5-8 fold transcript level increase in EP1, while PAG4 and PAG5 downstream of EP1 are upregulated 8-15 fold [266]. Also, depletion of the linker histone H1 causes a decondensation of chromatin and a concomitant EP1 and EP2 transcriptional upregulation [238].

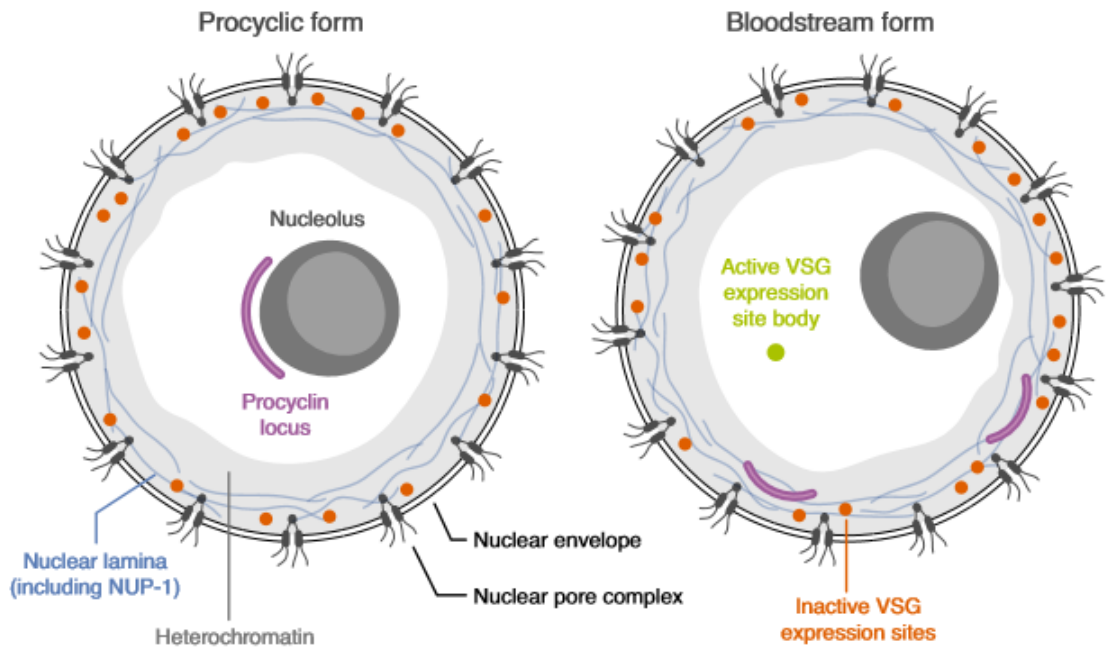


Figure 6 – Nuclear localization of RNA Polymerase I transcribed loci. Heterochromatin is represented in light grey around nuclear periphery. Silent Bloodstream (and possibly Metacyclic) Expression sites in both procyclic and bloodstream forms as well as procyclin loci solely in bloodstream forms are tethered to the nuclear periphery probably by NUP-1. Active Bloodstream Expression Site in bloodstream forms assumes a more central position to be transcribed, whereas procyclin loci in procyclic forms are transcribed perinucleolarly (Image taken from [279]).

Epigenetic control of Expression Sites in *T. brucei*

Control of BES in BSFs is of outmost importance in order to display VSG monoallelic expression. For this, all except one BES are required to be in a silent state leaving a single BES in an active status. In PFs, all BES need to become silent so that only procyclins become expressed. As for MES, these only become activated specifically in metacyclic forms, being tightly repressed in all other life-cycle stages.

Factors involved in the maintenance of silent BESs

Silent BESs possess a very compact and nucleosome enriched chromatin [9, 10], which most likely ensures a layer of control regarding monoallelic expression. In the recent years, several studies identified and characterized epigenetic factors that, once depleted, result in transcriptional derepression of the promoter region or the telomeric

region or even the entire silent BESs (Table 3). Below I describe these factors, although it is not yet understood how these factors are concerted to maintain a silent BES status.

Interestingly, for the majority of such factors, transcriptional derepression is limited to the BES promoter regions without detectable changes in the transcript levels of the terminally located VSG genes or low VSG transcripts are detected but no detectable VSG protein is expressed. Depletion of the histone H1 leads to a 4-14 fold transcriptional upregulation including silent VSG genes but protein expression occurred only from promoter region genes [238, 239]. In accordance, depletion of histone H3 causes expression of a promoter-proximal GFP::NPT reporter cassette [240]. The subunits Spt16 and Pob3 of the histone chaperone FACT are important to stabilize compact chromatin structure at silent BES promoter regions [252, 253], while the histone chaperones ASF1A, CAF-1b and the histone deacetylase DAC3 present a marked promoter-proximal derepression [12, 240]. The knock-down of the chromatin remodeler ISWI or any of its interacting factors RCCP, FYRP and NLP which overall compose the ISWI complex also cause derepression of a reporter gene that localizes just downstream the promoter [248-251]. Additionally histone H1, Spt16 and DAC3 were shown to derepress silent BES also in PFs [12, 239, 252]. The role of some of these factors in BES silencing is also tightly associated with cell cycle progression since derepression of silent BES promoter regions occurs only in S-phase (CAF-1b depleted mutant) or G2/M (histone H3, Spt16 and CAF-1b depleted mutants) [252, 280].

Other epigenetic factors are important to silence BESs all the way until the telomere, as loss-of-function mutants result in derepression not only of genes nearby the BES promoter but also of silent VSG genes. These factors include the histone H3K76 methyltransferase DOT1B, the telomeric protein RAP1, the lamina-like NUP-1, two proteins of the replication machinery, ORC1/CDC6, MCM-BP and VEX1 (Table 3). Moreover, deficiency in any of these factors results in a significant percentage of trypanosomes expressing more than one VSG at the cell surface [245, 258, 266, 267, 270]. In contrast, some factors such as the histone deacetylase Sir2rp1 and the putative histone acetlytransferase HAT1, appear to be important exclusively for maintaining telomere-proximal silencing but not for silencing the BESs [242, 281] indicating that the mechanisms of epigenetic regulation acting at BESs and telomeres do not simply overlap.

Although silent BESs chromatin has a closed conformation, transcription is not entirely silenced. In fact, the presence of low levels of CITFA in silent BES promoters [236] may explain the very low transcript levels observed from regions that can reach ESAG7 and ESAG6 [282], even at the single cell level [283]. Moreover, these studies showed that multiple silent BESs possess this low transcription level. Interestingly, this same phenotype is observed in PFs, where several BES (which are all in a silent state) show a slight transcription just downstream the promoter [284].

Localization of silent BESs within the nucleus may also play an important role for BES silencing. FISH analysis suggests that most telomeres co-localize with NUP-1, a lamina-like protein localized at the nuclear periphery that is likely a nuclear lamina component [263] (Figure 6). Depletion of NUP-1 disrupts nuclear morphology and leads to re-localization of telomeres from megabase chromosomes to nuclear blebs, with a concomitant derepression of silent BESs, indicating that nuclear structure is important for silencing gene expression at BESs. Two enzymes of the inositol phosphate pathway (PIP5Pase and PIP5K) were recently shown to be required for BES silencing and for the correct positioning of telomeres (together with the telomeric protein RAP1) and Pol I in the nucleus, supporting the impact of nuclear architecture in BES silencing [264]. It is however unknown how these enzymes or their associated metabolites epigenetically affect the telomeres and BESs.

Factors involved in the maintenance of the active BES

The structure and dynamics of the active BES is poorly understood and very few epigenetic regulators have been shown to control it, which will now be described (Table 3).

Unlike silent BESs, the active BES possesses an open and nucleosome-depleted chromatin [9, 10]. Four proteins have been identified as being important for maintaining the transcriptional status of the active BES: TDP1, NLP, the Spt16 subunit of FACT and a cohesin component (SCC1). TDP1 is a high mobility group box protein that binds along the entire BES and its depletion drives chromatin condensation and a decrease in active VSG

transcript levels [14]. Depletion of NLP also reduces gene expression at the promoter region of active BES, although it is not clear if this results from an indirect consequence of silent BES derepression also resulting from NLP depletion or from a double role of this putative chaperone in active and silent BESs. Knockdown of SPT16 also reduces the active VSG transcript levels besides its role in silent BES promoter-proximal silencing [252]. On the other hand, the histone deacetylase DAC1 antagonizes telomeric silencing but has no attributed role in gene expression at the active BES [12].

The active BES localizes in a Pol I extranucleolar body termed Expression Site Body (ESB), which appears to be a stable nuclear subcompartment in BSFs [285]. Besides the previous factors, VEX1 is sequestered in the ESB albeit the transcriptional status of the active BES is not significantly altered upon KD [270]. Interestingly, depletion of SSC1, a component of the cohesin complex, revealed that cohesion is necessary for the epigenetic inheritance of the active transcriptional status of BES and mediates a delayed separation of the sister chromatids from the active BES locus during early mitosis, indicating that both events are probably functionally related [265]. Although the exact composition and the dynamics of this ESB remain elusive, it is possible that it retains factors that are essential for the chromatin and transcriptional status necessary for VSG expression, thus contributing to monoallelic expression of the unique BES that locates in the ESB.

Post-translational modifications also seem important to maintain the status of the active BES. Accumulation of SUMOylated chromatin-associated proteins occurs throughout the BES including in the RPA1 subunit of Pol I. Reduced levels of this modification by depletion of a SUMOylation enzyme, the E3 ligase SIZ1, induces a drop in the recruitment of Pol I to the active BES and a consequent transcriptional downregulation [247].

Additional layers of BESs control

During the life cycle of *T. brucei*, parasites shift between the mammalian host and the insect vector [286]. To survive in such different hosts, significant changes need to be

undergone in gene expression, which include replacement of the VSG by the procyclins surface coat [287]. Silencing of the active BES during differentiation is characterized by a progressive downregulation of transcription along the BES [288], together with a re-localization of the BES to the heterochromatic nuclear periphery [11]. This BES re-localization during differentiation to the nuclear periphery is highly affected by NUP-1 depletion [263]. It has also been shown that the chromatin of the originally active BES becomes less open to T7 polymerase, suggesting that it acquires a more compact structure during differentiation [289].

How is a MES controlled during the life-cycle?

Besides the major difference in genome organization between BES and MES in which the former is polycistronic transcription unit and the latter is a monocistronic transcription unit, the control of MES is also unique. MES are activated *in situ* and are, so far, the only genomic region which is subjected to a transcriptional control rather than post-transcriptional regulation as transcript levels of mVSGs appear strictly in metacyclic forms and in no other life-cycle stage [290, 291]. However, genetic elements such as the promoters are involved in this control. When a MES promoter was introduced in a non-transcribed rDNA spacer in PFs, the promoter remained transcriptionally silent [292]. On the contrary, in BSFs, the promoter activity increased 50-fold [292]. Moreover, the lamina-like NUP-1, when depleted by RNA interference (RNAi) in PFs, leads to mVSGs upregulation of ~6-22 fold, which suggests that MES localize in the nuclear periphery and this localization may be important for complete silencing in other life-cycle stages than metacyclic forms [263]. Also, downregulation of ORC1/CDC6 also leads to a ~5-13 fold derepression of mVSGs in PFs, whose biological significance remains unknown [268].

HIGH MOBILITY GROUP BOX PROTEINS IN RNA POLYMERASE I TRANSCRIPTION

HMGB proteins are characterized by their small size, rapid mobility and by containing at least one high mobility group box domain, responsible for DNA binding in either a sequence or non-sequence specificity [293, 294]. The binding of the HMG-box domain to bent, kinked or unwounded DNA allows to bend, distort or loop DNA structures [295, 296]. HMGB proteins have been reported to fulfill several important roles that range from extracellular mediator of inflammatory responses, angiogenesis or wound repair [297-299], to nuclear regulator of essential processes such as transcription, replication, DNA repair or V(D)J recombination [300-303].

HMGB proteins were shown to influence transcription by three different mechanisms: they facilitate nucleosome sliding by chromatin remodeling proteins [304, 305], they serve as a transient chaperone for stable binding of transcription factors [306] or they participate in transcription blockage as part of a complex that inhibits the assembly of a preinitiation complex on promoters [307]. HMGB proteins also participate in the regulation of rDNA genes, which are transcribed by RNA polymerase I (Pol I).

HMO1

Hmo1 is one of the ten High Mobility Group proteins in *S. cerevisiae* that has a role in rDNA transcription machinery [308]. It contains two HMG box domains: box A with low DNA binding affinity and some structural specificity and box B with high DNA binding affinity but lower structural specificity [309]. Still, both boxes were shown to be important for DNA binding *in vivo* and *in vitro* [310]. Moreover, Hmo1 contains a lysine-rich C-terminal region which is important for the DNA bending by the whole protein [311]. Hmo1 binds throughout the entire rDNA sequence of the actively transcribed repeats, which are devoid of histones, although it also binds to most promoters of ribosomal protein genes that are transcribed by Pol II [312, 313]. In these ribosomal protein gene promoters, Hmo1 interacts with TFIID and also delineates the 5’ and 3’ boundaries of the PIC assembly zone to direct the transcription initiation in the appropriate site [314, 315]. In rDNA genes, Hmo1 coordinates Pol I transcription under the control of the TOR

pathway [316] and is able to stabilize the 35S chromatin in an open conformation in the absence of Pol I transcription preventing nucleosome assembly outside cell-cycle S stage [221].

Hmo1 has been implicated in other roles other than transcriptional regulation, including directing replication-associated DNA lesions to the error-free DNA damage tolerance pathway [302], stimulating the nucleosome sliding activity of SWI/SNF and ISWIa chromatin remodelers [305] and stabilizing the chromatin during occurrence of DNA double-strand breaks (DSBs) [317]. Hmo1 null-mutants are viable but possess limited growth defects [308]. On the other hand, overexpression of Hmo1 leads to a vegetative growth [318], although it is able to rescue the growth defects and lethality presented by a null-mutant of Rpa49, a conserved subunit of Pol I [308].

UBF

UBF is a *bona fide* nucleolar specific HMGB protein common to mammalian species and it contains six HMG box domains [185, 294]. There are two isoforms, UBF1 and UBF2, with 97 kDa and 94 kDa, respectively, as a consequence of alternative splicing [319] and UBF exists in homodimer or heterodimer of these isoforms. However, UBF2 is not as efficient as UBF1 since it lacks 37 aminoacids in box domain 2 that are important for binding in rDNA promoter [320]. Another study showed that box domain 1 is necessary and sufficient to DNA sequence specificity while box domain 4 might be important to establish interactions with SL1 [321]. UBF is distributed throughout the entire rDNA active repeats [322]. Multiple roles have been assigned to UBF that include involvement not only in transcription, but also in chromatin shaping. In transcription, UBF roles encompass the interaction with SL1 to recruit Pol I to the promoter to form the PIC [184, 185], activation of transcription by stimulating Pol I promoter escape in the transition between initiation and the elongation where the interaction with the hPAF53/CAST subunit of Pol I is crucial [186, 323] and promoting elongation by remodeling and unfolding the enhancesomes for Pol I transcription, an action that is dependent of the phosphorylation of box domains 1 and 2 by ERK [324].

UBF is a potent chromatin regulator as depletion of UBF1 leads to an increase of inactive rDNA genes in a methylation-independent manner by promoting higher content of linker histone H1 in this inactive chromatin [325]. Moreover, targeting UBF via a lac repressor fusion protein to a heterochromatic region induces large scale decondensation of the chromatin [326]. Since overexpression of UBF increases rDNA transcription in neonatal cardiomyocytes [327], it suggests UBF overexpression and consequent chromatin decondensation may lead to the recruitment of SL1 and Pol I to induce increased transcription.

TDP1

TDP1 stands for Trypanosome DNA-binding Protein 1 and it was identified as a HMG protein by screening an expression cDNA library of *T. b. rhodesiense* [13]. TDP1 is encoded by a single gene but it results in two stable mRNAs with two different lengths, 1.6 kb and 2.3 kb. The difference in length is due to differential polyadenylation sites. Interestingly, these two mRNA are differentially expressed since the 1.6 kb mRNA accounts for 90% of total TDP1 mRNA in BSFs but only 40% in PFs.

Using RNAi against the coding sequence, it was previously shown that TDP1 depletion causes growth defects in BSFs but not in PFs [14, 240]. More recently, the Rudenko lab has characterized TDP1 as a HMGB protein with two HMG-box domains and a DEK-C DNA binding domain, although lacking the typical acidic C-terminal tail [14]. TDP1 is enriched in rDNA genes in BSFs and PFs, active BES in BSFs and procyclin genes in both PFs and even in BSFs, which is consistent with the low transcriptional differences of procyclin between these two life cycle stages (the high post-transcriptional control of procyclin mRNA in BSFs greatly reduces these levels). Depletion of TDP1 leads to an increase of histones H1, H2A and H3 content in active BES and rDNA repeats and a concomitant decrease of transcripts derived from these loci. Thus, TDP1 is considered a Pol I transcription facilitator functionally related with Hmo1 and UBF.

ANTIGENIC VARIATION

Antigenic variation is a sophisticated mechanism that several pathogens employ to evade the host immune system, which only provides a temporary natural immunity. Understanding this process is of outmost importance as the most successful vaccines are effective against low antigenic diversity pathogens and new therapeutic strategies are required.

The ability to change the immunogenic epitopes recognized by the immune system is found in viruses, bacteria and several eukaryotic pathogens. In *Influenza* virus, the surface glycoproteins hemagglutinin (HA) and neuraminidase (NA) may suffer changes through specific processes termed antigenic drift and antigenic shift [328]. The former is based on aminoacid substitutions in HA and/or NA and the latter shows a new subtype of mixed HA and NA. Selection pressure in the environment is involved these processes. *Borrelia hermsii*, a bacterial spirochete that causes relapsing fever, expresses only one variable major protein (Vmp) from a repertoire of genes at the outer surface [329]. The switch of Vmps occurs by gene conversion, specifically a silent promoter-less *vmp* gene cassette is recombined into the single vmp expression site near the telomere of a linear plasmid [330, 331]. *Plasmodium falciparum*, the causative agent of malaria, uses antigenic variation of *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) during blood stage to maintain a chronic infection [332]. For this, only 1 out of 60 *var* genes is expressed at a time and during each red blood cell invasion a epigenetic-dependent switch to another *var* gene may occur.

All *T. brucei* subspecies undergo antigenic variation of the VSG to evade the host immune system and establish persistent infections [333]. Other trypanosome species like *T. congolense*, *T.vivax* and *T. equiperdum* are also able to undergo antigenic variation using VSGs [334, 335]. Nonetheless, *T. cruzi*, the causative agent of Chagas disease, is an intracellular parasite which does not use antigenic variation although expressing variable surface proteins [336]. However, it was recently shown that *in vivo* antigenic variation is very complex with several variants of VSGs present in a single host to maintain chronicity [337]. In *T. brucei* antigenic variation relies on two main features, monoallelic expression of VSG and periodic switching of VSG, which will now be addressed.

Monoallelic expression

Monoallelic gene expression or allelic exclusion refers to the process by which only one allele of a given gene (or gene family) is expressed. One of the most well studied examples of allelic exclusion is the expression of olfactory receptor (OR) genes in the olfactory sensory neurons (OSNs). Each OSN expresses a single allele of a single gene out of a family of ~1100 (in mice) or ~630 (in humans) *OR* genes [338]. In an immature OSN, *OR* gene clusters are marked with the constitutive heretochromatin marks H3K9me3 and H4K20me3 [339]. Histone demethylase LSD1 and additional unidentified chromatin remodellers remove the histone modifications of only one *OR* gene, allowing its derepression and activation [340]. Transcription and expression of a non-functional or mutated *OR* leads to switching to another *OR* transcription until translation of a functional OR elicits a feedback based mechanism to stably express only that *OR* [341]. This feedback mechanism starts by activation of the unfolded protein response (UPR), which ultimately increases translation of ATF5 [342]. ATF5 will induce the transcription of adenylyl cyclase 3 that relieves the UPR and downregulates LSD1, preventing other *ORs* from being expressed. This processes will ultimately establish expression of only one *OR* and trigger the differentiation to mature OSN [340].

Monoallelic expression of *var* genes in *Plasmodium falciparum* leads to the expression of a single PfEMP1 protein, the surface receptor encoded by a *var* gene. However, this does not strictly mean that PfEMP1, needs to be produced. In fact, replacing a *var* gene by a drug-selectable marker abolishes expression of PfEMP1 but parasites are viable and switching may still occur in the absence of drug pressure [343]. Thus, upstream elements in the promoter are the "counting" mechanism that controls the transcription of a single *var* gene locus. All *var* genes, except the active one, in heterochromatic perinuclear repressive centers enriched in H3K9me3 [344]. The active *var* gene is also positioned in the perinuclear region but segregated from the silent *var* genes and the promoter is enriched in H3K4me2/3, H3K9ac and H2AZ which are permissive marks for transcription [332]. Within the red blood cells and prior to DNA replication, the active *var* gene promoter becomes poised and enriched in H3K4me2 and only becomes active after new invasion of red blood cells [345]. This is accomplished by the histone methyltransferase PfSET10, which might serve as a mark for inheritable

expression of the same *var* gene [346]. Nonetheless, *Plasmodium* can switch the expression of the *var* genes although it is not understood in which stage of the cell cycle this is accomplished.

T. brucei also ensures monoallelic expression by expressing a single *VSG* gene amongst more than 2500 genes or pseudogenes. However, the control of monoallelic expression does not imply that only one VSG is stably expressed. When a second VSG gene is inserted into the active BES, both VSGs are expressed and homogenously distributed at the cell surface in equal amounts [347]. On the other hand, when parasites were forced to simultaneously transcribe two tagged BESs using drug pressure, rapid switching between the two BES occurred instead of obtaining a stable double expressor [348]. These studies suggest that *T. brucei* has a mechanism to avoid multiple VSGs from being simultaneously expressed in natural conditions at the cell surface. Recently, VEX1 protein is both a negative and a positive regulator of monoallelic expression, being sequestered in the active BES to ensure monoallelic expression in a "winner takes all" process [270]. RNAi against VEX1 leads to loss of silent BES silencing in a homology sequence-dependent manner, while overexpression of VEX1 leads to its accumulation in silent BES leading in both cases to upregulation of silent VSGs and expression of several VSGs at the cell surface. On the other hand, RNAi against the VSG mRNA leads to a precytokinesis cell-cycle arrest [349].

VSG switching

To evade the host's immune system and maintain a chronic infection, *T. brucei* changes the dense coat of homodimeric VSGs, which is believed to be the only parasite antigen against which an immune response is mounted. The two mechanisms by which the expressed VSG is switched are switching by recombination and the epigenetic BES transcriptional *in situ* switching (Figure 7). Several types of recombination may occur: exchange of sequences between the active and a silent BES, duplication of VSG genes from the internal silent arrays or minichromosomes to replace the active VSG and assembly of fragments of VSG pseudogenes into the active VSG to create mosaic VSGs. In terms of activation of these routes, *in situ* switching and BES exchanges arise early in infection, being replaced by the usages of the archival VSGs from the internal arrays and

minichromosomes and only late in infection mosaic VSGs are assembled [337, 350]. I will briefly revisit recombinational switching mechanisms while transcriptional switching will be more detailed.

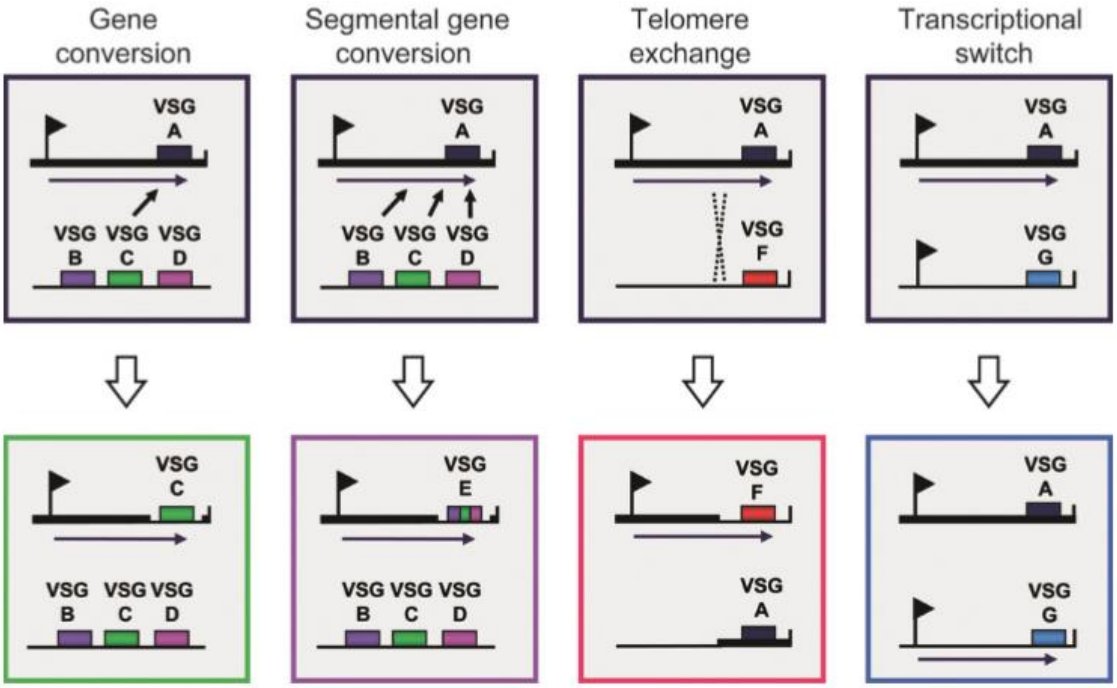


Figure 7 – Different mechanisms for VSG switching in *T. brucei*. Gene conversion is based on the duplication of a VSG from the internal arrays or minichromosomes in order to replace the VSG in the active BES. Segmental gene conversion generates mosaic VSGs from fragments of different VSG genes or pseudogenes that will ultimately replace the VSG in the active BES. Telomere exchange results in DNA sequence cross-over from an active and a silent BES, which can span from recombining only the VSG gene to the whole BES. Transcriptional switch consists in transcriptionally inactivating the active BES, with the concomitant activation of a previously silent BES. (Image taken from [351])

Switching by recombination

Recombinational switching is dependent of DNA DSBs, which is facilitated by the fact that subtelomeric regions (where BESs are located) are hotspots for recombination processes. Homologous recombination (HR) is a pathway to repair DSB with an important role towards VSG switching. This pathway is dependent of RAD51, an enzyme that, when inactivated, results in impaired VSG switching [352]. BRCA2 protein which interacts with RAD51 also impairs HR [353]. Break-induced replication (BIR) is a HR mechanism that is used for exchange of sequences between the active and a silent BES (telomere exchange

in Figure 7) [7]. BIR is frequently used when a DSB is generated in the 70-bp repeat upstream of the VSG gene, allowing translocation of the VSG gene together with the telomeric repeats. Extended regions even until the BES promoter may also be translocated, although this event is somewhat rare [6]. However, the major HR mechanism is gene conversion [354]. In gene conversion, the archival intact VSGs are duplicated together with the flanking 70-bp repeats and a second recombination junction is generated in the VSG 3'UTRs contain where highly conserved elements may serve as recombination locus [5]. This allows translocation of a new VSG into the active BES but maintaining telomeric sequences. Segmental gene conversion is termed for the generation of mosaic VSGs since only fragments of silent pseudogenes are recombined to yield new VSG genes [355, 356]. This ability to generate mosaic VSGs allows to exponentially increase the potential VSG archive size turning it in the core of antigenic variation and chronic infections [357]. Besides HR, the microhomology mediated end-joining, a non-RAD51 dependent mechanism, can also contribute for recombinational switching possibly through the 70-bp repeats [358].

DSBs in the active BES were shown to be natural intermediates of VSG gene conversion [359], which may be a consequence of replication fork collapse or the fragile nature of subtelomeric regions [358, 360]. However, the probability and mechanism of switching employed after a DSB may depend on the break site: a promoter-adjacent break does not lead to switching, a 70-bp repeat associated break frequently results in a RAD51-dependent gene conversion with a G_2/M cell-cycle arrest while a break in telomeric sequences bypasses the G_2/M cell-cycle checkpoint and consequently leads to a loss or replacement of the entire BES [358]. Nonetheless, it was recently proposed that it is not a self-induced DSB that initiates a switching event but rather a replicating-dependent process where collisions of transcription and replicating machineries create an instable environment that is favorable to induce switching [361]. Also, a probabilistic order in terms of the hierarchy of VSG activation and replacement has been shown to influence recombination to maintain chronic infections [362, 363].

The length of the telomeric repeats has also been shown to work as a cue to induce VSG switching. Thus, in telomerase reverse transcriptase (TERT) null mutants,

which are not able to maintain extend telomeric repeats, the shortening of telomeric sequence to very few repeats leads to VSG gene conversion [257, 259]

***In situ* switching**

BES transcriptional switching (or *in situ* switching) consists in silencing the active BES with the concomitant activation of a silent BES, thus maintaining monoallelic expression [8]. Because chromatin of the active and silent BESs is dramatically different, BES switching involves important modifications in chromatin structure, which are likely very well coordinated with changes in the transcriptional machinery.

All of the factors important for this type of switching also play a role in VSG monoallelic expression consistent with a interdependent control of both processes: cohesin represses *in situ* switching without affecting recombination-mediated switching [265]; ORC1 represses preferentially BES transcriptional switching [267]; NUP1 was proposed to keep silent BESs away from the ESB therefore preventing switching events [263]; DOT1B is considered a central factor in promoting fast *in situ* switching events since its depletion causes cells to express more than one VSG at the cell surface up to several weeks until VSG coat replacement is completed [245]. Moreover, inositol phosphate pathway regulates switching apparently through a telomere-dependent manner, since PIP5K represses both recombination-dependent and *in situ* switching [264]. How these factors act together to allow a fast and well-coordinated transcriptional switch remains totally unknown.

When two different BES marked with drug resistance genes were forced to be simultaneously expressed, rapid and permanent back and forth switches between the two BES were apparently prompted to guarantee drug resistance [348]. No parasites were obtained that simultaneously expressed two BESs, suggesting the presence of a mechanism that prevents such situation from happening and thus ensuring VSG monoallelic expression. The authors postulated the existence of natural unstable switching intermediates in which silent BESs are in a pre-active state that facilitates a transcriptional interchange.

The co-transposed region (CTR), a non-coding sequence located between the 70-bp repeats and the VSG gene, gave insight on BES stability. Deletion of the CTR led to an increase of more than 100-fold in BES switching frequency which became more severe if the start codon of the active VSG is also deleted together with the CTR [364]. Recently, a study showed that the ectopic expression of a second *VSG* from a non-BES locus promotes attenuation of transcription from the whole active BES [365]. This *VSG* silencing/attenuation is reversible and mediated by DOT1B. This study proposes that activation of a silent VSG may act as a trigger for transcriptional switching, suggesting that transcriptional attenuation is an intermediate state that allows probing for the integrity and functionally of a new BES before full commitment for switching. However, in light of current knowledge one cannot define what is the true trigger that drives switching.

An *in situ* switching is much poorly characterized compared to recombinational switching. Nonetheless, the former mechanism cannot be disregarded as it assumes high importance in initial stages of the infection. In this dissertation, I propose to unveil the role of chromatin conformation of BESs not only in the context of an *in situ* switching, but also in the control of monoallelic VSG expression.

THESIS AIMS

Antigenic variation is a powerful mechanism used by several pathogens to evade the host immune system [351]. The extracellular parasite, *Trypanosoma brucei,* periodically changes the variant surface glycoproteins (VSGs) present at the cell surface to maintain an infection for long periods of time [97]. Each parasite expresses a single VSG gene from one of ~15 bloodstream expression sites (BES), which are transcribed by RNA polymerase I (Pol I) [4, 93].

While the majority of the pathogens change the exposed surface proteins by employing recombination-based processes, *T. brucei* parasites use not only recombination-based but also epigenetic processes [351]. During an *in situ* switch, the active BES is silenced and a previously silent BES is activated, resulting in the expression of a new VSG [8]. Work from Figueiredo and other labs had previously shown that the highly transcribed active BES possesses a very open and nucleosome depleted chromatin structure while silent BES possesses a compact nucleosome enriched chromatin [9, 10].

These observations led me to ask how chromatin and transcriptional changes intertwine during an *in situ* switching. Since switching is rare and very difficult to observe, we used a reporter system that greatly induces the frequency of *in situ* switching [12].

At the onset of this work we established a number of specific aims. First, to generate and to characterize a BES transcriptional silencing inducible system with a transcriptional reporter. Second, to observe how the chromatin behaves upon inducing a BES transcriptional silencing. Finally, to identify and characterize the protein component(s) that mediate the interplay between chromatin and transcription during VSG *in situ* switching.

This work relied on the use several methodologies ranging from established *in vitro* assays, generation of parasite transgenic cell-lines, double-stranded RNA (dsRNA) assay, formaldehyde-assisted isolation of regulatory elements (FAIRE), chromatin immunoprecipitation (ChIP), western blotting and flow cytometry were employed.

With this project, we expect to have a better understanding of the regulation of an *in situ* switching, which could ultimately lead to the identification of intervention strategies that may block this mechanism and impair the progression of the infection.

CHAPTER 2

A TRANSCRIPTION-INDEPENDENT EPIGENETIC MECHANISM IS ASSOCIATED WITH ANTIGENIC SWITCHING IN *TRYPANOSOMA BRUCEI*

Francisco Aresta-Branco[1], Sílvia Pimenta[1], Luísa M. Figueiredo[1]

[1] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal

Author Contributions

FAB and LMF designed the experiments. FAB performed all the experiments. SP helped in the construction of GLB1. FAB wrote the chapter. LMF supervised the work and revised the text.

INTRODUCTION

RNA Polymerase I (Pol I) transcribes ribosomal DNA genes (rDNA), which account for over 60% of total nuclear transcription [163]. Most organisms have large tandem arrays of rDNA genes, but only a fraction is transcriptionally active. Consistent with their transcriptional activity, rDNA genes can be found in one of two chromatin states: a compact nucleosome-rich "closed state" and an accessible nucleosome-depleted chromatin "open state" [195]. In yeast, replication of DNA converts the chromatin of most rDNA genes into the closed state. When replication is complete and transcription is reinitiated, a stochastic fraction of rDNA genes regains the open state, in a process that is dependent on Pol I. Once chromatin has an open state, its status is maintained by a high-mobility box protein, HMO1, independent from Pol I [221].

Trypanosoma brucei, a unicellular parasite responsible for sleeping sickness, is an unusual eukaryote that also uses Pol I to transcribe genes that encode two classes of abundant surface proteins, the Variant Surface Glycoprotein (VSG) and Procyclins [4, 366]. Periodic exchange of the exposed VSG allows the parasites to evade the host immune system, a process known as antigenic variation. Although the *T. brucei* genome has more than 2000 VSG genes [91], only one is transcribed at any given time. *VSG*s are transcribed from Bloodstream Expression Sites (BESs), specialized polycistronic units in which a Pol I promoter drives transcription towards the telomere. Between the promoter and the telomere there is a variable number of expression-site-associated genes (ESAGs) followed by an array of tandem 70-bp repeats that precede the telomere-proximal *VSG* [93]. Among the approximately 15 BESs present in the genome, only one is functionally active, ensuring the monoallelic expression that is the heart of antigenic variation.

VSG switching can happen by two main mechanisms, by recombining a new VSG into the active BES or by switching off a BES and activating another one (*in situ* or transcriptional switch) [367]. VSG transcriptional switching involves two BESs, one that is silenced and another that is concomitantly activated. Cross-talk among BESs has been proposed to explain the phenotype observed when two BESs were simultaneously selected with drug selectable markers [348]. Davies *et al*. also detected that deletion of a non-coding DNA sequence upstream of the active *VSG*, increased switching frequency to a new BES [364]. More recently, Batram *et al.* showed that upon overexpression of an

exogenous VSG, the active BES is partially attenuated into an intermediate stage, which may allow probing of silent BESs before commitment to one [368].

While in the mammalian-infective bloodstream stage, the chromatin of the active and silent BESs is dramatically different. The actively transcribed BES is nucleosome-depleted (open state), while silent BESs are organized in regularly spaced nucleosomes [9, 10]. The promoters of both active and silent BESs are bound by the multi-subunit class I transcription factor A (CITFA) [233], although to a different extent [236], which is necessary to regulate BES transcription initiation. Several epigenetic factors have been shown to be necessary to prevent transcription from silent BESs [369]; others are necessary to ensure fast switching between BESs [245], but only TDP1 has been shown to be a core component of the active BES. TDP1 is a high mobility group box protein that is present in the chromatin of active BES, and at rDNA genes, and it is necessary for their transcription [13, 14]. The interplay among these factors and the mechanisms by which they affect *VSG* transcription are essentially unknown.

During the life cycle of *T. brucei*, parasites shift between the mammalian host and the Tsetse fly [286]. To survive in such different hosts, parasites undergo significant changes in gene expression, which include replacement of the VSG by the similarly abundant procyclins [287]. Silencing of the active BES during differentiation is characterized by a progressive downregulation of transcription along the BES [288], together with a re-localization of the BES to the heterochromatic nuclear periphery [11]. It has also been shown that the chromatin of the originally active BES becomes less open to T7 polymerase, suggesting that it acquires a more compact structure during differentiation [289].

The goal of this chapter was to understand the interplay between transcription and chromatin during antigenic variation. We characterized the early events that take place when the active BES is silenced. We found that trypanosomes have a cell-cycle and transcription-independent mechanism to maintain the open chromatin conformation of the active BES, which is dependent upon TDP1. We also observed that, in the first two days after transcriptional silencing is induced, parasites experience an intermediate stage in which several previously silent BESs are temporarily transcribed at higher levels,

suggesting that cells probe different BESs before commitment to a new single BES. Our findings provide evidence that regulating chromatin conformation is tightly associated to antigenic switching, an important virulence mechanism of this pathogen.

MATERIALS AND METHODS

Trypanosome cell-lines and plasmid construction

T. brucei bloodstream form (BSF) parasites (strain Lister 427, antigenic type MiTat 1.2, clone 221a) [370] were cultured in HMI-11 as described in [371]. PL1A cell-line was described in [258] and grown with 2.5 μg/ml of G418 and 1 μg/ml of Puromycin (Invivogen). All transfections were made with a nucleofector (AMAXA), program X-001, using the previously optimized homemade Tb-BSF buffer (90 mM Na_2HPO_4, 5 mM KCl, 0.15 mM CaCl2, 50 mM HEPES, pH 7.3) [372]. GLB1 and its derivative cell-lines were modified from the parental 2T1.ESPiGFP:NPT [12] and grown with 2.5 μg/ml of Phleomycin, 0.1 μg/ml of Puromycin, 10 μg/ml of Blasticidin-S (Invivogen) and 1 μg/ml of Tetracycline (Fisher Scientific). The *NPT* gene in the 2T1.ESPiGFP:NPT cell-line was replaced by luciferase reporter by transfecting a PCR product containing a stop codon for the *GFP* gene, an *Aldolase* 3' UTR, *Luciferase* and *BSR* genes (previously cloned in pFAB2). PCR was performed using primers with long tails (Table 4) which serve as target recombination regions for *GFP* ORF (open reading frame) and the BES sequence downstream of *GFP::NPT* 3' UTR.

GLB1-TDP1::TY1 was generated by transfecting pFAB11, which inserts a TY1 tag in the 3' end of one of the *TDP1* endogenous alleles. pFAB11 contains a (i) *Hygromycin resistance* gene, (ii) a 5'-end truncated *TDP1*, (iii) a TY1 tag in the 3'-end of *TDP1*, (iv) an *Aldolase* 3'UTR. pFAB11 was linearized with *Sma*I (New England Biolabs), which digests in the middle of *TDP1* ORF. GLB1-TDP1::3xcMyc was generated by transfecting pFAB14 which inserts a triple MYC tag in the 3' end of one of the *TDP1* endogenous alleles. pFAB14 contains a (i) *Hygromycin resistance* gene, (ii) a 5'-end truncated *TDP1*, (iii) a triple MYC tag in the 3'-end of *TDP1*, (iv) the annotated endogenous 3'UTR of *TDP1*. pFAB14 was linearized with *Sma*I (New England Biolabs), which digests in the middle of *TDP1* ORF.

GLB1-R15 was obtained by first generating switchers of GLB1 using the BES silencing inducible assay and subsequently transfecting pFAB17 in switchers. pFAB17 contains a (i) a 405 nucleotide sequence upstream BES promoter, (ii) BES promoter, (iii) an *RFP* ORF lacking NLS, PEST and stop codon amplified from pCAGGS, (iv) a *NPT* ORF

lacking the start codon, (v) an *Actin* 3'UTR, (vi) a 475 nucleotide sequence downstream BES promoter. *RFP* and *NPT* ORFs formed the fused *RFP::NPT* gene. pFAB17 was digested with *Asc*I and *Nde*I (New England Biolabs) prior to transfection. Selection of clones was performed with 25 μg/ml of G418. Cloning of VSG transcripts and sequencing allowed identification of the BES in which *RFP::NPT* had integrated. Transfected switchers were reverted to the original BES by removing G418 and adding 1 μg/ml of Tetracycline to the medium. Afterwards, serial dilutions were performed and 10 μg/ml of Blasticidin-S was added to each dilution to select revertants.

All cloning was performed using the In-Fusion® HD Cloning system (Clontech) following to the manufacturer's instructions.

Differentiation assay

Parasites were collected at a density of 1 – 1.5 × 10^6 cells/ml and centrifuged at 650 g for 10 min at room temperature. Cells were resuspended in DTM medium with 6 mM of *cis*-aconitate at a density of 1.5 – 2 × 10^6 cells/ml and grown at 27ºC without CO_2.

Transcript quantification

Parasites were harvested by centrifugation at 650 g for 10 min, 4°C and immediately resuspended in PureZOL (BioRad) or TRIzol (Invitrogen). RNA was isolated following the manufacturer's instructions and RNA quantity and quality was assessed on a NanoDrop 2000 (Thermo Scientific). cDNA was generated using a Superscript cDNA Synthesis Kit (Invitrogen), according to manufacturer's protocol. Quantitative PCR (qPCR) was performed using 1× SYBR Green PCR Master Mix (Applied Biosystems). Negative controls lacking reverse transcriptase (RT-) were confirmed by quantitative PCR (qPCR). Amplification reactions were performed in duplicates. The ΔΔCt method as used to determine transcript levels relative to normalizing gene.

Tetracycline-inducible BES silencing assay

Parasites were centrifuged at 650 g and washed three times with warm HMI-11. Pellets were resuspended in medium with drugs except Blasticidin-S and tetracycline and density was adjusted to 0.5 × 10^6 cells/ml. Cells were split in two flasks with and without Blasticidin-S and tetracycline (Tet+ and Tet-, respectively).

Survival frequency and commitment assay

To determine the percentage of cells that survive the BES-silencing assay, after washing away drugs, cells were diluted to a density of 10 cells/ml and plated in two 96-well plates with or without tetracycline (Tet+ and Tet-, respectively). Seven days after plating, surviving clones were counted in Tet- and Tet+ plates and its ratio yielded the survival frequency. For commitment assay, cell density of Tet- cultures was determined at 8, 24, 48 and 72 hours after BES silencing and subsequently diluted to a density of 10 or 50 cells/ml containing 1 μg/ml of tetracycline. Dilutions were performed with the same drugs and complemented with 1 μg/ml of tetracycline. Diluted cells were then plated in two 96-well plates. All plates containing tetracycline were replenished with fresh tetracycline three days after plating to maintain excess concentration. Six days later, around 20 surviving wells were passaged to new 96-well plates and were analyzed by a FACS High Throughput Sampler (BD Biosciences) to score for expression of GFP.

Cell-cycle profile after BES silencing

At each time point after inducing BES silencing, 2 million cells were centrifuged for 10 min at 1300 g, 4ºC and washed once with ice-cold PBS. Cells were resuspended in PBS with 2 mM EDTA and slowly fixed with 2.5 ml of absolute ethanol. After fixing for at least one hour, cells were washed once and resuspended in 1 ml PBS / EDTA. Cells were incubated with 10 μg RNase A and 1 μg of propidium iodide for 30 min at 37ºC and further analyzed by flow cytometry for DNA content.

Luciferase assay

1.5×10^6 cells were harvested for 5 min at 2800 g, 4ºC and washed once with 1 ml cold TDB. Pellets were resuspended in Lysis Buffer (Biotium) and protocol was followed according to the manufacturer's instructions. Luminescence was measured by a microplate reader (Tecan).

FAIRE

Formaldehyde-assisted isolation of regulatory elements (FAIRE) was performed as described previously [238]. Briefly, 20 million cells were centrifuged for 20 min at 3000 g, resuspended in 20 ml of HMI-11 and crosslinked with 1.1% Formaldehyde (Sigma) for 10

min. In cell-cycle sorted fractions, protocol was performed with only 1-2 million cells. Crosslinking was stopped with 0.125 M of Glycine (Sigma) and cells were washed once with Trypanosome Dilution Buffer (TDB) (5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$, 20 mM Na_2HPO_4, 2 mM NaH_2PO_4, 20 mM glucose, pH 7.4). Cell pellets were lysed with Lysis Buffer (50 mM TrisHCl, 10 mM EDTA, 1% SDS) complemented with Protease Inhibitor Cocktail (Sigma) and Solution P (PMSF, Pepstatin). Lysates were incubated on ice for 15 min and an external plasmid DNA spike was added before sonicating with 10 cycles of 30 sec on / off. Lysates were centrifuged at 16000 g, 4ºC to pellet cell debris and a 100 μl aliquot of the supernatant was taken to check sonication efficiency. The remaining fraction was subjected to two Phenol-Chloroform-Isoamyl Acid (25:24:1) (Sigma). DNA present in the aqueous phase, corresponding to open chromatin, was precipitated and washed with 2 volumes of absolute ethanol, sodium acetate 0.3 M and 20 μg/ml glycogen overnight at -20ºC DNA was washed once with ethanol 70% and dried before resuspension in Elution Buffer (Quiagen). Samples were then treated with 100 μg/ml of RNase A (ROTH) and purified with a PCR purification kit (Quiagen). A non-crosslinked sample was always included to normalize qPCR data for gene copy number. Quantification of the FAIRE and total DNA samples were analyzed by qPCR. Amplification reactions were performed in duplicates.

Chromatin Immunoprecipitation

ChIP was carried out essentially as described elsewhere [9] but with several modifications. A total of 3-5 × 10^7 cells were harvested for 10 min at 3000 g and fixed for 20 min in 1% formaldehyde solution (50 mM HEPES-KOH pH7.5, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA and 11% formaldehyde) in HMI-9 medium. Crosslinking was stopped with 0.125 M of Glycine and cells were washed once with PBS. Cells were first lysed with 10 ml lysis buffer 1 (50 mM HEPES-KOH pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% Tergitol, 0.25% Triton X-100 plus protease inhibitor cocktail) and incubated for 10 min at 4ºC. Upon spinning for 20 min at 4000 g, 4ºC, the pellet was lysed with 10 ml of lysis buffer 2 (10 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA plus protease inhibitor cocktail) for 10 min at room temperature. After spinning for 20 min at 4000 g, 4ºC, the pellet was lysed with 2 ml of lysis buffer 3 (10 mM Tris-HCl pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine

plus protease inhibitor cocktail). DNA was sonicated in a Bioruptor UCD-200 (DIAGENODE) for 10 min total (30-s on and off cycles). After sonication, 200 μl of 10% Triton X-100 was added, samples were split into two microcentrifuge tubes and centrifuged for 10 min at 16100 g at 4ºC. 50 μl were taken for the input sample and 100 μl to assess sonication efficiency. To the remaining 2 ml, 0.7 ml of lysis buffer 3, 0.3 ml of 10% Triton X-100 plus protease inhibitor cocktail was added. Lysate was incubated with Dynabeads Protein G (Life Technologies) combined with 10 μg of rabbit anti-H3 antibody (kind gift from Christian Janzen) or mouse anti-TY1 antibody (clone BB2, The Rockefeller University) overnight at 4°C. Immunoprecipitated material was washed seven times with cold RIPA buffer (50 mM HEPES-KOH, 500 mM LiCl, 1 mM EDTA, 1% NP-40 and 0.7% Na-Deoxycholate) and once with 10 mM Tris-HCl pH8.0, 1 mM EDTA, 50 mM NaCl. Material was reverse crosslinked with 200 μl Elution Buffer (50 mM Tris-HCl pH 8.0, 10 mM EDTA, 1% SDS) for 12-14 hr at 65ºC. Samples were treated with 80 μg of RNase A for 2 hr at 37ºC and then with 80 μg of Proteinase K for 2 hr at 55ºC before being purified using a Gel Extraction Kit (Qiagen). Immunoprecipitated material was quantified by qPCR. Amplification reactions were performed in duplicates.

Cell-cycle sorting

10^8 cells were centrifuged for 10 min at 650 g, resuspended in 20 ml of HMI-11 and crosslinked with 1.1% formaldehyde (Sigma) for 10 min. Crosslinking was stopped with 0.125 M of Glycine (Sigma) and cells were washed once with PBS and resuspended in 1 ml PBS / 2 mM EDTA. Cells were permeabilized for 5 min with 40 mM digitonin, washed twice with PBS and resuspended in 1 ml PBS / 2 mM EDTA. Staining was performed with 1 μl FxCycle Violet Stain (Molecular Probes) for 30 min protected from light. Cells were sorted according to their DNA content using a FACS Aria sorter (BD Biosciences) and collected in PBS.

dsRNA production

dsRNA against TDP1 were made using the MEGAscript RNAi Kit (Ambion) and protocol was followed according to the manufacturer's instructions. The sequence used to deplete TDP1 mRNA is identical to the one published in [14]. 20 μg of dsRNA or the same buffer volume (for the mock control) were transfected into 100 million cells, using the X-001 program, and the homemade Tb-BSF buffer in an AMAXA nucleofector. The

primers used for amplification of DNA template include a T7 promoter for *in vitro* transcription and are presented in Table 4.

Table 4 - List of primers used in the present chapter.

Primer name	Primer sequence (5' ->3') forward / reverse
TDP1 dsRNA template	TAATACGACTCACTATAGGGAGCGGCTATAGATGACATTGTTG / TAATACGACTCACTATAGGGAGCTCGTAGACCTTCCTCTCC
Long primers to generate GLB1	AGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAGGTGCCTGAGGATCCTGCCC / CGACATAAATCGAATACGGGCACTCCGGTAACGACGGCACGCTCAGGCCAACTCACGATCGGTTCGAGACGTCCGCCACACCAGCTTGCATAGATAACAAACGCATC
Tb927.10.12970_qPCR	CCAGCCTTCTCAATCTCCAG / GGCCACAGTTGGATAGCTTG
Luciferase_qPCR	ATGTCCGTTCGGTTGGCAG / CATACTGTTGAGCAATTCACG
BSR_qPCR	CGGCTACAATCAACAGCATC / ACGATACAAGTCAGGTTGCC
VSG2_qPCR	AGCAGCCAAGAGGTAACAGC / CAACTGCAGCTTGCAAGGAA
VSG9_qPCR	ACTAAGCTCGTGGCGCAC / CGCGTAGTTGACGCATGAC
VSG3_qPCR	GCTTATTTTGTGTCTGTCGC / GACGCAGCAGAATCAACAC
VSG13_qPCR	ATAACGCATGGCCATCTTGAC / GTCGTTGCTGTGGATTGCTC
18S_qPCR	ACGGAATGGCACCACAAGAC / GTCCGTTGACGGAATCAACC
rDNA spacer_qPCR	GGTTACTCGTGGGACTAGG / TGCCTTGGCCCTGATGGC
β-tubulin_qPCR	TTCAGGCTGGCCAATGCG / TACGGAGTCCATTGTACCTG
GAPDH_qPCR	AGATTGATGTCGTTGCTGTTGTG / ATGGCTTGCTCTTCGTAGTCG
NPT_qPCR	CTTGCCGAATATCATGGTGGA / ACCGTAAAGCACGAGGAAGC
Ampicillin_FAIRE DNA spike_qPCR	ATCGTTGGGAACCGGAGC / AGCGCAGAAGTGGTCCTG

Western Blotting

Cells were lysed using Laemmli buffer, resuspended at 2×10^5 cells/μl and treated with 200 U/ml Benzonase (Sigma). A SDS-PAGE was performed at 4ºC and proteins were transferred to a nitrocellulose membrane using an iBlot Dry Blotting System (Invitrogen)

for 6 min and 30 sec. Membrane was blocked with 5% milk in PBS / 0.1% Tween for at least 1 hour and primary mouse anti-cMyc antibody (clone 9E10, The Rockefeller University) was incubated overnight at 4ºC in a 1:1000 dilution in 3% milk in PBS / 0.1% Tween. Membrane was washed three times with PBS / 0.1% Tween, and incubated with anti-mouse horseradish peroxidase-linked secondary antibody (Amersham) at room temperature for 1 hr in a 1:20000 dilution in 3% milk in PBS / 0.1% Tween. Membrane was further washed three times with PBS / 0.1% Tween and developed with a Western Lightning Plus-ECL (PerkinElmer). Picture acquisition was made by a Chemidoc XRS+ (Bio-Rad). For loading control histone H2A, membrane was stripped with the Restore PLUS Western Blot Stripping Buffer (Thermo Scientific) according to the manufacturer's instructions and blocked again with 5% milk in PBS / 0.1% Tween for at least 1 hr. Primary rabbit anti-H2A antibody (custom made) was incubated for 1 hr at room temperature in a 1:5000 dilution in 3% milk in PBS / 0.1% Tween. The remaining protocol was followed as above except for the secondary antibody which was an anti-rabbit horseradish peroxidase-linked antibody (Amersham).

Quantification was performed using the Image Lab software (Bio-Rad). Background intensity was subtracted from each band and intensity of TDP1::cMyc band was normalized to the loading control and to mock transfected control.

VSG staining of live cells

0.5×10^6 cells were harvested for 5 min at 2800 g, 4ºC. Cells were resuspended in 50µl of cold HMI-11 in which Alexa Fluor 647 anti-VSG13 conjugated antibody had been previously diluted (1:5000). After 15 min of incubation at 4ºC with gentle shaking, cells were washed three times in cold TDB, resuspended in cold TDB, and immediately analyzed on a FACS Fortessa (Becton Dickinson Biosciences). Data were processed with FlowJo software (FlowJo, LLC).

RESULTS

Transcription and chromatin dynamics during differentiation

When a bloodstream parasite undergoes transcriptional VSG switching or when it differentiates into procyclic forms, the active BES must be silenced and its chromatin closed [9, 10, 289]. In both processes, the interplay between transcription and chromatin is unknown. What is the order of these events: does chromatin condensation lead to a transcriptional silencing, or does transcription stop first and chromatin is condensed later? We began to tackle this problem by characterizing what happens during differentiation (Figure 8), and then in VSG switching (Figure 9).

To study the interplay between transcriptional silencing and chromatin condensation in the active BES during differentiation, we induced cell differentiation *in vitro* by adding the chemical trigger *cis*-aconitate and lowering the temperature from 37ºC to 27ºC. We measured levels of mRNA at different time points, using a cell-line in which BES1 is actively transcribed (Figure 8A). As expected, *VSG2* transcript levels rapidly decreased, reaching 2% of the bloodstream form levels, 24 hr after inducing differentiation (Figure 8B). Ribosomal DNA (rDNA) *18S* transcripts remained unchanged, as expected [288], and transcript levels of silent VSG9 were slightly reduced. As previously observed, transcript levels of RNA Polymerase II (Pol II)-transcribed genes, *β-tubulin* and *GAPDH*, also decreased [50, 288, 373], the second reflecting the reduced metabolic dependence of glycolysis in procyclic forms.

To test if this transcriptional silencing of the active BES was a consequence of chromatin conformational changes, we assessed nucleosome occupancy during differentiation by FAIRE [374] (Figure 8C). This technique allows the purification and quantification of DNA with low protein content from what we call a more “open” chromatin conformation [9, 238]. We observed that, throughout the first 24 hr of differentiation, chromatin of silent *VSG9*, *β-tubulin* and *GAPDH* kept the same FAIRE-enrichment, suggesting that the chromatin remained equally condensed. Importantly, at BES1, although chromatin condensed slightly (2.5-fold at *VSG2*), it remained 12-fold more open than a silent *VSG*.

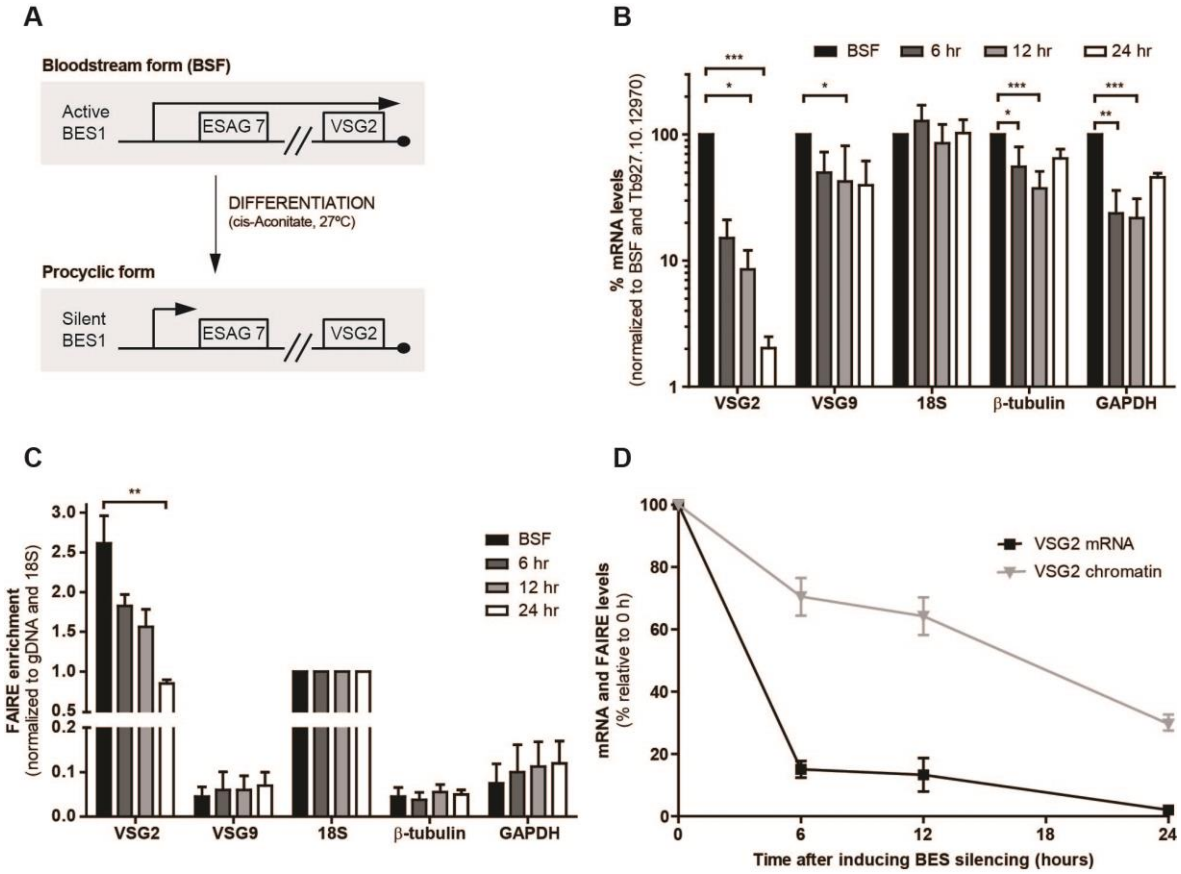


Figure 8 - BES transcriptional silencing precedes chromatin condensation during differentiation. **(A)** A bloodstream form cell-line that expressed *VSG2* (BES1) was differentiated to procyclic forms by adding cis-aconitate to the medium and changing temperature from 37ºC to 27ºC. During differentiation, BES1 is silenced. Procyclic forms do not express VSG at the surface. Quantification of mRNA levels and **(B)** chromatin conformation **(C)** 6, 12 and 24 hr after induction of differentiation. **(B)** Transcript levels were measured by qPCR and normalized to bloodstream form (BSF) levels and *Tb927.10.12970* [44], a gene previously shown to maintain constant transcript levels during differentiation. Four to six independent experiments were analyzed. **(C)** DNA purified from FAIRE was quantified by qPCR and normalized to *18S* as its transcript levels also remained constant throughout differentiation [11] and FAIRE signal was more intense than *Tb927.10.12970*. Three to five independent experiments were analyzed. **(D)** Comparison between transcript levels and FAIRE enrichment for *VSG2* gene. Values were extracted from analysis in **(B)** and **(C)**. Statistical significance was determined by 1-way ANOVA with Bonferroni post-test comparison. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

These results show that, during differentiation, the drop in mRNA levels of the active BES starts earlier and it is more pronounced than the condensation of chromatin (Figure 8D), suggesting that chromatin does not close immediately after transcription is halted.

Transcription and chromatin dynamics during BES switching

Next we hypothesized that, during BES switching, chromatin condensation may also be delayed relative to transcription silencing. We postulated that, if open chromatin of the active BES was exclusively the consequence of high transcription, induction of transcription silencing should lead to an immediate condensation of chromatin. To test this hypothesis, we used a BES-inducible silencing system to block the transcription at the active BES and we followed transcription and chromatin dynamics. Because VSG switching happens at very low frequency, we adapted a previously established reporter strain in which BES1 has a tetracycline operator sequence at the promoter followed by a GFP reporter [12]. When tetracycline is removed from the culture medium, the heterologous Tetracycline repressor is free and binds to the Tetracycline operator, thus sterically blocking Pol I transcription. We introduced a *luciferase* gene downstream of the *GFP* gene since this is a more sensitive transcriptional reporter (Figure 9A). This cell-line was named GLB1, for *GFP*, *luciferase* and *BSR* in BES1.

During the first 8 hr after tetracycline was removed from the medium (Tet-), GLB1 cells grew at the same rate as the control (Tet+), after which cell growth lagged until 48-72 hr (Figure 9B). This lag phase was characterized by abnormal cell morphology and motility (data not shown) and cell-cycle arrest in G_2/M, with a considerable accumulation of cells with polyploidy or abnormal DNA content (Figure 9C). Quantification of cell death using propidium iodide, which only stains dead cells or in late apoptosis [375] (Figure 10), showed a significant number of stained cells from 24 to 72 hr after tetracycline removal, reaching a peak of 7.5% at 48 hr. Consistent with this significant number of dead cells for several days, clonogenic assay confirmed that only 8% of the initial population of cells survive the silencing assay (Figure 9D). After 96 hr, surviving cells took over the culture and grew at normal rate of around 8 hr per population doubling. As expected, all surviving clones no longer expressed GFP, indicating they had successfully switched to a new VSG (Figure 9E).

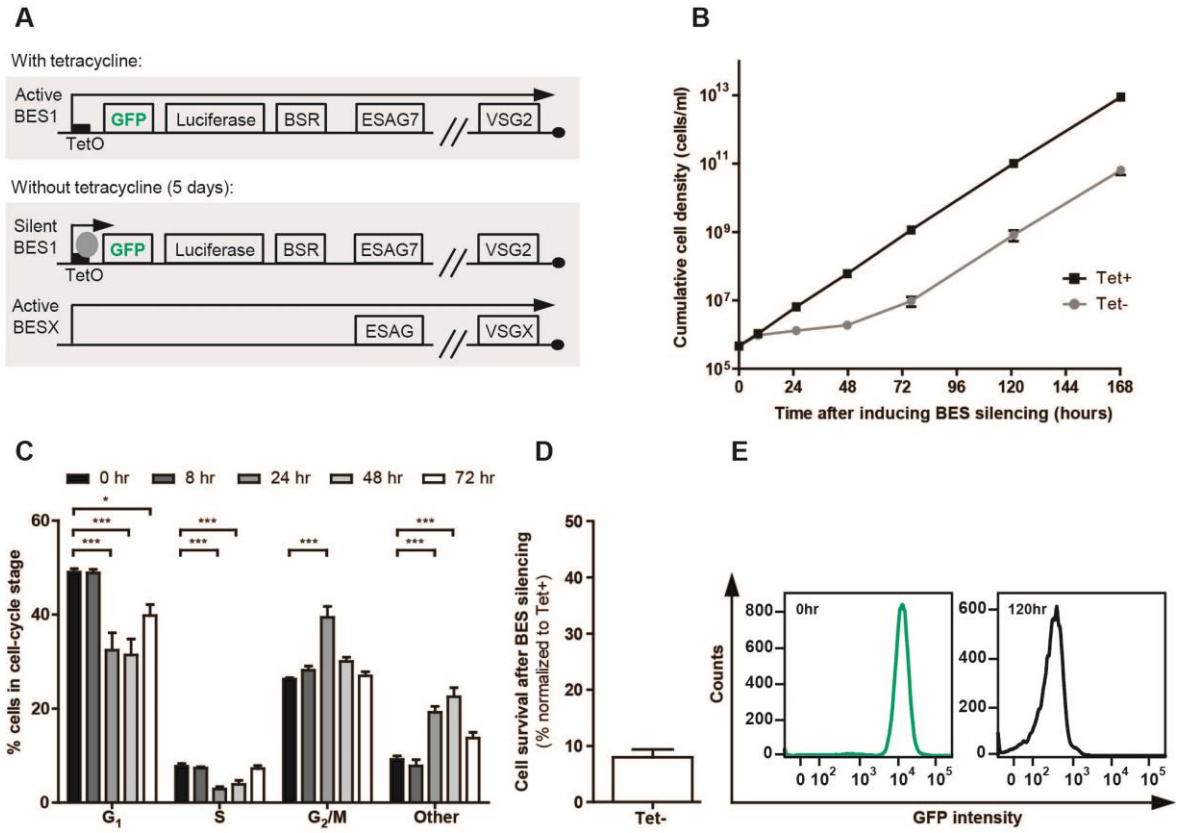


Figure 9 - BES silencing causes growth delay and G_2/M cell-cycle arrest but only after more than 8 hr of induction. (A) In the bloodstream reporter cell-line, GLB1, removing tetracycline from the medium induces BES1 silencing and subsequent activation of another BES. **Upper panel** shows that in the presence of tetracycline, BES1 is actively transcribed. BES1 contains a tet operator sequence (TetO, black rectangle), *GFP*, *luciferase* and *BSR* genes downstream of the promoter. **Lower panel** shows the outcome of tetracycline removal: BES1 becomes silent (because TetR (grey circle) binds TetO, sterically blocking Pol I transcription) and a new BES is activated. *BSR*, Blasticidin-S Resistance. **(B)** Growth curves of cells in the presence (Tet+, black curve) or after removal (Tet-, grey curve) of tetracycline. Four independent experiments were analyzed. **(C)** Cell-cycle profile of GLB1 at different time-points after removal of tetracycline. "Other" represents cells with abnormal DNA content. Four independent experiments were analyzed. Statistical significance was determined by a 2-way ANOVA with Bonferroni post-test comparison. *: $p < 0.05$; ***: $p < 0.001$. **(D)** Percentage of GLB1 cells that survive the BES silencing induction was determined by a clonogenic assay and normalized to Tet+ cells. Four independent experiments were analyzed. **(E)** Flow cytometry analysis of GFP expression of cells at 0 hr (**left panel**) and 120 hr (**right panel**) after tetracycline removal.

To test the efficiency of the steric blockade of Pol I upon tetracycline removal, we checked whether Pol I was evicted from BES1 chromatin fiber by conducting ChIP against the RPA31 subunit of Pol I (Figure 11).

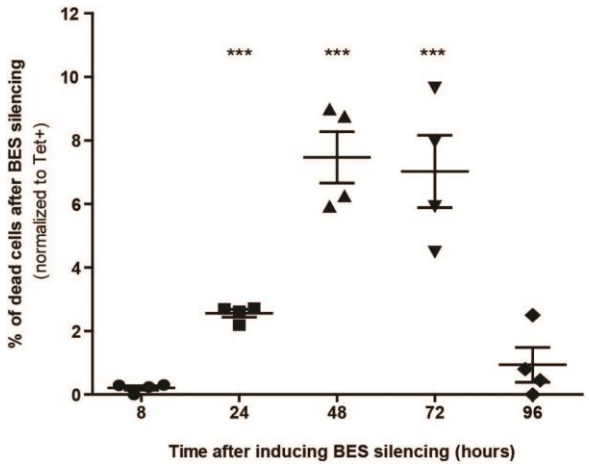


Figure 10 - Cell death during BES silencing assay. BES silencing was induced and cells were collected 8, 24, 48, 72 and 96 hr later. Induced cells were incubated with 5 µg/ml of propidium iodide (PI) prior to FACS analysis. Dead cells or in late apoptosis are stained by propidium iodide. The number of PI positive cells in 'Tet–' culture was normalized to the number of PI positive cells in 'Tet+' culture. Four independent experiments were analyzed. Statistical significance was determined by a t-test against a hypothetical mean value of 0, corresponding to no cell death. ***: $p < 0.001$.

As expected, before BES silencing, Pol I was present in the *18S* rDNA gene and in the active BES genes (*Luciferase* and *VSG2*) and essentially absent from Pol II loci (*β-tubulin* and *GAPDH)*. 8 hr after BES silencing (tetracycline removal), the levels of Pol I in the active BES decreased to similar background levels detected in the silent *VSG9, β-tubulin* and *GAPDH*. Hence, Pol I is indeed efficiently removed from chromatin fiber upon BES silencing.

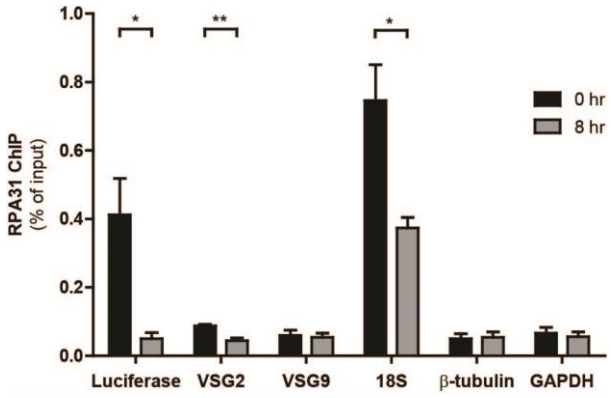


Figure 11 - A Pol I is absent from active BES 8 hr after inducing BES silencing. RNA Pol I occupancy was determines by RPA31 subunit ChIP at 0 and 8 hr after tetracycline removal. Immunoprecipitated DNA was compared to the total input material. Five independent

experiments were analyzed. Statistical significance was determined by a paired t-test against time-point 0 hr. *:$p < 0.05$; **: $p < 0.01$.

To characterize the dynamic of chromatin structure once transcription is halted, we focused on the first 8 hr because this is the period in which cells grow well and present a normal morphology (Figure 9B and 10). As a proxy of transcription, we followed luciferase activity 2, 5 and 8 hr after removing tetracycline (Figure 12A). Luciferase activity showed an exponential decrease to ~20% of the initial activity at 5 hr and to only 5% 8 hr after tetracycline removal. Luciferase activity was confirmed by quantifying mRNA transcript levels 8 hr after inducing BES silencing (Figure 12B). *18S* rDNA and *β-tubulin* did not suffer transcriptional changes, confirming that BES silencing only affected expression sites and it did not cause any other major indirect changes in the rest of the genome. *Luciferase* and *BSR* transcripts decreased 90% of the initial levels, while active *VSG2* decreased less (32%) probably due to its stability (half-life around 4.5 hr) [373]. Concomitantly, we observed a 4-8 fold transcriptional up-regulation of several silent VSGs, which is likely a result from either newly activated BESs or derepressed silent BESs.

To determine the chromatin structure of the inducibly silenced BES, we performed FAIRE (Figure 12C) and histone H3 chromatin immunoprecipitation (ChIP) (Figure 12D). By FAIRE, we observed that, both at 5 hr and 8 hr after BES silencing, the chromatin at the active BES remained highly enriched in the aqueous phase, indicating an open conformation. During this period, FAIRE-enrichment of silent BESs (*VSG9* and *VSG13* genes) remained unchanged. Although at 8 hr the chromatin of the active *VSG2* presented a 2-fold decrease in FAIRE-enrichment, it was still 123-fold higher than FAIRE-enrichment of the same gene six days post-silencing, when this expression site was completely silenced. Six days post-silencing, genes from previously silent BESs (*VSG9* and *VSG13*) showed an increase of three-fold in FAIRE enrichment. This increase was expected not to be maximal (up to around 100) because FAIRE was performed on a mixed population of switchers. As expected, *18S* rDNA and *β-tubulin* did not show major chromatin alterations at any time. These results were mirrored by histone H3 ChIP: at 8 hr, chromatin of originally active BES1 (*BSR* and *VSG2*) was still heavily depleted of histone H3, which is consistent with a high FAIRE-enrichment. *VSG2* had slightly more

histone H3, but the levels were still much lower than those detected at silent *VSG9* or *VSG13*. As expected for a healthy parasite population, ChIP of other control genes was not affected during the first 8 hr post-silencing induction, including at the *18S* rDNA.

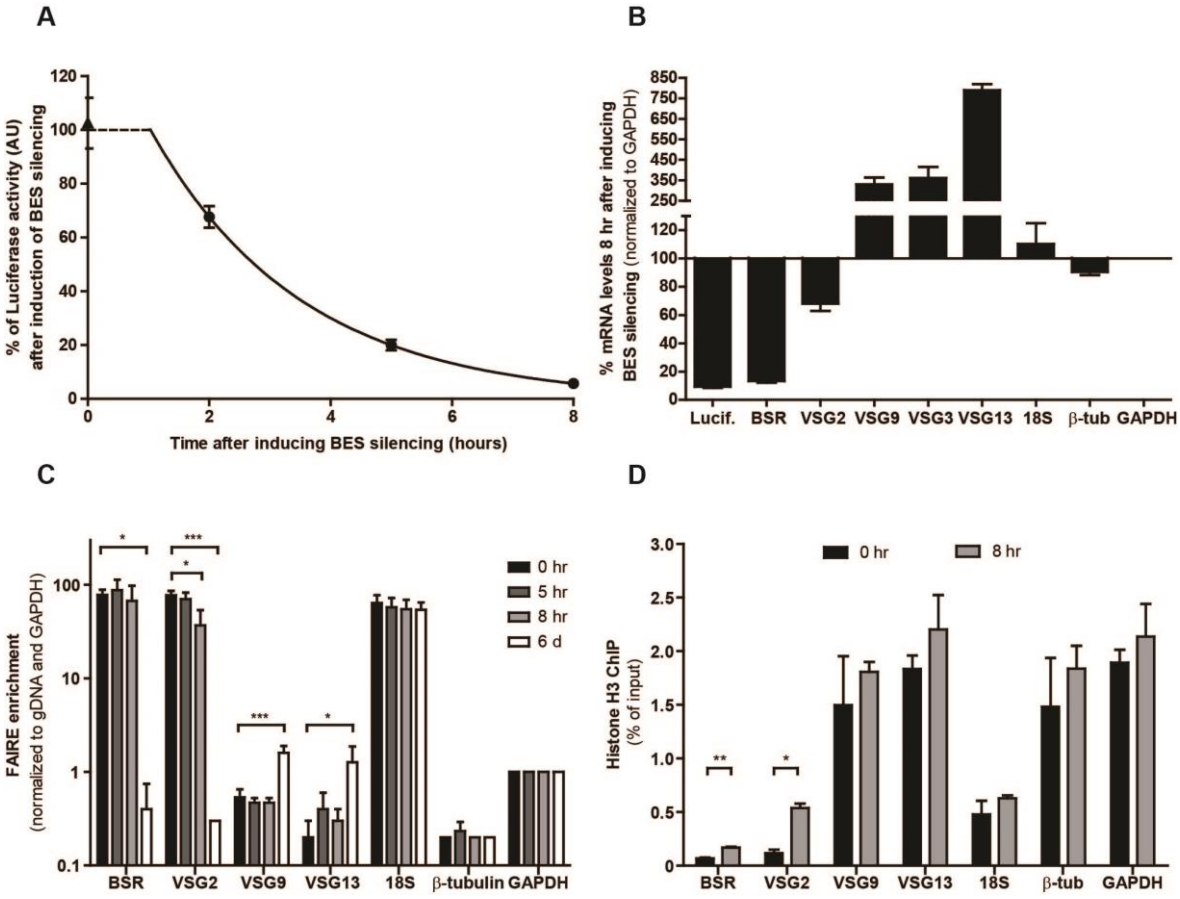


Figure 12 - BES chromatin retains an open conformation despite its transcription being reduced 90%. (A) % of luciferase activity of Tet- relative to Tet+ cells after 0, 2, 5 and 8 hr of tetracycline removal. Curve represents the best decay fit for time-points 2, 5 and 8 hr. Five to seven independent experiments were analyzed for each time-point. **(B)** % of mRNA levels after 8 hr of tetracycline removal relative to Tet+ cells, measured by qPCR and normalized to *GAPDH* transcripts. Five independent experiments were analyzed. **(C)** Chromatin conformation was measured by FAIRE at 0, 5, 8 hr and 6 days after tetracycline removal. DNA isolated by FAIRE was quantified by qPCR and normalized to gDNA copy number and *GAPDH*. Three independent experiments were analyzed. Statistical significance was determined by 1-way ANOVA with Bonferroni post-test comparison. **(D)** Nucleosome occupancy was determined by histone H3 ChIP at 0 and 8 hr after tetracycline removal. Immunoprecipitated DNA was compared to the total input material. Three independent experiments were analyzed. Statistical significance was determined by a paired t-test against time-point 0 hr. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Overall, our results show that, during VSG switching, chromatin condensation lags significantly behind transcriptional silencing, suggesting that *T. brucei* has a mechanism of maintaining chromatin open when transcription has been halted.

Chromatin conformation is cell-cycle independent

In yeast, the ratio between open and closed rDNA genes changes throughout the cell-cycle: entrance into S phase leads to repression of most rDNA genes and their chromatin becomes more compact, while transcription re-initiation in G_2 re-opens chromatin [221]. As BESs are transcribed by Pol I, we hypothesized that chromatin of active BES may also close as a function of the cell-cycle. To test this hypothesis, we stained GLB1 fixed cells with FxCycle Violet DNA stain and we FACS-sorted them, according to the DNA content, into G_1, S and G_2/M subpopulations (Figure 13A). The chromatin conformation of these subpopulations was subsequently assessed by FAIRE (Figure 13B).

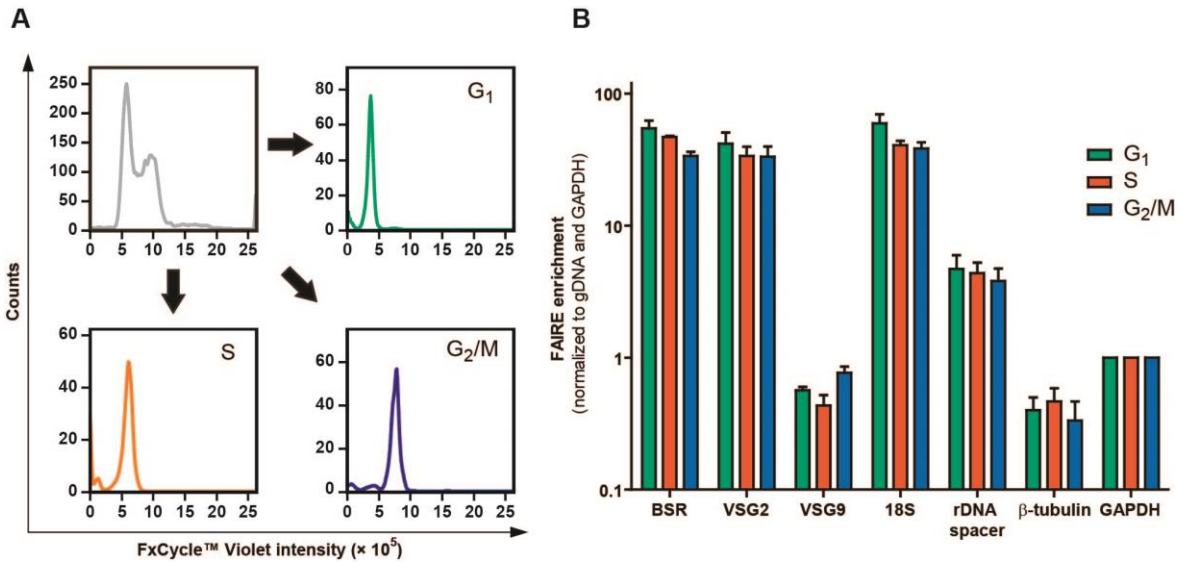


Figure 13 - Chromatin conformation of BESs is cell-cycle independent. (A) Nuclear DNA of GLB1 cells was stained by FxCycle Violet and sorted by flow cytometry in G_1 (red), S (orange) and G_2/M (blue) cell-cycle stages. Panel in grey represents original population, colored panels represent analysis of post-sorted populations. **(B)** Chromatin conformation of the three cell-cycle populations was measured by FAIRE. DNA isolated by FAIRE was quantified by qPCR and normalized to gDNA copy number and *GAPDH*. Three independent experiments were analyzed. Statistical significance was determined by a 1-way ANOVA with Bonferroni post-test comparison.

The first observation was that the FAIRE-enrichment of sorted subpopulations revealed patterns very similar to unsorted cells (*BSR*, *luciferase* and *18S* rDNA 50-100; *VSG9*, *β-tubulin* and *GAPDH* around 1) (Figure 12C), suggesting that the sorting procedure did not affect chromatin conformation of these genomic loci. Second, for each gene, the FAIRE-enrichment was constant overall for the three stages of the cell-cycle, suggesting that chromatin conformation is essentially insensitive to cell-cycle. We cannot exclude the possibility that very rapid and transient chromatin changes in conformation take place, which could not be captured at the time-resolution used here.

Unlike rDNA genes in yeast, we observed no major changes in the chromatin conformation of the active BES throughout the cell-cycle, which suggests that the mechanism that keeps BES chromatin open in the absence of transcription open is very likely cell-cycle independent.

TDP1 maintains chromatin open in the absence of transcription

In budding yeast, a Pol I transcriptional restriction for four hours under G_1 arrest does not affect the conformation of rDNA chromatin, which is maintained by HMO1, a high mobility group box (HMGB) protein [221]. In. *T. brucei*, TDP1, a high-mobility group box protein, facilitates transcription of Pol I transcribed genes [14]. Here, we tested if TDP1 was necessary to maintain chromatin open in the absence of Pol I transcription. For this, we tagged an endogenous allele of *TDP1* with a TY1 epitope and we performed ChIP. As previously reported [14], we found TDP1 highly enriched in the active BES (*Luciferase* and *VSG2* genes) and *18S* rDNA (0 hr) (Figure 14A). After 8 hr of transcriptional silencing, the chromatin of the promoter-proximal region of the BES contained the same amount of TDP1, whereas the telomeric region had two-fold less TDP1. Nonetheless, these levels were still higher (7 to 9-fold) than in silent *VSG*s. As TDP1 inversely correlates with histone H3 [14], this result is in accordance with the H3 distribution after 8 hr of BES silencing (Figure 12C and 12D), and it shows that TDP1 is still largely present in the active BES when its transcription is reduced by at least 90% (Figure 12A). The surviving switchers (six days post-silencing induction) present, as expected, very low levels of TDP1 in the previously active BES.

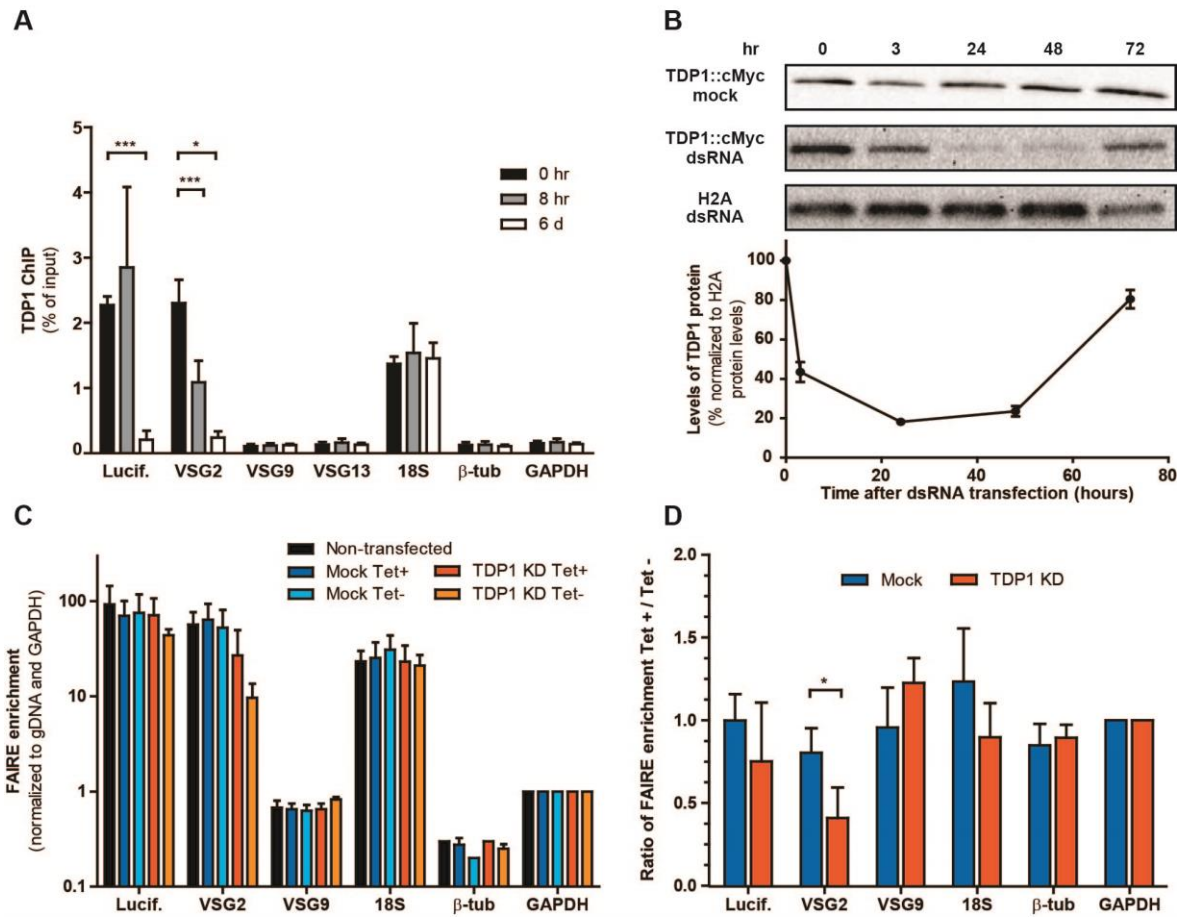


Figure 14 - When active BES is silenced, TDP1 keeps its chromatin open. (A) TDP1 ChIP at 0, 8 hr and 6 days after tetracycline removal in GLB1-TDP1::TY1, a cell-line in which one endogenous allele of TDP1 is fused with a TY1 tag. Immunoprecipitated DNA was compared to the total input material and normalized to 18S DNA. Statistical significance was determined by a 1-way ANOVA with Bonferroni post-test comparison. Three independent experiments were analyzed. **(B)** Western blotting analysis of TDP1 protein after 3, 24, 48 and 72 hr of transfection with buffer (mock) or anti-TDP1 dsRNA in GLB1-TDP1::3xcMyc, a cell-line in which one endogenous allele of TDP1 is fused with a triple c-MYC tag that is more sensitive for western blot. Time-point 0 hr indicates mock control cells transfected only with buffer. Each lane corresponds to lysates from 2 × 10^6 cells. Quantification of TDP1 signal is indicated in the **lower panel**. TDP1 protein levels were normalized for H2A protein levels and mock control. Four independent experiments were analyzed. **(C)** Chromatin conformation of GLB1-TDP1::3xcMyc cells after 5 hr of BES silencing and 24 hr of TDP1 depletion was measured by FAIRE. DNA isolated by FAIRE was quantified by qPCR and normalized to gDNA copy number and to *GAPDH*. Four independent experiments were analyzed. Statistical significance was determined by an unpaired t-test comparing Mock Tet+ and TDP1 KD Tet+. **(D)** Ratio of FAIRE enrichment (calculated from data in **panel C**) between Tet- and Tet+ for Mock and TDP1 KD conditions. Statistical significance was determined by an unpaired t-test. *: $p < 0.05$.

To test if TDP1 is necessary to maintain chromatin open when the active BES is silenced, we depleted TDP1 by transiently transfecting an anti-TDP1 dsRNA [14]. This method has been previously used to knock-down *α-tubulin* transcripts [376]. Transfection with buffer (mock control), showed no changes in TDP1 levels over time (Figure 14B, and 15A). However, transfection of anti-TDP1 dsRNA lead to a reduction of TDP1 protein levels to ~18% after 24 hr and remained low until 48 hr, after which TDP1 levels increased (Figure 5B). Consistent with a previous report [14], we observed that 24 hr of TDP1 depletion resulted in a ~43% reduction of luciferase activity, confirming the role of TDP1 as a transcriptional facilitator of the active BES (Figure 15B).

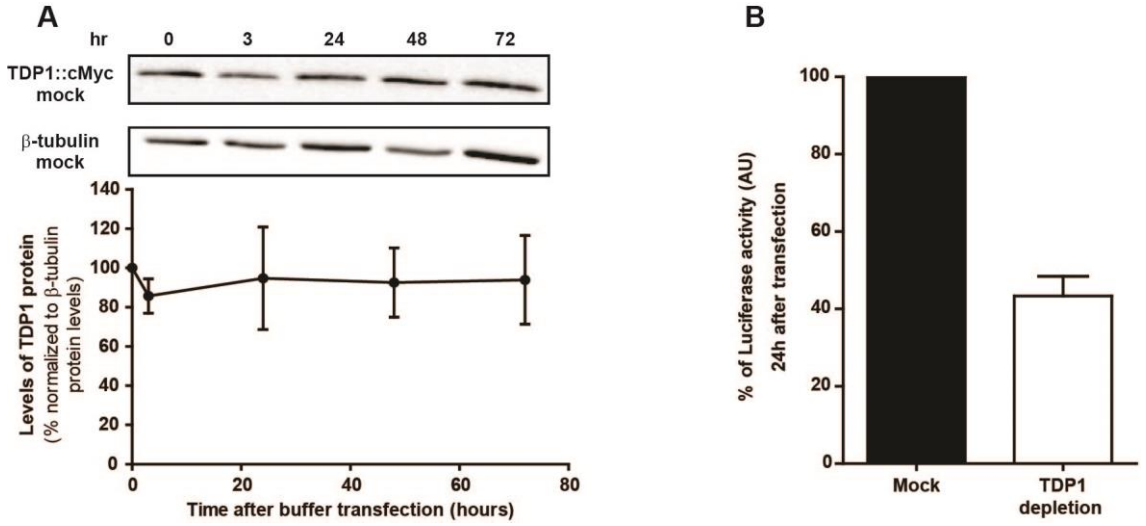


Figure 15 - TDP1 facilitates BES transcription. (A) Western blotting analysis of TDP1 protein after 3, 24, 48 and 72 hr of transfection with buffer (mock) in GLB1-TDP1::3xcMyc. Time-point 0 hr indicates cells which were not transfected. Each lane corresponds to lysates from 2×10^6 cells. Quantification of TDP1 signal is indicated in the lower panel. TDP1 protein levels were normalized for β-tubulin protein levels and to not transfected cells. Four independent experiments were analyzed. **(B)** Luciferase activity was measured 24 hr after TDP1 depletion and compared to the mock control in GLB1-TDP1::3xcMyc. Five independent experiments were analyzed.

To test if TDP1 is necessary to maintain open chromatin status in the absence of transcription, we transfected TDP1 dsRNA or buffer into the BES inducible reporter strain and, after 19 hr, we transcriptionally silenced the active BES by removing tetracycline from the medium. FAIRE was used to characterize chromatin status 5 hr post-silencing (which corresponds to 24 hr post TDP1 depletion) in four conditions: when BES silencing was induced or not, and in the presence or absence of TDP1 (Figure 14C). As expected,

chromatin changes were not detected for control genes: Pol II-transcribed genes, nor in silent *VSG9*. Chromatin of *18S* rDNA, although transcribed by Pol I, was not affected either, which is consistent with observations in yeast, in which Hmo1-null mutants do not lead to condensation of rDNA chromatin [221]. Relative to non-transfected control, mock control cells did not display changes in chromatin in any loci, including BES1, indicating that transfection *per se* did not significantly affect chromatin conformation (Figure 14C).

When TDP1 was depleted but BES1 remained active (TDP1 KD, Tet+), chromatin remained open in the promoter region and began to close at the telomere (*VSG2* is 2,6-fold more closed than mock transfection), which is consistent with the observations of Narayanan *et al*. upon depletion of TDP1 by RNA interference [14]. However, when TDP1 was depleted and BES1 was silenced (TDP1 KD, Tet-), chromatin of *VSG2* closed even further and chromatin of *luciferase* also closed slightly (Figure 14C). Next we compared the change of FAIRE-enrichment after and before silencing was induced (Tet- versus Tet+), in both conditions: mock transfection (presence of TDP1) and upon TDP1 knock-down (Figure 14D). For most genes in both conditions, this fold-change of FAIRE-enrichment is around 1, indicating that chromatin after 5 hr of silencing is not dramatically different and absence of TDP1 does not cause global changes in chromatin during this period. However, in the *VSG2* gene we detected a significant reduction in the fold-change of FAIRE-enrichment between mock and TDP1 depleted conditions. These results indicate that TDP1 is necessary to keep an open chromatin conformation when transcription of BES is halted.

We conclude that TDP1 is a key player of the chromatin of active BES, especially at the telomeric end. Narayanan *et al*. has previously shown that TDP1 maintains chromatin nucleosome-depleted and facilitates transcription. Here we show that TDP1 is crucial to keep an open chromatin in a BES that has been recently silenced (especially at the telomeric end of BES).

BES probing precedes commitment to switching

Next we investigated the advantages of having a mechanism that keeps chromatin open when transcription has been silenced. We hypothesized that this may be a mechanism that allows parasites to probe different silent BESs before committing to a

new BES. To test this hypothesis, first, we determined how long cells take to commit to a second BES. For that, we repeated the BES silencing assay but we re-added tetracycline to the cells 8, 24 or 48 hr post-silencing. Tetracycline relieves the transcriptional block of Tet repressor protein and thus allows cells that re-activate BES1 to survive (Figure 16A). Cells were cloned by limiting dilution and GFP-intensity of each clone was measured by FACS. GFP-positive clones indicate that cells re-activated BES1, while GFP-negative clones indicate cells that switched to a new BES. We observed that, none of the wells were GFP-negative at 8 hr, suggesting that all surviving cells could potentially reactivate BES1. Instead, adding tetracycline at 24 hr resulted in 49% of clones no longer expressing GFP, while at 48 hr this number was 96%. These results show that most surviving cells are already committed to a new BES two days after silencing was induced. These results also show that during the first 8 hr of switching, most cells are not committed to a VSG switch and can revert to transcribe the original BES.

We postulated that if commitment of most cells happens between 8 and 48 h, during this period we may be able to detect cells transiently probing new BESs at intermediate levels. Such cells were detected by Chaves *et al*. [348]. To test if transcription of silent BESs increases before commitment, we constructed another reporter strain, GLB1-R3, in which an RFP::NPT fusion gene was inserted downstream of the silent BES3 in GLB1 cell-line (Figure 16B). This reporter allowed us to test at single-cell level whether silent BES3 was being transiently more transcribed during switching (Figure 16C-D). Before silencing was induced, a small number of cells (1-2 %) expressed low levels of RFP (11-14-fold higher intensity than background levels), consistent with previous studies showing that silent BES are transcribed at low rate [282, 283]. 12 hr after silencing was induced, we observed an increase in the proportion of total cells expressing RFP (2-3%), suggesting that more cells are transcribing BES3. The number of cells probing BES3 increased with time up until 36 - 48 hr, in which around 20% of the cells showed elevated levels of RFP (Figure 16D).

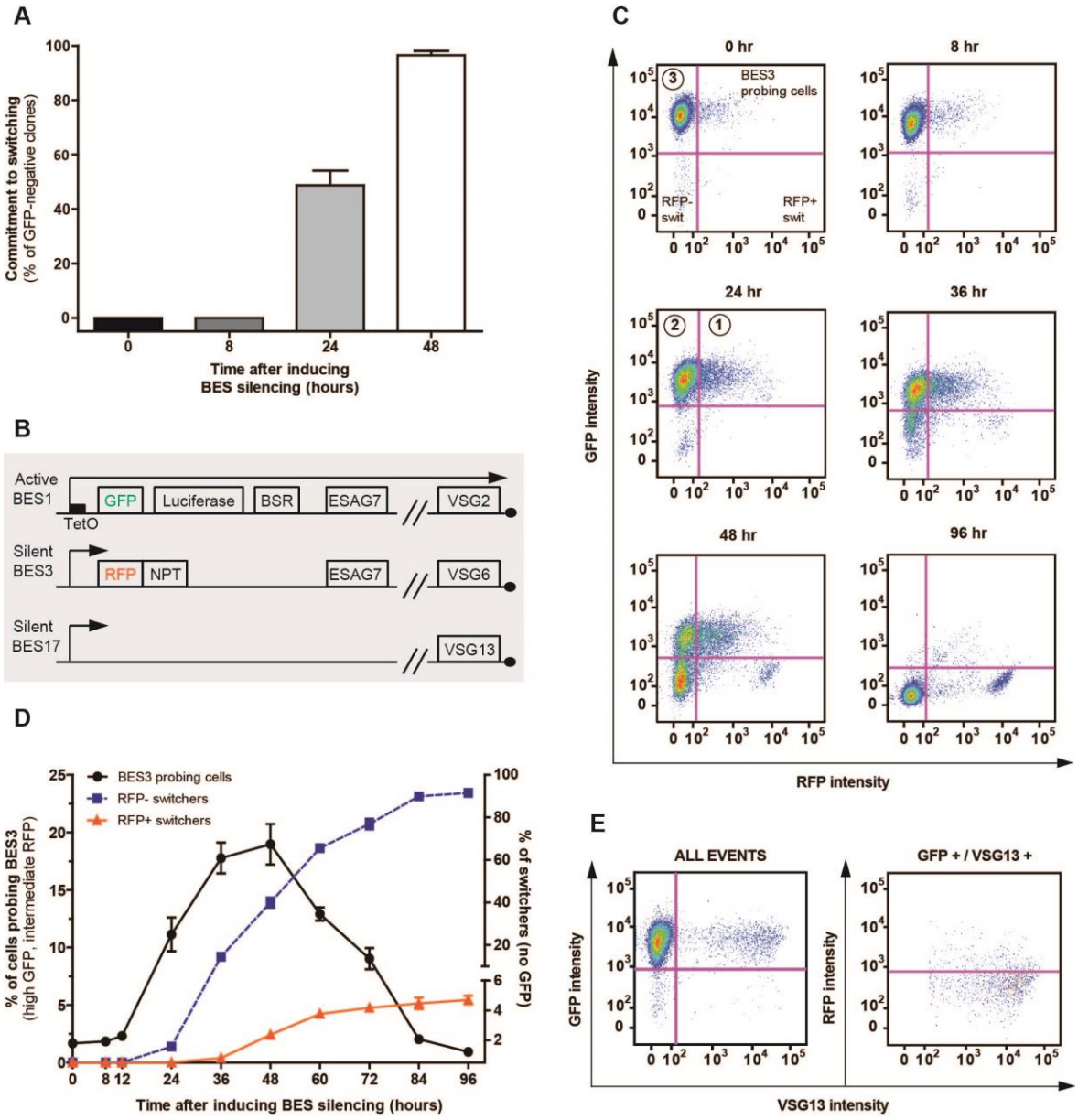


Figure 16 - Cells transcriptionally probe silent BESs for up to two days, when most cells are committed to switching. (A) Commitment assay. 8, 24 or 48 hr after inducing BES silencing, tetracycline was added back to the medium and cells were cloned. Six days later FACS was used to assess if clones were GFP-positive, indicating re-expression of original BES1. A minimum of 95 total individual wells was analyzed for each time-point between five individual experiments. **(B)** The cell-line GLB1-R3 is a derivative of GLB1, in which the fused gene *RFP::NPT* was introduced downstream the promoter of a silent BES. *NPT*, Neomycin Phosphotransferase. **(C)** Representative examples of FACS plots at several time-points post-BES silencing showing GFP and RFP expression. **(D)** Proportion of probing cells was assessed by measuring number of cells expressing RFP at intermediate levels and still present high levels of GFP (black line, left Y axis); switchers were defined as GFP-negative cells and either RFP- or RFP+ (blue and orange lines, respectively, right Y axis) after tetracycline removal. Three independent experiments were analyzed. Circled 1,2 and 3 labels indicate the sorted populations described in **Figure 17**. **(E) Left panel** - Representative example of a FACS plot at 24 hr post-BES silencing showing VSG13 and GFP expression. **Right**

panel - Representative example of a FACS plot showing VSG13 and RFP expression in the cells present on the top right gate in **left panel**. Four independent experiments were analyzed.

At 36 hr, two new and distinct GFP-negative populations were detected: one expressed high levels of RFP (0,7%) (~200-fold higher intensity than RFP background levels) and the other was RFP-negative (14,5%), suggesting that these cells are switchers that silenced BES1 and activated BES3 (red curve) or another BES (blue curve), respectively (Figure 16D). In this mixed population of cells, the switcher subpopulations became more predominant with time, while the number of cells probing silent BES3 gradually decreased. At 96 hr, switchers were almost the sole populations in culture (around 5% expressed RFP and 92% did not).

Is probing restricted to the promoter region or does it span an entire BES? To understand this, we used an anti-VSG13 antibody to test if we could detect VSG13 (from BES17) at the cell surface of silencing-induced cells. Silencing was induced in GLB1-R3 and, at 24 hr, cells were stained with anti-VSG13 (BES17) and analyzed by FACS (Figure 16E). We observed that ~5% of cells that expressed GFP also expressed heterogeneous levels of VSG13. Of these, ~20% simultaneously expressed intermediate levels of RFP and VSG13, indicating that cells can simultaneously probe two BESs, BES3 (RFP) and BES17 (VSG13). These results also show that probing is not restricted to the promoter region and, at least in some cells, the entire BES is upregulated all the way until the telomeric end of BES17.

Taken together, we conclude that parasites transiently increase the transcription levels of silent BESs for around two days, when commitment to a new BES is almost complete; probing spans the entire BES, until the VSG gene and, albeit at a low frequency, two BESs can be simultaneously probed in individual cells.

BES probing is a reversible intermediate switching step

Next we investigated whether cells that partially upregulated RFP (probing BES3) are true switching intermediates. We chose to characterize the 24 hr time-point because its FACS profile is less heterogeneous (Figure 16C) and there are fewer dead cells (Figure 10). We induced BES silencing for 24 hr, FACS-sorted GFP+/RFP+ cells, and placed them in

culture in limiting dilutions in the presence or absence of tetracycline (Population 1, indicated in 24 hr FACS plot Figure 16C). We also FACS-sorted and plated GFP+/RFP- cells (Population 2) because, although BES3 was not upregulated, other BESs may be upregulated. As a control, a population of cells in which silencing was not induced (Population 3) was also sorted (Figure 17A). Cells that attempt to reactivate BES1 die in the absence of tetracycline and survive in the presence of this drug. Therefore, the number of surviving clones six days after sorting reflects the efficiency of BES switching of each sorted population at 24 hr post-silencing. Overall, in the absence of tetracycline, 3-5% of the clones survived the silencing assay, which is consistent with 8% of surviving clones measured in the clonogenic assay of the whole population (Figure 9D) and confirms that, with this inducible system, although we can dramatically increase the switching frequency from around 1 in a million cells to 1 in 10 cells, 9 in 10 cells still fail to switch to a new BES, even in the populations 1 and 2. When tetracycline was added to the culture medium after sorting, the number of surviving clones doubled (7-9%), which was expected because at 24 hr half of the population is still capable of returning to the originally active BES1 (Figure 16A).

Seven days post BES silencing, surviving clones were characterized by FACS for their expression of GFP and RFP, which we used as reporter of BES activity: GFP+/RFP- profile indicates clones that probably did not switch and reactivated BES1; GFP-/RFP+ indicates switchers to BES3; and GFP-/RFP- indicates switchers to another BES. No clones were obtained that simultaneously expressed GFP and RFP. For both populations 1 and 2, all post-sorting clones obtained in the absence of tetracycline no longer expressed GFP and thus represent switchers. Interestingly, from population 1, 57% (12 clones on average of three experiments) of switchers expressed RFP, while 43% (9 clones) didn't, indicating that cells can probe one BES at 24 hr post-silencing and eventually switch to another BES (Figure 17B). From population 2, all switcher clones were RFP-negative (13 clones), indicating that switchers activated a BES other than BES3 (Figure 17B). These results suggest that if a BES is not probed at 24 hr, apparently it is not fully activated later.

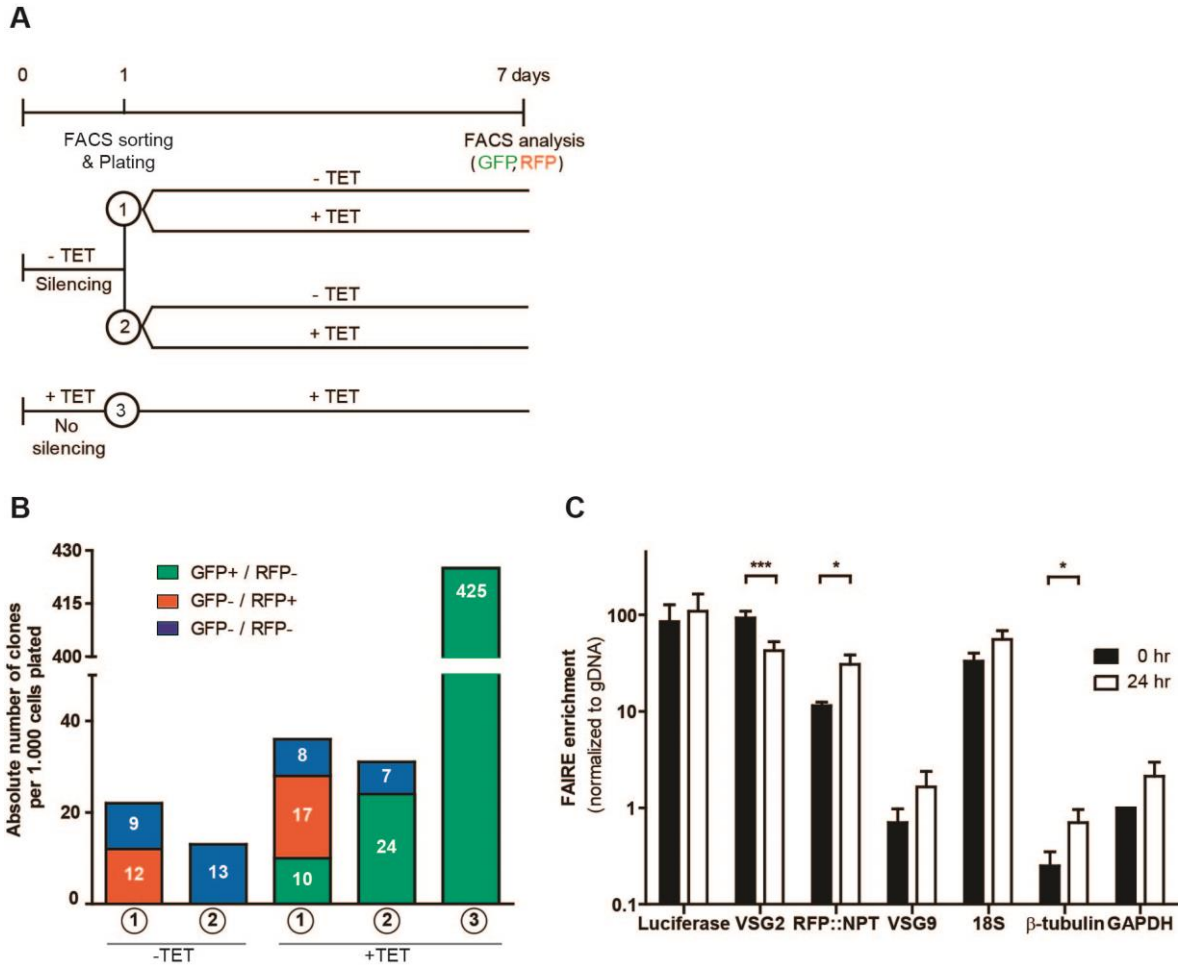


Figure 17 - Probing cells are switching intermediates that can choose different fates. (A) Experimental design of FACS sorting and subsequent phenotype characterization. Tetracycline was removed from the medium of GLB1-R3 cell-line. One day later, two populations 1 and 2 (indicated in **Figure 16C**) were sorted and plated by limiting dilution. Population 1 consists of cells that express GFP and RFP, while population 2 consists of cells that express GFP, but not RFP. As control, a third population of GFP+/RFP- cells was sorted from a culture kept under tetracycline pressure. Clones obtained after six days were characterized in terms of expression of GFP and RFP by FACS. **(B)** GFP and RFP expression of surviving clones was assessed by FACS. Numbers in white show mean number of clones with each specific phenotype from three individual experiments. **(C)** Chromatin conformation of GLB1-R3 cells after 24 hr of BES silencing was measured by FAIRE. DNA was quantified by qPCR and normalized to gDNA copy number and to *GAPDH* at 0 hr. Statistical significance was determined by a paired t-test against time-point 0 hr. *: $p < 0.05$; ***: $p < 0.001$.

When sorted populations were plated in the presence of tetracycline we obtained some clones that expressed GFP and not RFP, as expected and consistent with the fact that at 24 hr there are still many cells that are not committed for switching and reactivate BES1 (Figure 17A). Post-sorting clones from population 1 showed similar proportions of

the three different types of GFP/RFP expression profiles, indicating a large plasticity of this population before commitment (Figure 17B). In contrast, most post-sorting clones from population 2 reactivated BES1 (77%, 24 clones) and none activated BES3 (Figure 17B), suggesting once again that if a BES is not probed at 24 hr, it seems not to be activated later. Clones obtained after sorting population 3, in which silencing was never induced, resulted as expected in 100% of clones expressing GFP and not RFP (Figure 17B), consistent with BES1 remaining active. This shows that the stress associated with sorting did not induce any unexpected changes in BES expression.

Taken together, our results indicate that cells in a probing state are not a dead end product resulting from the silencing inducible system used in this work. Importantly, these data show that probing cells can revert to the original BES1, or switch and commit to the probed BES, or switch and commit to another BES. Probing is therefore a reversible step during BES switching, allowing cells to reactivate the originally active BES or activate previously silent BESs.

Given that at 24 hr, the population consists mainly of cells capable of switching or reverting to original BES (with only around 3% of dead cells, Figure 10), we decided to check whether chromatin of the originally active BES1 remains accessible during this period in this cell-line (GLB1-R3). For this, silencing was induced and cells were collected for FAIRE at 0 and 24 hr post-silencing (Figure 17C). At 24 hr, chromatin in *VSG2* became significantly more compact than at 0 hr, which is consistent with the trend previously observed by FAIRE in GLB1 cell-line at 8 hr (Figure 12C), but the conformation is still 25-fold more open than that of a silent *VSG9*. Consistent with probing phenotype, we detected a significant increase in chromatin accessibility of the *RFP::NPT* gene. All other tested genes showed a slight increase in accessibility, which may be due to the fact that, at 24 hr, the culture has some cells arrested in G_2/M or with an abnormal DNA content (Figure 9C), or already dead (Figure 10).

Taken together, our data reveal a novel association between the alterations of chromatin structure at the active BES and BES switching. Chromatin is kept essentially open for at least 24 hr, while silent BESs are reversibly probed before committing to a new BES.

DISCUSSION

In this study, we showed that the chromatin structure of the active BES is cell-cycle independent and TDP1 keeps its open conformation when transcription is halted, especially in the telomeric region. We propose that these properties are critical for the dynamics of switching between BESs because it gives time for cells to reversibly probe multiple silent BESs before committing to a new BES.

Transcription and chromatin dynamics in the active BES

The initial trigger that leads to VSG switching remains a mystery. However, several studies in which transcription of the active BES was somewhat interrupted or diminished, either due to loss of the VSG upstream sequence (CTR) [364] or replacement of a BES promoter by the T7 promoter [289], resulted in more frequent *in situ* switching, suggesting one of the earliest events during VSG switching is the silencing of the active BES. In this work, we first checked if halting transcription is also one of the earliest steps during differentiation from bloodstream to procyclic forms, a process in which the active BES needs to be silenced while procyclin genes are upregulated. We show that, 24 hr after inducing differentiation, transcription of the active BES is highly reduced to 2% (Figure 8B), while the chromatin of this locus remains essentially open (Figure 8C). We conclude that during differentiation, transcription silencing precedes chromatin changes, and thus it is likely that a similar mechanism happens during VSG switching in bloodstream forms.

In the tetracycline-inducible system developed by the Horn lab, when the Tet repressor binds the Tet operator, it sterically blocks Pol I transcription, which results in silencing of the initially active BES and activation of a new one (Figure 9). In this work, we confirmed that inducible BES silencing promotes a rapid decrease in transcript levels of genes from initially active BES (at least 90% drop in the first 8 hr, as accessed by transcript levels of the unstable *luciferase* reporter) and an eviction of Pol I from chromatin fiber, which confirms that transcription of active BES is efficiently and rapidly stopped (Figure 11, 12A and 12B). During this period, chromatin remained in an open state despite transcription being halted (Figure 12C and 12D). By ChIP, we observed a slight enrichment

of histone H3 in the *VSG2* gene and a decrease in TDP1, suggesting a more compact chromatin mainly at the telomeric end of BES.

Our study also shows that chromatin conformation of the active BES is cell-cycle independent. Using FAIRE, we detected no significant differences in chromatin conformation of active BES in cells in G_1, S and G_2/M. Given that, when parasites divide, most daughter cells use the "mother's" BES and only very few switch to a new BES, it is tempting to speculate that keeping an open chromatin structure serves as an epigenetic marker that, after cell division, signals which BES should be used by the daughter cells. Perhaps TDP1 interacts with the cohesion complex, which is necessary after S phase to maintain the two sister chromatids of the active BES associated to the single Expression Site Body while waiting for chromosome segregation during mitosis [265]. This hypothesis deserves further investigation in the future.

TDP1 maintains open chromatin when BES has been silenced

High-mobility group box (HMGB) proteins are essential nuclear components in chromatin structure, transcriptional activity and DNA damage repair [300-302]. In yeast, Hmo1 is a HMGB protein that is associated with Pol I transcription machinery [308], which can competitively displace histone H1 [295] and also maintain chromatin in an open conformation [221]. In *T. brucei*, TDP1 is an essential HMGB protein that is highly enriched in Pol I loci [13, 14]. Like yeast Hmo1, TDP1 is a facilitator of Pol I transcription and it is necessary to keep chromatin nucleosome-depleted in a steady-state situation. Here we show that when BES silencing was induced in the absence of TDP1, chromatin became significantly more compact, indicating that TDP1 is necessary to maintain an open chromatin structure in the absence of transcription. This role is consistent with what has been observed in yeast, in which Hmo1 maintains Pol I chromatin open when transcription is halted [221].

In eukaryotes, deposition of canonical histones is normally replication-dependent [377]. Although our data shows that maintenance of BES chromatin structure is cell-cycle independent (Figure 13), we predict that during the rare event of VSG switching, chromatin remodelling occurs in S phase of the cell-cycle. Thus, during replication, eviction of TDP1 from a silencing BES may be an important step to complete a BES

switching. Regulating the timing of TDP1 eviction may be a means to determine for how long a cell stays in a "probing" stage, before commitment.

It is interesting to observe that, in two different circumstances (BES transcriptional silencing in wild-type or TDP1 depleted conditions), the chromatin of the BES telomeric end is more rapidly closed than chromatin close to the BES promoter (which can be up to 50 kb upstream from telomere). Telomeres possess a highly controlled chromatin structure and, in trypanosomes, they are very important for antigenic variation [378]. It is possible that, upon transcriptional silencing, loading of nucleosomes onto chromatin happens faster at the telomeric region. Tiengwe *et al*. found that although most origins of replication are found in core regions of the chromosome, VSG genes also harbor origins of replication, which could drive inward replication and deposition of nucleosomes [268]. Spreading of such chromatin condensation may happen evenly towards the BES promoter or it may happen in steps if the BES has insulator elements, such as the 70-bp repeat array for example, which could prevent spreading of nucleosome loading into the upstream ESAGs.

Probing and commitment to a new BES

Chaves and colleagues proposed that, during a switching event, a cell pre-activates a silent BES before a natural intermediate rapidly and transiently express two VSGs and ultimately commits for switching or reverts to the original BES [348]. A more recent study, in which BES transcriptional attenuation was observed as a consequence of ectopically overexpressing a silent VSG, also proposed that this attenuation could give time for parasites to probe silent BES before switching to a fully competent BES [368]. The nature of what cells are probing is unknown. It is possible that cells are testing the order of VSG switching, or if ESAGs are functional, or if previously untested or mosaic VSGs are functional. Our results are in agreement with these models of "probing before decision" and we propose that regulation of chromatin structure may be the underlying epigenetic mechanism. We show here that trypanosomes test silent BES before a choice is made and this is associated to keeping chromatin of the active BES open (Figure 18).

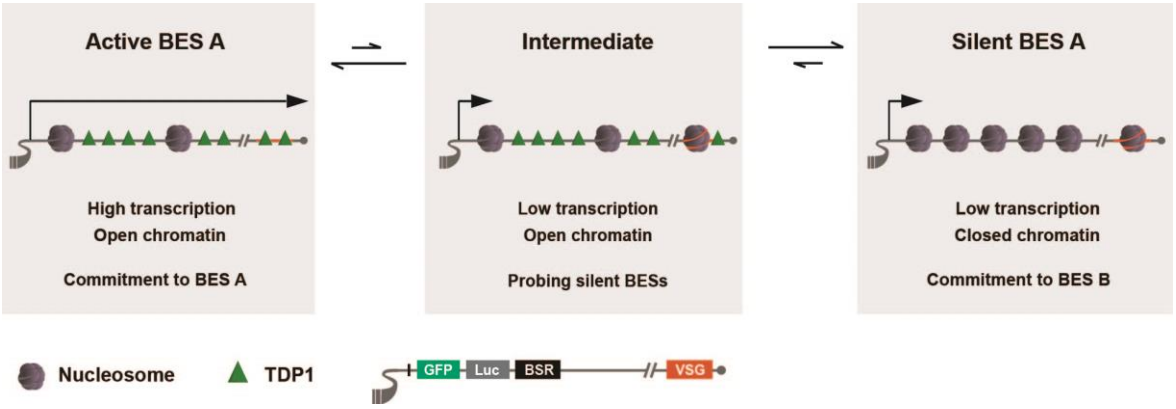


Figure 18 - Model for VSG expression site switching: transcriptional probing silent BESs before commitment is associated to a temporary maintenance of open chromatin by TDP1. Chromatin of the actively transcribed BES possesses an open chromatin enriched in TDP1. Upon transcriptional silencing of the active BES, cells undergo an intermediate state characterized by stabilization of the active BES open chromatin by TDP1 while probing silent BESs. Commitment to switch or not switch seems to take around two days. This decision can either be returning to the initial active BES (which, in our reporter cell-line, because BES1 is blocked by TetR, it would result in cell death) or switching to a new BES. In this later case, the chromatin of the originally active BES loses TDP1 and becomes nucleosome-enriched.

Our commitment assay showed that, 8 hr after silencing, no cells are committed to switching although silent VSG transcripts are already detected (Figure 12B). Half of the cells are committed 24 hr post-silencing, while most of them are committed at 48 hr. From 12-48 hr, more cells probe silent BESs, but only a fraction of the cells expresses higher levels of RFP, which indicates that not all BESs are being probed at the same time in each cell. Nevertheless, we found that at least two BESs can be probed simultaneously by the same cell (Figure 16E). At 24 hr, a small fraction of cells becomes GFP-negative, consistent with committed switching and this population becomes more predominant with time. These results show that probing of one or more BESs is a reversible process that lasts around two days, during which cells either switch to the probed BES, switch to a different BES or revert to BES1. Reverting to original BES1 is probably facilitated by the fact that BES1 retains an open chromatin structure for at least 24 hr (Figure 17C).

In some aspects, our model resembles the formation of poised chromatin. Poised chromatin at Pol II transcribed genes is characterized by the presence of bivalent marks; the co-existence of activating and repressive histone marks within the same domains

[379]. This epigenetic status has been associated to genes that are transcribed at low levels but need to be in a prepared state for developmental fates of rapid activation or repression, such as differentiation of embryonic stem cells or activation of T-cells [380, 381]. A poised chromatin state has also been found at *var* genes in *Plasmodium falciparum*, the causative agent of malaria. Switching between *var* genes is essential for antigenic variation [346]. It is possible that *T. brucei* uses a TDP1-dependent mechanism as a means of temporarily keeping the active BES in a poised state, ready for being activated or repressed. In the future, it will be interesting to test whether this mechanism of keeping chromatin temporarily open also facilitates VSG switching by recombination. In lymphocytes, for example, it has been already shown that the chromatin that is accessible for V(D)J recombination typically displays elevated acetylation of histones H3 and H4, which is a hallmark of a more accessible chromatin [382].

In this study we showed that the chromatin structure is kept open by TDP1 probably to facilitate probing, before cells commit to a new BES. It is intriguing that, during differentiation, chromatin of the originally active BES is also hold open for some time while transcription has already been reduced. This suggests that differentiating cells may also undergo an intermediate stage, which may be reversible. This is consistent with findings by Batram *et al*., in which attenuation of the active BES results in an intermediate stage that can reversibly progress in two directions: differentiation to procyclic or returning to proliferation in the mammalian form. The implications of chromatin dynamics in differentiation should be further investigated, and preferably in a pleomorphic strain, whose differentiation process is more similar to what happens *in vivo*.

ACKNOWLEDGMENTS

The authors would like to thank Christian Janzen and David Horn for the kind gifts of histone H3 antibody and 2T1.ESPiGFP:NPT cell-line, respectively. The authors would also like to thank Ruy M. Ribeiro and Sandra Trindade for help with statistical analysis, Daniel Neves for drawing the model in Figure 18 and George A.M. Cross for critically reading the manuscript of the published paper.

CHAPTER 3

OVEREXPRESSION OF A HIGH MOBILITY GROUP BOX PROTEIN COMPROMISES VSG MONOALLELIC EXPRESSION IN *TRYPANOSOMA BRUCEI*

Francisco Aresta-Branco[1], Luísa M. Figueiredo[1]

[1] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Avenida Professor Egas Moniz, 1649-028 Lisboa, Portugal

Author Contributions

FAB and LMF designed the experiments. FAB performed all the experiments. FAB wrote the chapter. LMF supervised the work and revised the text.

INTRODUCTION

High Mobility Group box (HMGB) proteins are characterized by their small size, rapid mobility and by containing at least one high mobility group box domain, responsible for DNA binding [293, 294]. HMGB proteins have been reported to fulfill several important roles that range from extracellular mediator of inflammatory responses, angiogenesis or wound repair [297-299], to nuclear regulator of essential processes such as transcription, replication, DNA repair or V(D)J recombination [300-303]. The HMG-box domain is able to bind to bent, kinked or unwounded DNA structures in order to bend, distort or loop DNA in either a sequence or non-sequence specificity [295].

HMGB proteins were shown to influence transcription by three different mechanisms: they facilitate nucleosome sliding by chromatin remodeling proteins [304, 305], they serve as a transient chaperone for stable binding of transcription factors [306] or they participate in transcription blockage as part of a complex that inhibits the assembly of a preinitiation complex on promoters [307]. HMGB proteins also participate in the regulation of ribosomal DNA (rDNA) genes, which are transcribed by RNA polymerase I (Pol I). Upstream binding factor (UBF) in mammals and Hmo1 in yeast are two HMGB proteins important for the chromatin structure and transcription of rDNA loci [308, 383]. Both UBF and Hmo1 bind throughout the rDNA actively transcribed genes [313, 322], while only UBF binds to the rDNA promoter as part of the preinitiation complex (PIC) [383]. In yeast, a similar role is fulfilled by upstream-associated factor (UAF) in the promoter [173]. Nonetheless, UBF and Hmo1 are important to establish and maintain open chromatin conformation in Pol I loci [166].

During the life cycle of *T. brucei*, parasites shift between the mammalian host and the Tsetse fly [286]. To survive in such different hosts, parasites undergo significant changes in gene expression, which include replacement of the VSG by the similarly abundant procyclins [287] and then to metacyclic VSGs before injection back to the mammal [33]. Both bloodstream and metacyclic VSG genes are transcribed from subtelomeric loci called Bloodstream and Metacyclic Expression Sites (BESs and MESs), respectively. BESs are polycistronic units with a variable number of expression-site-associated genes (ESAGs), whereas MES are transcribed monocistronically [93, 98, 384]. On the other hand, procyclin genes are located in the internal part of chromosomes

distributed among two loci [271]. Strikingly, both VSGs and procyclins coding genes are transcribed by Pol I [4, 366]. Although the *T. brucei* genome has more than 2000 VSG genes and pseudo genes [91], in bloodstream forms only one VSG is transcribed from one of the approximately 15 BESs present in the genome at a specific time. This ensures the VSG monoallelic expression, one of the hallmarks of antigenic variation.

While in the mammalian-infective bloodstream forms (BSF), the chromatin of the active and silent BESs is dramatically different. The actively transcribed BES is nucleosome-depleted (open state), while silent BESs are organized in regularly spaced nucleosomes [9, 10]. Procyclin loci possess a more open chromatin structure than silent BESs, since procyclin genes are transcribed in BSFs but down-regulated post-transcriptionally, whereas MESs chromatin structure is completely closed [238]. Besides the multi-subunit class I transcription factor A (CITFA), which binds only to active BES promoters and is necessary to regulate BES transcription initiation [233], only TDP1 has been shown to be a core structural component of the active BES [14]. TDP1, one of the few HMGB genes in the *T. brucei* genome and so far the only identified nuclear HMGB protein, is highly enriched in Pol I actively transcribed regions, namely the active BES and rDNA genes [14]. TDP1 depletion results in a striking growth arrest, chromatin repression and VSG transcriptional downregulation. We have shown in the previous chapter that TDP1 is important to maintain open chromatin structure in the active BES after inducing a transcriptional silencing. This functional role allows the probing of silent BES before commitment for switching to a new BES or returning to the initial BES is undertaken, demonstrating an impact of TDP1 in antigenic variation.

In this chapter, we characterized the phenotype of a TDP1 gain-of-function mutant. We found that upon TDP1 overexpression, growth rate was slightly reduced but chromatin of silent BES, MES and procyclin loci became decondensed leading to transcriptional derepression of these loci. Moreover, monoallelic expression becomes disrupted as expression of a silent VSG was observed. Although these results are still preliminary and require further experiments before publication, we show that the tight control of TDP1 expression is important for monoallelic VSG expression.

MATERIALS AND METHODS

Trypanosome cell-lines and plasmid construction

T. brucei bloodstream form (BSF) parasites (strain Lister 427, antigenic type MiTat 1.2, clone 221a) [385] were cultured in HMI-11 as described in [371]. PL1S cell-line was described in [258]. POT clones were generated by transfecting pFAB9 in PL1S with an AMAXA nucleofector (Lonza), program X-001, using the previously optimized homemade Tb-BSF buffer (90 mM Na_2HPO_4, 5 mM KCl, 0.15 mM CaCl2, 50 mM HEPES, pH 7.3) [372]. pFAB9 contains two bidirectional promoters: a T7 promoter that promotes transcription of the *BLE* gene to allow selection of positive clones with phelomycin and rDNA promoter followed by two Tet Operator sequences and a TDP1 gene with C-terminal TY1 tag and an aldolase 3'UTR. pFAB9 was linearized with *Not*I (New England Biolabs), for integration in of the rDNA spacer sequences. Overexpression was induced by adding 1 μg/ml of Tetracycline to the medium.

All cloning was performed using the In-Fusion® HD Cloning system (Clontech) following to the manufacturer's instructions.

Western Blotting

Cells were lysed using Laemmli buffer, resuspended at 1×10^5 cells/μl and treated with 200 U/ml Benzonase (Sigma). A SDS-PAGE was performed at 4ºC and proteins were transferred to a nitrocellulose membrane using an iBlot Dry Blotting System (Invitrogen) for 6 min and 30 sec. Membrane was blocked with 5% milk in PBS / 0.1% Tween for at least 1 hour and primary mouse anti-TY1 antibody [386] and mouse anti-β-tubulin [387] were incubated overnight at 4ºC in a 1:2000 and 1:1000 dilutions, respectively, in 3% milk in PBS / 0.1% Tween. Membrane was washed three times with PBS / 0.1% Tween, and incubated with anti-mouse horseradish peroxidase-linked secondary antibody (Amersham) at room temperature for 1 hr in a 1:20000 dilution in 3% milk in PBS / 0.1% Tween. Membrane was further washed three times with PBS / 0.1% Tween and developed with a Western Lightning Plus-ECL (PerkinElmer). Film acquisition was made using Fuji Medical X-ray films (Fujifilm) in a Hypercassette (Amersham) in a CURIX 60 tabletop processor low-volume radiology (AGFA).

FAIRE

Formaldehyde-assisted isolation of regulatory elements (FAIRE) was performed as described previously [238]. 20 million cells were centrifuged for 20 min at 3000 g, resuspended in 20 ml of HMI-11 and crosslinked with 1.1% Formaldehyde (Sigma) for 10 min. In cell-cycle sorted fractions, protocol was performed with only 1-2 million cells. Crosslinking was stopped with 0.125 M of Glycine (Sigma) and cells were washed once with Trypanosome Dilution Buffer (TDB) (5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$, 20 mM Na_2HPO_4, 2 mM NaH_2PO_4, 20 mM glucose, pH 7.4). Cell pellets were lysed with Lysis Buffer (50 mM TrisHCl, 10 mM EDTA, 1% SDS) complemented with Protease Inhibitor Cocktail (Sigma) and Solution P (PMSF, Pepstatin). Lysates were incubated on ice for 15 min and an external plasmid DNA spike was added before sonicating with 10 cycles of 30 sec on / off. Lysates were centrifuged at 16000 g, 4ºC to pellet cell debris and a 100 □l aliquot of the supernatant was taken to check sonication efficiency. The remaining fraction was subjected to two Phenol-Chloroform-Isoamyl Acid (25:24:1) (Sigma). DNA present in the aqueous phase, corresponding to open chromatin, was precipitated and washed with 2 volumes of absolute ethanol, sodium acetate 0.3 M and 20 μg/ml glycogen overnight at -20°C DNA was washed once with ethanol 70% and dried before resuspension in Elution Buffer (Quiagen). Samples were then treated with 100 μg/ml of RNase A (ROTH) and purified with a PCR purification kit (Quiagen). A non-crosslinked sample was always included to normalize qPCR data for gene copy number. Quantification of the FAIRE and total DNA samples were analyzed by qPCR with primers presented in Table 5. Amplification reactions were performed in duplicates.

Transcript quantification

Parasites were harvested by centrifugation at 650 g for 10 min at 4°C and immediately resuspended in PureZOL (Bio-Rad) or TRIzol (Life Technologies). RNA was isolated following the manufacturer's instructions and RNA quantity and quality was assessed on a NanoDrop 2000 (Thermo Fisher Scientific). cDNA was generated from 200 ng of RNA using a Superscript cDNA Synthesis Kit (Life Technologies), according to manufacturer's protocol. Quantitative PCR (qPCR) was performed using 1× SYBR Green PCR Master Mix (Applied Biosystems) using primers presented in Table 5. Negative controls lacking reverse transcriptase (RT-) were confirmed by qPCR. Amplification

reactions were performed in duplicates. The ΔΔCt method as used to determine transcript levels relative to normalizing gene.

Luciferase assay

1.5×10^6 cells were harvested for 5 min at 2800 g, 4ºC and washed once with 1 ml cold TDB. Pellets were resuspended in Lysis Buffer (Biotium) and protocol was followed according to the manufacturer's instructions. Luminescence was measured by a microplate reader (Tecan).

VSG staining of live cells

0.5×10^6 cells were harvested for 5 min at 2800 g, 4ºC. Cells were resuspended in 50μl of cold HMI-11 in which Alexa Fluor 647 anti-VSG13 conjugated antibody had been previously diluted (1:5000). After 15 min of incubation at 4°C with gentle shaking, cells were washed three times in cold TDB, resuspended in cold TDB, and immediately analyzed on a FACS Fortessa (Becton Dickinson Biosciences). Data were processed with FlowJo software (FlowJo, LLC).

Table 5 – List of primers used in the present chapter.

Primer name	Primer sequence (5' ->3') forward / reverse
BSR_qPCR	CGGCTACAATCAACAGCATC / ACGATACAAGTCAGGTTGCC
VSG9_qPCR	ACTAAGCTCGTGGCGCAC / CGCGTAGTTGACGCATGAC
Luciferase_qPCR	ATGTCCGTTCGGTTGGCAG / CATACTGTTGAGCAATTCACG
VSG2_qPCR	AGCAGCCAAGAGGTAACAGC / CAACTGCAGCTTGCAAGGAA
VSG13_qPCR	ATAACGCATGGCCATCTTGAC / GTCGTTGCTGTGGATTGCTC
18S_qPCR	ACGGAATGGCACCACAAGAC / GTCCGTTGACGGAATCAACC
mVSG639_qPCR	TCGCACTTTCAGCTCTGCTC / GCCGACCACTCGCTGTCC
GPEET2_qPCR	ACGGGACCAGAGGAAACTG / TAGAATGCGGCAACGAGAG
β-tubulin_qPCR	TTCAGGCTGGCCAATGCG / TACGGAGTCCATTGTACCTG
GAPDH_qPCR	AGATTGATGTCGTTGCTGTTGTG / ATGGCTTGCTCTTCGTAGTCG
Ampicillin_FAIRE DNA spike_qPCR	ATCGTTGGGAACCGGAGC / AGCGCAGAAGTGGTCCTG

RESULTS

TDP1 overexpression is well tolerated

The high mobility group box proteins (HMGB) UBF and Hmo1 are structural chromatin components with an essential role in RNA Polymerase I (Pol I) elongation [308, 324]. Overexpression of UBF1, one of the proteins that compose UBF complex, leads to increased levels of rDNA genes transcription [327]. On the other hand, the overexpression of Hmo1 causes yeast cells to enter a vegetative growth [318]. In *T. brucei*, TDP1 is a core component of Pol I-loci chromatin and its depletion leads to chromatin repression and transcriptional downregulation, especially in the actively transcribed bloodstream expression site (BES) [14, 388]. How overexpression of TDP1 can impact monoallelic VSG expression poses an important question for antigenic variation.

A tetracycline-inducible system was first generated to conditionally overexpress C-terminal TY1-tagged TDP1 gene from the rDNA spacer locus in PL1S cell-line (Figure 19A) [258]. PL1S expresses VSG9 from BES2 and contains a *luciferase* reporter gene in the promoter region of silent BES1. Two independent clones from this PL1S Overexpressing TDP1 cell-line were named POT1 and POT2. Overexpression was first confirmed in two POT clones by western blot after 24 and 48 hr of induction (Figure 19B). A control cell-line was used which possesses a TY1-tag in a histone H1 gene to evaluate the functionality of the anti-TY1 antibody. The levels of the TY1-tagged allele are undetectable before induction, indicating a tight-control of the inducible expression in both clones. At 24 hr, a clear band of the expected size (~32kDa) could be detected in both clones and its intensity was higher at 48hr, indicating that overexpression was successful. Both clones show a normal growth within the first 48 hr of overexpression, after which a small reduction in growth is observed (Figures 19C and 19D).

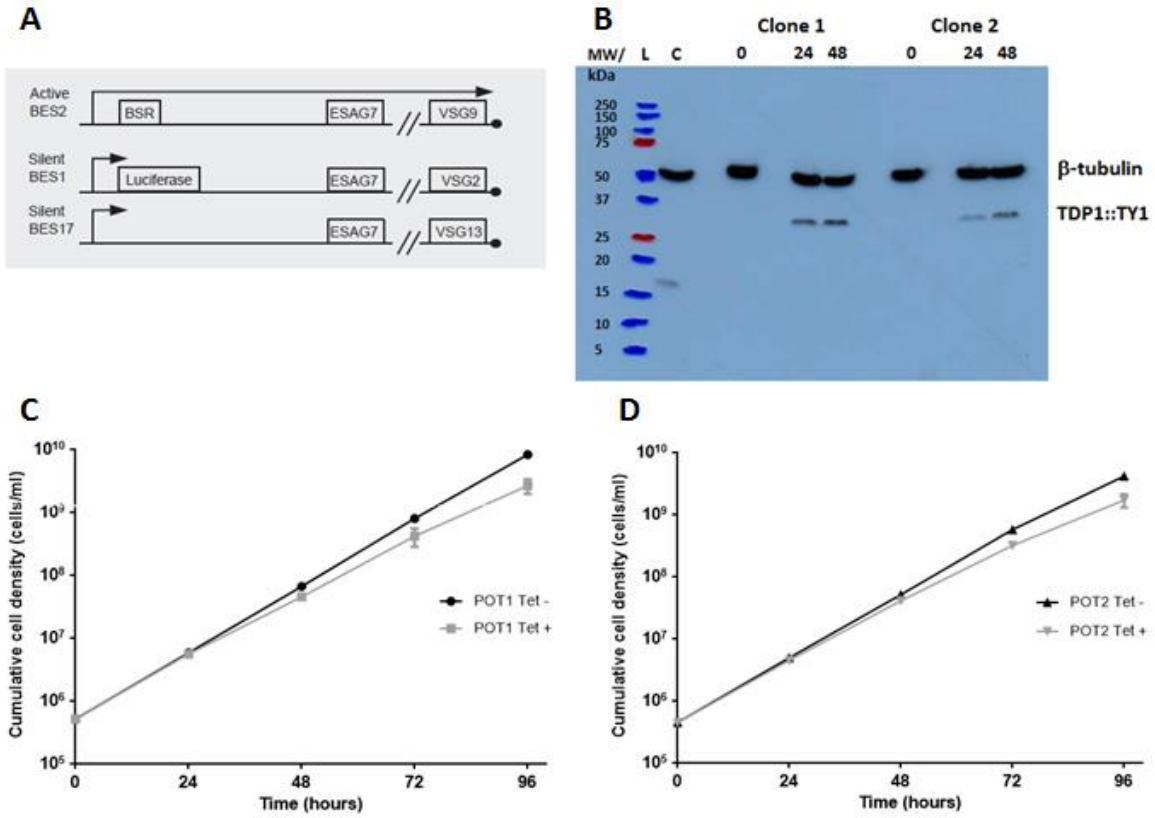


Figure 19 – Inducible overexpression of TDP1 in bloodstream forms. (A) The cell-line POT is a derivative of PL1S, in which a tetracycline-dependent overexpression construct for TDP1::TY1 was integrated in a rDNA spacer locus. **(B)** Western blotting analysis of TDP1 protein after 24 and 48 hr of overexpression in two POT clones. Time-point 0 hr indicates non-induced condition and shows no TDP1::TY1 protein. Each lane corresponds to lysates from 1 × 10^6 cells. L stands for protein standards ladder (with the corresponding molecular weights to the left) and C stands for a control cell-line which has a TY1-tag in a histone H1 gene. **(C and D)** Growth curves of POT1 **(C)** and POT2 **(D)** before (Tet-, grey curve) or after (Tet+, black curve) induction of overexpression. Three independent experiments were analyzed.

TDP1 opens chromatin of many different loci but only silent BESs, MESs and procyclin loci become upregulated

TDP1 in *T. brucei* keeps chromatin of active BES open in the presence [14] and absence of transcription as observed in the previous chapter. Given that TDP1 is a chromatin component the actively transcribed BES, it was hypothesized that overexpression of TDP1 would lead to its loading in sites typically devoid of TDP1, such as silent BESs. We used FAIRE to measure the global impact in chromatin conformation 24 and 48 hr after TDP1 overexpression (Figure 20A and 20B) [9, 238, 388]. In both POT clones, we observed that the chromatin of silent BES became more open with a more

dramatic change for POT1. At 48 hr, the promoter region of silent BES (*Luciferase*) opened ~50-100 fold relative to the non-induced condition, whereas the correspondent VSG in the same silent BES (*VSG2*) became only 4-9 fold more open when compared to non-induced cells, indicating silent BESs are not fully derepressed. Interestingly, silent *VSG13* was more derepressed (21-33 fold more open) than *VSG2*, suggesting differential regulation between the silent BES upon overexpression induction. Also, not only the procyclin loci (*GPEET2*) chromatin become 5-8 fold more open as the highly regulated and silent MES (*mVSG639*) become derepressed to some extent (4-5 fold more open). Chromatin of genes from the active BES (*BSR* and *VSG9*) and rDNA (*18S*) remained unchanged, suggesting that these loci are fully loaded with TDP1. It was also detected a slight but significant increase in the chromatin accessibility in *β-tubulin* (~5-fold), a Pol II transcribed gene, a situation which is similar in *GAPDH* although not with statistical significance. This suggests a slight promiscuous DNA binding of TDP1, but it is not unexpected as TDP1 was detected in Pol II and Pol III transcribed genes notwithstanding low levels [14].

Can this increase in chromatin accessibility in silent BES lead to higher levels of transcription? To answer this question, the transcript levels of several genes were measured 24 and 48 hr after inducing TDP1 overexpression (Figure 20C and 20D). Once again, the phenotype presented by both clones is similar but more robust for POT1. Transcript levels of genes in silent BESs were increased, although the increase was higher in the promoter region (*luciferase*, ~70-200 fold increase) relative to the telomeric region (*VSG2* and *VSG13*, 3-5 fold and 20-30 fold increase respectively). These results are consistent with the chromatin accessibility measured by FAIRE, i.e. those genes close to the BES promoter become more derepressed. Transcripts of life-cycle regulated MES (*mVSG639*, 5-16 fold increase) and procyclin (*GPEET2*, 3-4 fold increase) genes were also slightly upregulated, indicating that all Pol I low-transcribed genes were derepressed. Conversely, active BES (*BSR* and *VSG9*) and rDNA (*18S*) genomic regions maintain their transcription status, which is consistent with no change in chromatin conformation (Figure 20A and 20B). Interestingly, transcript levels of Pol II transcribed genes (*β-tubulin* and *GAPDH)* were not affected by TDP1 overexpression (although POT2 presents a slight

significant *β-tubulin* downregulation), suggesting that the slight decondensation of chromatin at these loci is not sufficient to result in higher transcript levels.

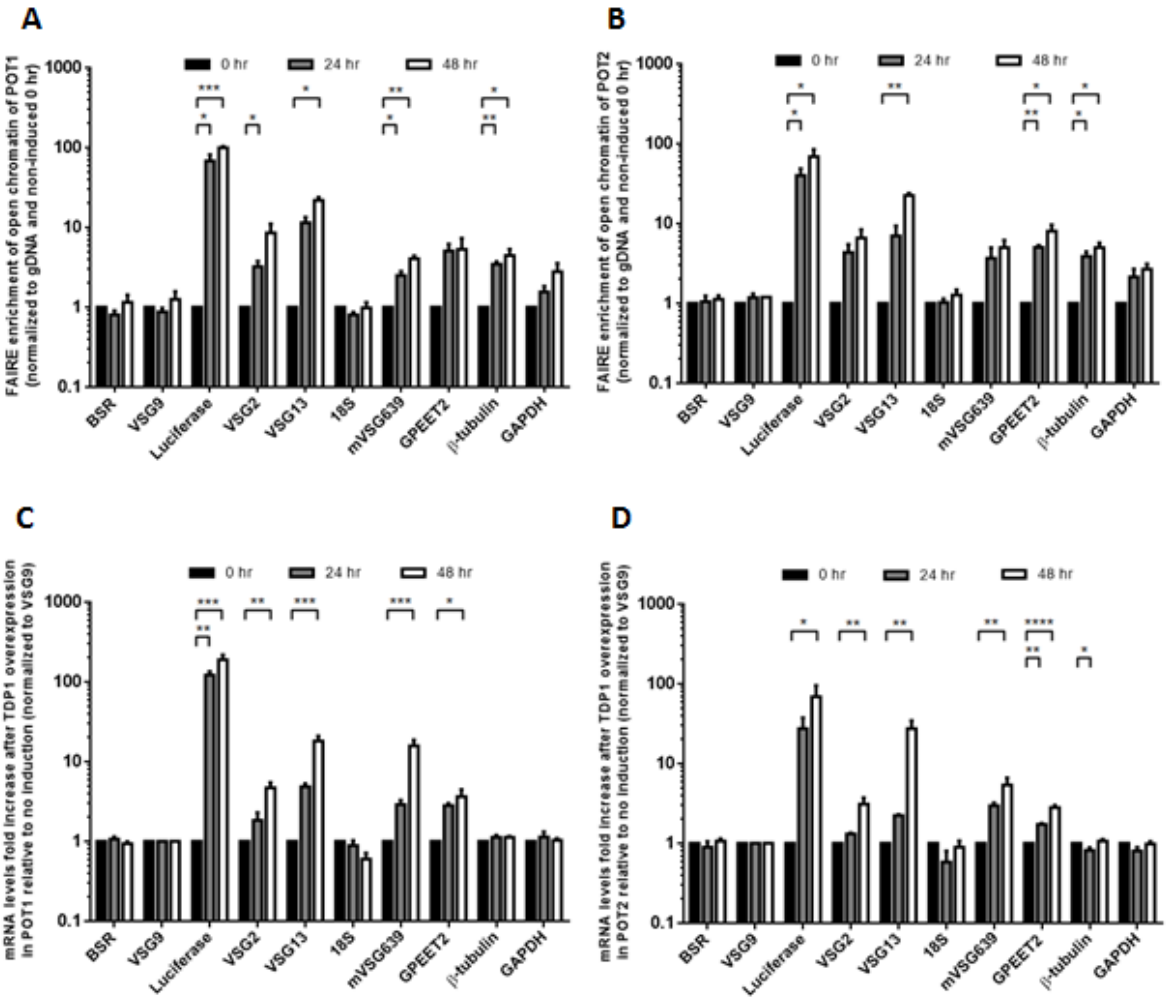


Figure 20 - TDP1 overexpression opens chromatin structure of many loci, but only Pol I silent genes become significantly up-regulated. (A and B) Chromatin conformation was measured by FAIRE at 24 and 48 hr after overexpression induction in POT1 **(A)** and POT2 **(B)**. DNA isolated by FAIRE was quantified by qPCR and normalized to gDNA copy number, ampicillin gene from DNA spike and non-induced cells. Four independent experiments were analyzed. **(C and D)** Fold-increase of mRNA levels fold after 24 and 48 hr of overexpression induction in POT1 **(C)** and POT2 **(D)**, measured by qPCR and normalized to non-induced cells and VSG9 gene. 3-4 independent experiments were performed. Statistical significance was determined by 1-way ANOVA with Bonferroni post-test comparison. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Thus, TDP1 is capable to open chromatin structure in all tested genes whose chromatin is organized in regularly spaced nucleosomes, but its effect is more pronounced in silent BESs. Consistently, TDP1 overexpression resulted in transcriptional

upregulation of Pol I genes but not of Pol II genes, consistent with the model that TDP1 is a transcriptional facilitator of Pol I but not of Pol II.

VSG monoallelic expression is disrupted when TDP1 is overexpressed

VSG monoallelic expression is a hallmark of antigenic variation. Although silent BES derepression has been previously observed in several *T. brucei* loss of function mutants, the expression of several VSGs at the cell surface has been detected only in a few cases [245, 258, 263, 264, 389]. Hence, it was tested if the transcriptional upregulation in silent BESs caused by TDP1 overexpression could be detected at the protein level.

First, the activity of the luciferase reporter, which is located downstream of the promoter of the silent BES1 was measured. Because the mRNA levels of this gene were highly increased (Figure 20C), it was expected its luminescence activity to be concomitantly higher (Figure 21A). Indeed, after 48 hr of TDP1 overexpression, the luciferase activity was higher in both clones and proportional to the increase at the mRNA levels, with POT1 displaying, as expected, a stronger phenotype. The fold-increase of luciferase activity is close to 300-fold for POT1, which is much higher than what has been detected in previous studies (maximum of 40-fold in TbRAP1 knockdown cells, maximum of 10-fold in histone H1 knockdown cells) [238, 258]. Although this increase is very significant, it is still below the levels luciferase exhibits when its gene is localized in the active BES, which typically are 1000-4000 fold higher [258].

To test the presence of derepressed VSGs at the surface upon TDP1 overexpression, an antibody specific for VSG13 was used [245] (Figure 21B). At 24 hr post-induction the expression of VSG13 is not significant on either clone (0.8-2.4%) but, at 48 hr, 9-24% of the cells already have intermediate levels of VSG13 (Figure 21C).

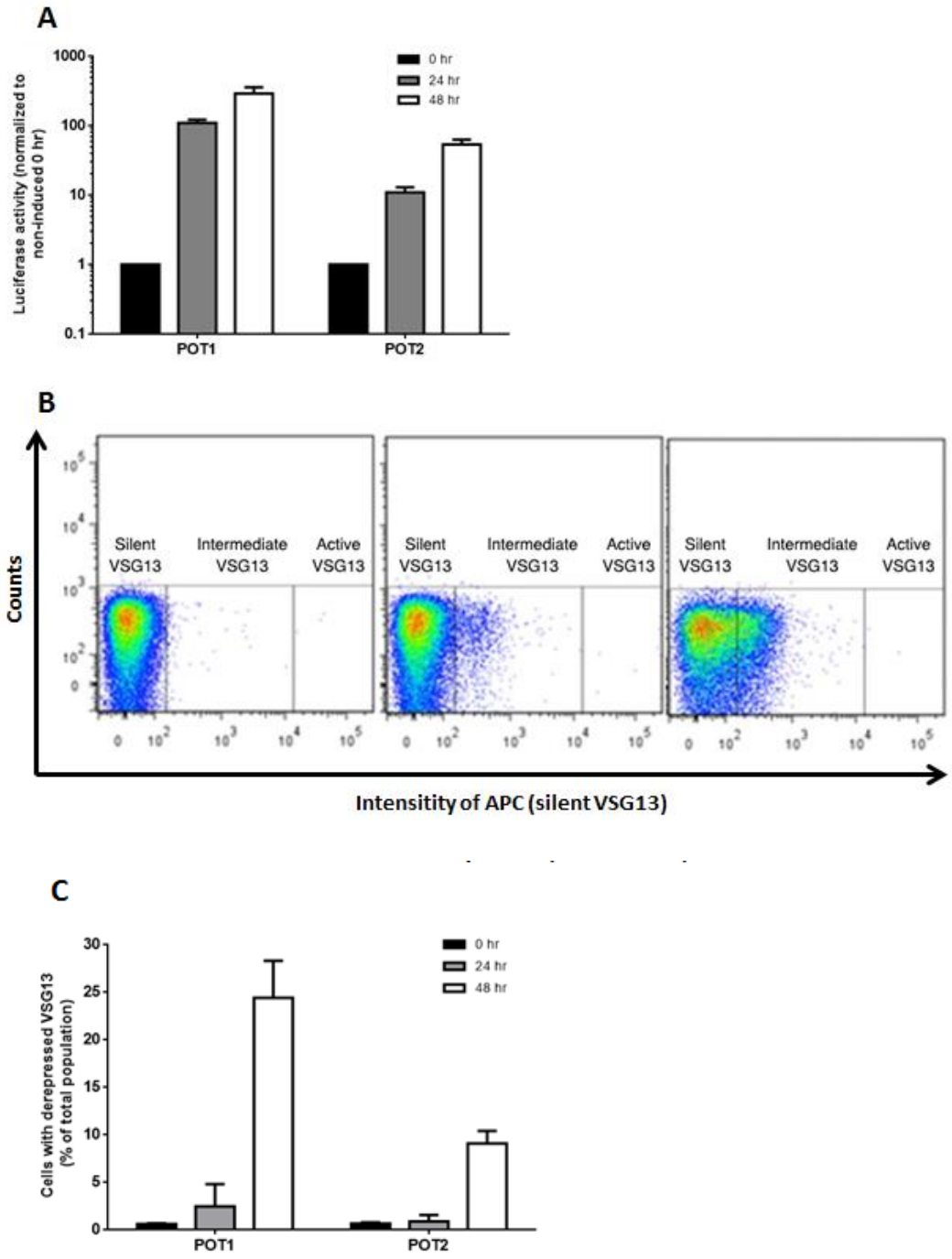


Figure 21 - TDP1 overexpression derepresses silent BESs and it compromises VSG monoallelic expression. (A) Luciferase activity fold increase of TDP1 overexpressed cells relative to non-induced cells after 24 and 48 hr of induction for POT1 and POT2. Five to six independent experiments were analyzed for each time-point. **(B)** Representative examples of FACS plots showing VSG13 expression at 24 and 48 hr of TDP1 overexpression in POT1. Definition of gates was done according to POT non-induced cells (for silent VSG13 gate), 17.13 cell-line [245] (for active VSG13 gate, data not shown), and intensities between (for intermediate VSG13 gate). **(C)** Quantification of VSG13 derepressed cells (intermediate gate) for each POT clone. Two independent experiments were analyzed

We conclude that overexpressing TDP1 leads to expression of more than one VSG at a given time, which means that monoallelic expression is compromised. To our knowledge, it is the first time that overexpression of a factor in *T. brucei* interferes with monoallelic expression.

DISCUSSION

It was previously shown that TDP1 is a component of the chromatin of actively transcribed Pol I loci, such as the active BES and rDNA repeats, and it is necessary to prevent deposition of nucleosomes at those loci [9, 10]. In this chapter, we describe the first steps in the characterization of a gain-of-function mutant of TDP1 and we showed that when TDP1 is overexpressed, the chromatin of silent BESs, MESs and procyclin genes becomes more open (Figures 20A and 20B) and transcription of these loci increases (Figure 20C and 20D). These results suggest that TDP1 can displace nucleosomes from chromatin fibers and this is sufficient to facilitate RNA polymerase I transcription.

The chromatin decondensation effect observed upon TDP1 overexpression is not homogeneous throughout the whole length of silent BESs. By FAIRE, we observed that the promoter region opened ~100-fold, while the decondensation at the corresponding VSG was less pronounced (around 5-10 fold). Consistently, we observed that the increase in transcript levels was higher close to the promoter than at the telomeric end of the silent BES (70-190 fold vs 3-27 fold, respectively). This suggests that more TDP1 was loaded in the promoter region of the silent BES chromatin than in the telomeric region, although ChIP experiments need to be conducted to show that TDP1 becomes loaded in these loci. One possible explanation is that chromatin in the BES promoter region is normally more open than telomeric region, which has been proposed before [9, 10]. Indeed, FAIRE and ChIP experiments have shown that BES promoter regions have fewer nucleosomes, which could facilitate the incorporation of TDP1 at these loci upon overexpression. Alternatively, some of the non-nucleosome components of chromatin that are enriched at the telomeric end of silent BESs (RAP1 and NLP) could hinder the binding of TDP1 in these regions [251, 258]. It is also interesting that the transcript fold increase upon TDP1 overexpression for the silent VSGs is comparable to the transcript fold increase for almost all silent VSGs that are upregulated upon VEX1 overexpression [270].

It is worth noticing that the transcript fold increase for the silent VSG is still low compared to the typical more than 10.000 fold increase that active VSG transcripts present relative to the silent VSGs (qPCR C_T ~14~-15 for active VSG9 and 29-30 for silent VSG2). These results were confirmed at the protein level. By FACS, we observed that TDP1 overexpression results in partial expression of a “silent” VSG13 (Figure 21). These

results are reminiscent to what was observed in DOT1B KO [245] and shows that TDP1 is sufficient to disrupt monoallelic expression. Unfortunately, due to absence of an antibody against VSG9, we could not test if the expression of the active VSG is maintained, as observed at the mRNA level. Moreover, it would be interesting to, not only try to observe expression of other silent VSGs, but also perform microscopy experiments against VSGs to assess that expression of silent VSG(s) occurs at the cell surface. Why silent BESs are never fully derepressed in mutant cell-lines remains unknown. In this study, it is possible that the levels of TDP1 were not high enough to displace all histones from silent BESs. But it is also possible that additional levels of regulation prevent a silent BES from being transcribed at the same levels as the active BES. For example, perhaps the derepressed silent BESs are not transcribed from the ESB and thus lack a limiting component for full transcription. Maybe overexpression of both VEX1 and TDP1 could lead to not only loss of the "winner takes all" mechanism allowing silent BES to be also transcribed (through VEX1) but in a more competent and efficient fashion (through TDP1) causing higher levels of derepression.

Other silent Pol I transcribed loci (procyclins and metacyclic VSGs) also become derepressed. For *metacyclic VSG639*, the chromatin accessibility increased 4-5 fold whereas transcript levels increased 5-16 fold. Metacyclic VSGs expression is regulated at the level of transcription [290] where its endogenous promoter is inactive in bloodstream forms [292], thus showing that TDP1 alone is able to drive transcription of life-cycle regulated surface protein genes. For *GPEET2* procyclin gene, the chromatin accessibility increased 5-8 fold whereas transcript levels increased 3-4 fold. Procyclins are around 9-fold less transcribed in bloodstream forms than in procyclic forms [274]. However, as procyclin mRNAs possess a highly unstable 3'UTR that reduces the stability of these transcripts [156], we cannot be certain if the transcription rate of procyclins due to TDP1 overexpression is higher than the detected through the steady-state mRNA levels and compensated by post-transcriptional degradation. These results show that TDP1 loading is able to open chromatin and facilitate transcription of multiple Pol I transcribed regions [14]. The transcript levels fold increases on both these loci are also within the same range as the ones observed for several *metacyclic VSG* and *procyclin* genes when VEX1 is overexpressed [270].

Upon TDP1 overexpression, the chromatin of Pol II transcribed genes *β-tubulin* and *GAPDH* became more accessible, especially *β-tubulin*. This suggests that TDP1, besides binding to Pol II transcribed DNA [14], may also remodel Pol II loci chromatin. In contrast to Pol I loci, these small alterations of chromatin conformation in Pol II loci, did not result in major detectable changes of transcript levels. This is consistent with a very slight reduced growth rate in POT clones, indicating that excess of TDP1 is not toxic to the cell and that loss of monoallelic expression does not compromise growth. Yeast Hmo1 also binds to Pol II transcribed loci [312, 390], but in contrast to *T. brucei* this binding contributes to Pol II transcription initiation [314, 315] and leads to a major growth defect [318]. Overexpression of TDP1 does not lead to chromatin accessibility and transcription changes in *18S* genes. It has also been previously showed that depletion of H1 leads to chromatin opening of several Pol I loci but not *18S* [238]. Thus, these observations suggest that rDNA genes in *T. brucei* may be all, or almost all, in an open chromatin conformation. Overall, these results indicate that TDP1 shares several functional features with UBF (structural component and transcriptional facilitator of Pol I loci) and Hmo1 (structural component of both Pol I and Pol II loci) but is not a *bona fide* orthologue of neither of them. It is possible that TDP1, being essential and one of the few HMGB proteins in *T. brucei*, it fulfills several essential roles not only in transcription, but also in other nuclear processes described for HMGB proteins, such as replication and DNA repair.

Expression of only one VSG at the cell surface of *T. brucei* is believed to be crucial to reduce the repertoire of antibodies generated by the host immune system. Switching to a new VSG enables survival and maintenance of infection. Using FACS, we detected no switchers in POT clones, suggesting overexpression of TDP1 does not lead to a dramatic increase in switching frequency. However, because procyclin and metacyclic VSGs were transcriptionally upregulated at the mRNA level, these non-variant proteins might be exposed at the cell surface of bloodstream forms. If these proteins are present at high enough levels and exposed to antibody recognition by the host immune system, we would expect the virulence of POT1 clones in mice to be compromised. It is also possible that differentiation of parasites from slender to procyclic forms in this monomorphic strain may also be compromised.

In the upcoming months, we expect to have strengthened functional and biological data to support our hypothesis that TDP1 overexpression leads to VSG monoallelic disruption as a result of the deregulation of silent BES chromatin status, which may impact the progression of an infection in mice. Thus, we anticipate the preparation of a manuscript that reveals the gain-of-function phenotype of the antigenic variation key player TDP1.

CHAPTER 4

GENERAL DISCUSSION

GENERAL DISCUSSION

Trypanosoma brucei parasites are able to establish an infection in the mammalian host for several months or even years. Such persistence is dependent of antigenic variation. One of the mechanisms used by the parasite to exchange the VSG surface coat consists in transcriptionally silence the active BES and activate a new BES [8]. The mechanism of VSG *in situ* switching is poorly understood. The steady-state chromatin structure of the active and silent BES has been characterized [9, 10] and several chromatin-associated factors involved in the maintenance of monoallelic expression have been identified (reviewed in [369], updated in Table 2). However, the sequence of events and the key players involved in an *in situ* switching were fully missing at the beginning of this dissertation. I studied how changes in chromatin structure and transcription are coordinated during *in situ* switching, which led to the identification of TDP1 (trypanosome DNA-binding protein 1) as a critical chromatin factor in switching and monoallelic expression.

The natural trigger that drives a VSG switching event is still currently unknown. It is unknown if changes occur first in the active or if it is a silent BES that becomes upregulated. It is also unknown what changes first: at the level of transcription or chromatin structure. Due to the low switching frequency both *in vivo* or *in vitro* [391, 392], it is extremely rare to observe a natural switching event, so the first question becomes difficult to tackle. Nonetheless, the active BES must also be shut down during an *in situ* switching. Similarly, during the differentiation process from bloodstream forms (BSFs) to procyclic forms (PFs), the active BES is shut down and another set of surface proteins, the procyclins, is expressed [59]. When differentiation in BSFs is triggered *in vitro*, the transcript levels decrease very steeply in the first 24 hr, but the chromatin conformation is still essentially open indicating that transcriptional silencing precedes chromatin changes. Although in a different biological context, it is tempting to think that in an *in situ* switching event the same observation could occur.

To induce a transcriptional silencing, a cell-line containing a tetracycline inducible system in the active BES promoter was employed. Similarly to differentiation, when BES silencing was induced, we observed that the transcription was rapidly blocked and

transcript levels were reduced in the first 8 hr. However, the active BES chromatin remained open as late as 24 hr post-silencing. Other features could also be observed at this specific time-point: i) only a small fraction of cells had already switched to another VSG; ii) half of transcriptionally silenced cells were already committed to switched although not completed the process; iii) cells that are not committed to switch are able to revert and re-initiate transcription of the active BES; iv) cells are probing silent BESs and can test more than one BES at the same time; v) Silent BES probing occurs not only in the promoter region but is extended until the VSG gene with silent VSG protein being detected. All these observations allowed us to propose a model in which the maintenance of the open chromatin in the active BES upon transcriptional silencing is important to allow parasites to test different BES for switching. If switching is not successful, this chromatin status may serve as facilitator to rapidly fire active BES transcription. At 24 hr, we observed an accumulation of G_2/M cells, which could be a consequence of reduced VSG transcript levels [349]. The parasites are probably urged to normalize VSG transcript and also protein levels.

During switching, what would be the advantage for the chromatin of the originally active BES to stay open for several hours? The maintenance of an open chromatin structure prior to a full commitment to a new VSG may serve to "test" silent BESs: i) compatibility of the ESAGs with the infected host as some ESAGs possess isoforms that may regulate growth in different mammalian hosts [94]; ii) VSG functionality, similarly to what is observed in the mammalian olfactory receptor (OR) genes in which the transcription and expression of a non-functional or mutated OR leads to switching to another *OR* until a functional OR protein elicits a feedback loop mechanism to select this single *OR* and prevent all the other from being expressed [341]. Although selected *OR* gene locus does not possess an open chromatin structure as BES, it is noteworthy that alleviation of heterochromatin structure (by means of chromatin marks) of silent *OR* genes is necessary for these to become activated and tested until a final choice is done [340]. However, while probing other BES, it is possible that the parasite does not *in situ* switch because it has already switched BES (and its corresponding VSGs) a few times. Thus, it would be interesting to unveil if upon an *in situ* switch, the newly silenced BES contains a specific mark or modification to avoid switching back to a VSG that has already

been expressed. For instance, base J is known to be present in silent BES but its functional role in these loci remains unclear. Thus, it would be interesting to evaluate by chromatin immunoprecipitation (ChIP) if base J serves as a marker for the non-transcribed intermediate state which is maintained if the BES becomes silenced or removed if the BES becomes re-transcribed. Cell-cycle stage dependency was also considered as a putative feature that could maintain the active BES in an open chromatin state. Contrary to yeast, which shows a replication-dependent nucleosome assembly in Pol I loci [221], the active BES is maintained in an open chromatin state throughout all cell-cycle stages. As the active BES is replicated earlier relative to the silent BES [361], synchronization and collection of cells at different time-points of the S stage to observe chromatin conformation in narrower windows of this cell-cycle stage should be inspected. Considering that for the parasite it might not be profitable to *in situ* switch BES too many times, it is possible that this mechanism of keeping BES chromatin open may also be useful for VSG switching by recombination. Thus, the fact that the active BES is being maintained in an open chromatin conformation under transcriptional silencing might serve to initiate gene conversion mechanisms to copy intact VSGs from the internal arrays or to generate mosaic VSGs.

A factor that presents a pivotal role in antigenic variation is TDP1. This dissertation presents evidence that TDP1 is not only important to maintain chromatin in an open conformation during an *in situ* switching, but also for monoallelic expression of a single VSG. TDP1 is a high mobility group box protein that had been previously shown to be a Pol I transcription facilitator [13, 14]. Depletion of TDP1 leads to compaction of chromatin and decreased transcript levels of Pol I loci [14]. In this work, we observed that in the absence of transcription, TDP1 is crucial to stabilize chromatin in an open conformation in the absence of transcription, similar to the yeast ortholog Hmo1 [221]. On the other hand, overexpression of TDP1 induced three different phenotypes: i) Pol II-transcribed regions possessed a slightly more open chromatin conformation but with no increase in the transcript levels; ii) the active BES and rDNA loci chromatin conformation remained unaltered as well as its transcript levels, suggesting that these loci remain with their chromatin fully open in wild-type parasites; iii) BESs, MESs and procyclin loci acquire a more open chromatin conformation and its transcript levels were also increased.

Nonetheless, expressed silent VSGs never reach high transcript levels, nor the active VSG transcript levels decrease, which indicates a crosstalk between BESs that maintains only one BES as the active one. Thus, the role of TDP1 in establishing and maintaining an open chromatin conformation of BES in BSFs may explain why this protein is essential in this life-cycle stage [14, 393].

However, TDP1 may not be as important in PFs as in BSFs. RNA interference (RNAi) against TDP1 in PFs does not lead to cell death [393], which may suggest that low levels of TDP1 are sufficient for the transcription of rDNA genes (if not all, at least some of the repeats) and for the procyclin locus. Nonetheless, as TDP1 is the only identified nuclear HMGB protein, other roles of TDP1 cannot be excluded which can even be differential across the trypanosome life-cycle stages.

Another interesting observation is how chromatin structure changes within the BES locus. When TDP1 was depleted using dsRNA, the telomeric region of the active BES becomes more affected and becomes repressed earlier than the promoter region. Likewise, when TDP1 was overexpressed the promoter region of the silent BESs acquired a more open chromatin conformation than the telomeric regions. This suggests that BES remodeling tends to protect the telomeric region of the BESs probably through the action of several identified factor such as RAP1, TRF or NUP1.

Interestingly, Baltram *et al*. have shown that upon overexpression of an ectopic VSG from an rDNA locus, the active BES becomes transcriptionally attenuated in a gradual pattern, i.e. VSG and more telomeric ESAGs are more transcriptionally down-regulated than the promoter region ESAGs [368]. This attenuation is dependent upon the chromatin-associated DOT1B histone methyltransferase, since ΔDOT1B cells do not suffer this gradual transcriptional effect and only the active VSG transcripts become depleted. It would be interesting to understand if the open chromatin maintenance in the active BES upon transcriptional silencing is also DOT1B-dependent. Thus, considering that *in situ* switching may be a concerted process of factors intervening in the active and silent BESs at the same time, it would be important to better unveil which chromatin factors may be associated with the active BES. Although previous attempts to immunoprecipitate endogenous TDP1 were unsuccessful to identify putative partners [14], TDP1 loading in

silent BES during overexpression should be confirmed by ChIP, possibly complemented by co-immunoprecipitation (co-IP) of TDP1 under the same conditions to attempt identifying its partners. It is interesting to notice that chromatin is maintained in an open conformation in two very different biological processes: differentiation and VSG switching. The Navarro lab showed that upon differentiation of BSFs, the active BES spatially migrates to the nuclear periphery after 6 hr of induction where heterochromatin and also silent BESs are localized in BSFs [11]. However, the chromatin is still maintained in an open conformation. It is thus possible that TDP1 also has a role in maintaining chromatin open in the active BES upon differentiation. However, considering that when stumpy forms are subjected to differentiation cues it only takes ~2-3 hr for full commitment to occur [49], it remains a mystery why after 24 hr of differentiation (even considering that we used monomorphic BSFs), the active BES still possesses this chromatin structure. It would also be interesting to perform immunofluorescence against TDP1 during the first 24 hr after inducing the transcriptional silencing to understand if the spatial changes also occur in the active BES.

What would be the result of overexpressing TDP1 in parasites during a mouse infection? With several VSGs expressed at the same time, how does the immune system respond? Can the infection be controlled or even eliminated? More than one VSG is expressed at the cell surface upon TDP1 overexpression but it is unknown how much is the contribution of the “silent” expressed VSGs relative to the total content of VSG. Is it small such that the “silent” VSGs are still untraceable to the immune system? Or are the levels of “silent” VSGs high enough to be recognized by antibodies? Experimentally addressing these questions would allow us to understand if the effect of TDP1 on antigenic variation actually compromises this survival mechanism.

Besides antigenic variation, it should also be assessed how overexpression of TDP1 affects differentiation to stumpy forms and progressively to procyclic forms and even to metacyclic forms is affected. Procyclin and metacyclic VSGs become transcriptionally upregulated, but it is still unclear if these proteins are also expressed at the parasite cell surface. If so, it would be interesting to determine if the densely packed VSG coat at the cells surface becomes disrupted with the presence of surface proteins other than bloodstream VSGs. Also, all BESs are silenced during differentiation to procyclic forms and

it is plausible that this process becomes undermined during TDP1 overexpression if BESs remain in transcriptionally upregulated, which could potentially interfere with parasite survival in the Tsetse midgut.

The explosion of the epigenetic field has made apparent that, in most eukaryotes, chromatin structure and its dynamics impose profound effects on almost all DNA-related processes, including transcription. This dissertation showed that chromatin structure plays a crucial role in the dynamics of antigenic variation of African trypanosomes, by allowing parasites to go through an intermediary state before parasites choose which BES/VSG will be activated. Although many questions remain unanswered, we have come a step closer to understanding the regulation of BES expression and switching.

Antigenic variation is central for African Trypanosomiasis. Advances in the understanding of the molecular mechanism underlying VSG gene regulation brings us a step closer to identifying potential drug targets. It would be interesting if we could lock parasites in a VSG switching intermediate state. Impossibility of switching, or incapacity of choosing a VSG may be detrimental for the parasite and make it more susceptible to elimination by the immune system.

REFERENCES

1. World Health Organization. *Epidemiology situation*. [cited 2016 8 Aug].
2. Trindade, S., et al., *Trypanosoma brucei Parasites Occupy and Functionally Adapt to the Adipose Tissue in Mice.* Cell Host Microbe, 2016. **19**(6): p. 837-48.
3. Franco, J.R., et al., *Epidemiology of human African trypanosomiasis.* Clin Epidemiol, 2014. **6**: p. 257-75.
4. Gunzl, A., et al., *RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in Trypanosoma brucei.* Eukaryot Cell, 2003. **2**(3): p. 542-51.
5. Bernards, A., et al., *Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3' end.* Cell, 1981. **27**(3 Pt 2): p. 497-505.
6. Pays, E., et al., *Telomeric reciprocal recombination as a possible mechanism for antigenic variation in trypanosomes.* Nature, 1985. **316**(6028): p. 562-4.
7. Bernards, A., et al., *Two modes of activation of a single surface antigen gene of Trypanosoma brucei.* Cell, 1984. **36**(1): p. 163-70.
8. Myler, P., et al., *Two mechanisms of expression of a predominant variant antigen gene of Trypanosoma brucei.* Nature, 1984. **309**(5965): p. 282-4.
9. Figueiredo, L.M. and G.A. Cross, *Nucleosomes are depleted at the VSG expression site transcribed by RNA polymerase I in African trypanosomes.* Eukaryot Cell, 2010. **9**(1): p. 148-54.
10. Stanne, T.M. and G. Rudenko, *Active VSG expression sites in Trypanosoma brucei are depleted of nucleosomes.* Eukaryot Cell, 2010. **9**(1): p. 136-47.
11. Landeira, D. and M. Navarro, *Nuclear repositioning of the VSG promoter during developmental silencing in Trypanosoma brucei.* J Cell Biol, 2007. **176**(2): p. 133-9.
12. Wang, Q.P., T. Kawahara, and D. Horn, *Histone deacetylases play distinct roles in telomeric VSG expression site silencing in African trypanosomes.* Mol Microbiol, 2010. **77**(5): p. 1237-45.
13. Erondu, N.E. and J.E. Donelson, *Differential Expression of 2 Messenger-Rnas from a Single Gene Encoding an Hmg1-Like DNA-Binding Protein of African Trypanosomes.* Molecular and Biochemical Parasitology, 1992. **51**(1): p. 111-118.
14. Narayanan, M.S. and G. Rudenko, *TDP1 is an HMG chromatin protein facilitating RNA polymerase I transcription in African trypanosomes.* Nucleic Acids Res, 2013. **41**(5): p. 2981-92.

15. Brun, R. and J. Blum, *Human African trypanosomiasis.* Infect Dis Clin North Am, 2012. **26**(2): p. 261-73.
16. Steverding, D., *The history of African trypanosomiasis.* Parasit Vectors, 2008. **1**(1): p. 3.
17. Truc, P., et al., *Atypical human infections by animal trypanosomes.* PLoS Negl Trop Dis, 2013. **7**(9): p. e2256.
18. Brandenberger, G., et al., *Disruption of endocrine rhythms in sleeping sickness with preserved relationship between hormonal pulsatility and the REM-NREM sleep cycles.* J Biol Rhythms, 1996. **11**(3): p. 258-67.
19. Buguet, A., et al., *Sleep structure: a new diagnostic tool for stage determination in sleeping sickness.* Acta Trop, 2005. **93**(1): p. 107-17.
20. Bonnet, J., C. Boudot, and B. Courtioux, *Overview of the Diagnostic Methods Used in the Field for Human African Trypanosomiasis: What Could Change in the Next Years?* Biomed Res Int, 2015. **2015**: p. 583262.
21. Douzery, E.J., et al., *The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils?* Proc Natl Acad Sci U S A, 2004. **101**(43): p. 15386-91.
22. Cavalier-Smith, T., *Kingdoms Protozoa and Chromista and the eozoan root of the eukaryotic tree.* Biol Lett, 2010. **6**(3): p. 342-5.
23. Berriman, M., et al., *The genome of the African trypanosome Trypanosoma brucei.* Science, 2005. **309**(5733): p. 416-22.
24. Ferguson, M.A., M.G. Low, and G.A. Cross, *Glycosyl-sn-1,2-dimyristylphosphatidylinositol is covalently linked to Trypanosoma brucei variant surface glycoprotein.* J Biol Chem, 1985. **260**(27): p. 14547-55.
25. Benne, R., et al., *Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA.* Cell, 1986. **46**(6): p. 819-26.
26. Boothroyd, J.C. and G.A. Cross, *Transcripts coding for variant surface glycoproteins of Trypanosoma brucei have a short, identical exon at their 5' end.* Gene, 1982. **20**(2): p. 281-9.
27. Murphy, W.J., K.P. Watkins, and N. Agabian, *Identification of a novel Y branch structure as an intermediate in trypanosome mRNA processing: evidence for trans splicing.* Cell, 1986. **47**(4): p. 517-25.
28. Sutton, R.E. and J.C. Boothroyd, *Evidence for trans splicing in trypanosomes.* Cell, 1986. **47**(4): p. 527-35.

29. Gommers-Ampt, J., J. Lutgerink, and P. Borst, *A novel DNA nucleotide in Trypanosoma brucei only present in the mammalian phase of the life-cycle.* Nucleic Acids Res, 1991. **19**(8): p. 1745-51.
30. Gommers-Ampt, J.H., et al., *beta-D-glucosyl-hydroxymethyluracil: a novel modified base present in the DNA of the parasitic protozoan T. brucei.* Cell, 1993. **75**(6): p. 1129-36.
31. Opperdoes, F.R. and P. Borst, *Localization of nine glycolytic enzymes in a microbody-like organelle in Trypanosoma brucei: the glycosome.* FEBS Lett, 1977. **80**(2): p. 360-4.
32. Opperdoes, F.R., et al., *Localization of glycerol-3-phosphate oxidase in the mitochondrion and particulate NAD+-linked glycerol-3-phosphate dehydrogenase in the microbodies of the bloodstream form to Trypanosoma brucei.* Eur J Biochem, 1977. **76**(1): p. 29-39.
33. Vickerman, K., *On the surface coat and flagellar adhesion in trypanosomes.* J Cell Sci, 1969. **5**(1): p. 163-93.
34. Ferrante, A. and A.C. Allison, *Alternative pathway activation of complement by African trypanosomes lacking a glycoprotein coat.* Parasite Immunol, 1983. **5**(5): p. 491-8.
35. Modespacher, U.P., W. Rudin, and H. Hecker, *Surface coat synthesis and turnover from epimastigote to bloodstream forms of Trypanosoma brucei.* Acta Trop, 1991. **50**(1): p. 67-78.
36. Caljon, G., et al., *The Dermis as a Delivery Site of Trypanosoma brucei for Tsetse Flies.* PLoS Pathog, 2016. **12**(7): p. e1005744.
37. Langousis, G. and K.L. Hill, *Motility and more: the flagellum of Trypanosoma brucei.* Nat Rev Microbiol, 2014. **12**(7): p. 505-18.
38. Reuner, B., et al., *Cell density triggers slender to stumpy differentiation of Trypanosoma brucei bloodstream forms in culture.* Mol Biochem Parasitol, 1997. **90**(1): p. 269-80.
39. Vassella, E., et al., *Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway.* J Cell Sci, 1997. **110 (Pt 21)**: p. 2661-71.
40. Mony, B.M., et al., *Genome-wide dissection of the quorum sensing signalling pathway in Trypanosoma brucei.* Nature, 2014. **505**(7485): p. 681-5.
41. Dean, S., et al., *A surface transporter family conveys the trypanosome differentiation signal.* Nature, 2009. **459**(7244): p. 213-7.
42. Vickerman, K., *Polymorphism and mitochondrial activity in sleeping sickness trypanosomes.* Nature, 1965. **208**(5012): p. 762-6.
43. Jensen, B.C., et al., *Widespread variation in transcript abundance within and across developmental stages of Trypanosoma brucei.* BMC Genomics, 2009. **10**: p. 482.

44. Kabani, S., et al., *Genome-wide expression profiling of in vivo-derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in Trypanosoma brucei.* BMC Genomics, 2009. **10**: p. 427.

45. Nilsson, D., et al., *Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of Trypanosoma brucei.* PLoS Pathog, 2010. **6**(8): p. e1001037.

46. Gunasekera, K., et al., *Proteome remodelling during development from blood to insect-form Trypanosoma brucei quantified by SILAC and mass spectrometry.* BMC Genomics, 2012. **13**: p. 556.

47. Robertson, M., *Notes on the polymorphism of Trypanosoma gambiense in the blood and its relation to the exogenous cycle in Glossina palpalis.* Proceedings of the Royal Society of London Series B-Containing Papers of a Biological Character, 1912. **85**(582): p. 527-539.

48. Czichos, J., C. Nonnengaesser, and P. Overath, *Trypanosoma brucei: cis-aconitate and temperature reduction as triggers of synchronous transformation of bloodstream to procyclic trypomastigotes in vitro.* Exp Parasitol, 1986. **62**(2): p. 283-91.

49. Domingo-Sananes, M.R., et al., *Molecular control of irreversible bistability during trypanosome developmental commitment.* J Cell Biol, 2015. **211**(2): p. 455-68.

50. Queiroz, R., et al., *Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons.* BMC Genomics, 2009. **10**: p. 495.

51. Brems, S., et al., *The transcriptomes of Trypanosoma brucei Lister 427 and TREU927 bloodstream and procyclic trypomastigotes.* Mol Biochem Parasitol, 2005. **139**(2): p. 163-72.

52. Siegel, T.N., et al., *Genome-wide analysis of mRNA abundance in two life-cycle stages of Trypanosoma brucei and identification of splicing and polyadenylation sites.* Nucleic Acids Res, 2010. **38**(15): p. 4946-57.

53. Veitch, N.J., et al., *Digital gene expression analysis of two life cycle stages of the human-infective parasite, Trypanosoma brucei gambiense reveals differentially expressed clusters of co-regulated genes.* BMC Genomics, 2010. **11**: p. 124.

54. Urbaniak, M.D., M.L. Guther, and M.A. Ferguson, *Comparative SILAC proteomic analysis of Trypanosoma brucei bloodstream and procyclic lifecycle stages.* PLoS One, 2012. **7**(5): p. e36619.

55. Butter, F., et al., *Comparative proteomics of two life cycle stages of stable isotope-labeled Trypanosoma brucei reveals novel components of the parasite's host adaptation machinery.* Mol Cell Proteomics, 2013. **12**(1): p. 172-9.

56. Acosta-Serrano, A., et al., *The surface coat of procyclic Trypanosoma brucei: programmed expression and proteolytic cleavage of procyclin in the tsetse fly.* Proc Natl Acad Sci U S A, 2001. **98**(4): p. 1513-8.

57. Jackson, D.G., M.J. Owen, and H.P. Voorheis, *A new method for the rapid purification of both the membrane-bound and released forms of the variant surface glycoprotein from Trypanosoma brucei.* Biochem J, 1985. **230**(1): p. 195-202.

58. Gruszynski, A.E., et al., *Regulation of surface coat exchange by differentiating African trypanosomes.* Mol Biochem Parasitol, 2006. **147**(2): p. 211-23.

59. Roditi, I., et al., *Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of Trypanosoma brucei.* J Cell Biol, 1989. **108**(2): p. 737-46.

60. Vassella, E., et al., *A major surface glycoprotein of trypanosoma brucei is expressed transiently during development and can be regulated post-transcriptionally by glycerol or hypoxia.* Genes Dev, 2000. **14**(5): p. 615-26.

61. Van Den Abbeele, J., et al., *Trypanosoma brucei spp. development in the tsetse fly: characterization of the post-mesocyclic stages in the foregut and proboscis.* Parasitology, 1999. **118 (Pt 5)**: p. 469-78.

62. Urwyler, S., et al., *A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of Trypanosoma brucei.* Mol Microbiol, 2007. **63**(1): p. 218-28.

63. Gibson, W., et al., *Analysis of a cross between green and red fluorescent trypanosomes.* Biochem Soc Trans, 2006. **34**(Pt 4): p. 557-9.

64. Peacock, L., et al., *Identification of the meiotic life cycle stage of Trypanosoma brucei in the tsetse fly.* Proc Natl Acad Sci U S A, 2011. **108**(9): p. 3671-6.

65. Tetley, L., et al., *Onset of expression of the variant surface glycoproteins of Trypanosoma brucei in the tsetse fly studied using immunoelectron microscopy.* J Cell Sci, 1987. **87 (Pt 2)**: p. 363-72.

66. Brun, R. and M. Schonenberger, *Stimulating effect of citrate and cis-Aconitate on the transformation of Trypanosoma brucei bloodstream forms to procyclic forms in vitro.* Z Parasitenkd, 1981. **66**(1): p. 17-24.

67. Breidbach, T., E. Ngazoa, and D. Steverding, *Trypanosoma brucei: in vitro slender-to-stumpy differentiation of culture-adapted, monomorphic bloodstream forms.* Exp Parasitol, 2002. **101**(4): p. 223-30.

68. Turner, C.M., N. Aslam, and C. Dye, *Replication, differentiation, growth and the virulence of Trypanosoma brucei infections.* Parasitology, 1995. **111 (Pt 3)**: p. 289-300.

69. Kolev, N.G., et al., *Developmental progression to infectivity in Trypanosoma brucei triggered by an RNA-binding protein.* Science, 2012. **338**(6112): p. 1352-3.

70. Ramey-Butler, K., et al., *Synchronous expression of individual metacyclic variant surface glycoprotein genes in Trypanosoma brucei.* Mol Biochem Parasitol, 2015. **200**(1-2): p. 1-4.
71. Ogbadoyi, E., et al., *Architecture of the Trypanosoma brucei nucleus during interphase and mitosis.* Chromosoma, 2000. **108**(8): p. 501-13.
72. Melville, S.E., et al., *The molecular karyotype of the megabase chromosomes of Trypanosoma brucei stock 427.* Mol Biochem Parasitol, 2000. **111**(2): p. 261-73.
73. Van der Ploeg, L.H., et al., *Chromosomes of kinetoplastida.* EMBO J, 1984. **3**(13): p. 3109-15.
74. Alsford, N.S., et al., *The identification of circular extrachromosomal DNA in the nuclear genome of Trypanosoma brucei.* Mol Microbiol, 2003. **47**(2): p. 277-89.
75. Imboden, M.A., et al., *Transcription of the intergenic regions of the tubulin gene cluster of Trypanosoma brucei: evidence for a polycistronic transcription unit in a eukaryote.* Nucleic Acids Res, 1987. **15**(18): p. 7357-68.
76. Wickstead, B., K. Ersfeld, and K. Gull, *The small chromosomes of Trypanosoma brucei involved in antigenic variation are constructed around repetitive palindromes.* Genome Res, 2004. **14**(6): p. 1014-24.
77. Berriman, M., et al., *The architecture of variant surface glycoprotein gene expression sites in Trypanosoma brucei.* Mol Biochem Parasitol, 2002. **122**(2): p. 131-40.
78. Van der Ploeg, L.H., A.Y. Liu, and P. Borst, *Structure of the growing telomeres of Trypanosomes.* Cell, 1984. **36**(2): p. 459-68.
79. Van der Ploeg, L.H., et al., *Antigenic variation in Trypanosoma brucei analyzed by electrophoretic separation of chromosome-sized DNA molecules.* Cell, 1984. **37**(1): p. 77-84.
80. Fairlamb, A.H., et al., *Isolation and characterization of kinetoplast DNA from bloodstream form of Trypanosoma brucei.* J Cell Biol, 1978. **76**(2): p. 293-309.
81. Chen, K.K. and J.E. Donelson, *Sequences of two kinetoplast DNA minicircles of Tryptanosoma brucei.* Proc Natl Acad Sci U S A, 1980. **77**(5): p. 2445-9.
82. Sloof, P., et al., *The nucleotide sequence of the variable region in Trypanosoma brucei completes the sequence analysis of the maxicircle component of mitochondrial kinetoplast DNA.* Mol Biochem Parasitol, 1992. **56**(2): p. 289-99.
83. Eperon, I.C., et al., *The major transcripts of the kinetoplast DNA of Trypanosoma brucei are very small ribosomal RNAs.* Nucleic Acids Res, 1983. **11**(1): p. 105-25.
84. Benne, R., et al., *The nucleotide sequence of a segment of Trypanosoma brucei mitochondrial maxi-circle DNA that contains the gene for apocytochrome b and some unusual unassigned reading frames.* Nucleic Acids Res, 1983. **11**(20): p. 6925-41.

85. Hensgens, L.A., et al., *The sequence of the gene for cytochrome c oxidase subunit I, a frameshift containing gene for cytochrome c oxidase subunit II and seven unassigned reading frames in Trypanosoma brucei mitochondrial maxi-circle DNA.* Nucleic Acids Res, 1984. **12**(19): p. 7327-44.

86. Rohrer, S.P., et al., *Transcription of kinetoplast DNA minicircles.* Cell, 1987. **49**(5): p. 625-32.

87. Corell, R.A., et al., *Trypanosoma brucei minicircles encode multiple guide RNAs which can direct editing of extensively overlapping sequences.* Nucleic Acids Res, 1993. **21**(18): p. 4313-20.

88. Williams, R.O., J.R. Young, and P.A. Majiwa, *Genomic environment of T. brucei VSG genes: presence of a minichromosome.* Nature, 1982. **299**(5882): p. 417-21.

89. Becker, M., et al., *Isolation of the repertoire of VSG expression site containing telomeres of Trypanosoma brucei 427 using transformation-associated recombination in yeast.* Genome Res, 2004. **14**(11): p. 2319-29.

90. Taylor, J.E. and G. Rudenko, *Switching trypanosome coats: what's in the wardrobe?* Trends Genet, 2006. **22**(11): p. 614-20.

91. Cross, G.A., H.S. Kim, and B. Wickstead, *Capturing the variant surface glycoprotein repertoire (the VSGnome) of Trypanosoma brucei Lister 427.* Mol Biochem Parasitol, 2014.

92. Turner, C.M., et al., *An estimate of the size of the metacyclic variable antigen repertoire of Trypanosoma brucei rhodesiense.* Parasitology, 1988. **97 (Pt 2)**: p. 269-76.

93. Hertz-Fowler, C., et al., *Telomeric expression sites are highly conserved in Trypanosoma brucei.* PLoS One, 2008. **3**(10): p. e3527.

94. Pays, E., et al., *The VSG expression sites of Trypanosoma brucei: multipurpose tools for the adaptation of the parasite to mammalian hosts.* Mol Biochem Parasitol, 2001. **114**(1): p. 1-16.

95. Xong, H.V., et al., *A VSG expression site-associated gene confers resistance to human serum in Trypanosoma rhodesiense.* Cell, 1998. **95**(6): p. 839-46.

96. De Greef, C. and R. Hamers, *The serum resistance-associated (SRA) gene of Trypanosoma brucei rhodesiense encodes a variant surface glycoprotein-like protein.* Mol Biochem Parasitol, 1994. **68**(2): p. 277-84.

97. Pays, E., et al., *The molecular arms race between African trypanosomes and humans.* Nat Rev Microbiol, 2014.

98. Alarcon, C.M., et al., *A monocistronic transcript for a trypanosome variant surface glycoprotein.* Mol Cell Biol, 1994. **14**(8): p. 5579-91.

99. Ginger, M.L., et al., *Ex vivo and in vitro identification of a consensus promoter for VSG genes expressed by metacyclic-stage trypanosomes in the tsetse fly.* Eukaryot Cell, 2002. **1**(6): p. 1000-9.

100. Pedram, M. and J.E. Donelson, *The anatomy and transcription of a monocistronic expression site for a metacyclic variant surface glycoprotein gene in Trypanosoma brucei.* J Biol Chem, 1999. **274**(24): p. 16876-83.

101. Paindavoine, P., et al., *A gene from the variant surface glycoprotein expression site encodes one of several transmembrane adenylate cyclases located on the flagellum of Trypanosoma brucei.* Mol Cell Biol, 1992. **12**(3): p. 1218-25.

102. Pays, E., et al., *The genes and transcripts of an antigen gene expression site from T. brucei.* Cell, 1989. **57**(5): p. 835-45.

103. Barker, A.R., et al., *Bioinformatic insights to the ESAG5 and GRESAG5 gene families in kinetoplastid parasites.* Mol Biochem Parasitol, 2008. **162**(2): p. 112-22.

104. Steverding, D., et al., *ESAG 6 and 7 products of Trypanosoma brucei form a transferrin binding protein complex.* Eur J Cell Biol, 1994. **64**(1): p. 78-87.

105. Hoek, M., M. Engstler, and G.A. Cross, *Expression-site-associated gene 8 (ESAG8) of Trypanosoma brucei is apparently essential and accumulates in the nucleolus.* J Cell Sci, 2000. **113 (Pt 22)**: p. 3959-68.

106. Hoek, M., T. Zanders, and G.A. Cross, *Trypanosoma brucei expression-site-associated-gene-8 protein interacts with a Pumilio family protein.* Mol Biochem Parasitol, 2002. **120**(2): p. 269-83.

107. Barnwell, E.M., et al., *Developmental regulation and extracellular release of a VSG expression-site-associated gene product from Trypanosoma brucei bloodstream forms.* J Cell Sci, 2010. **123**(Pt 19): p. 3401-11.

108. Gottesdiener, K.M., *A new VSG expression site-associated gene (ESAG) in the promoter region of Trypanosoma brucei encodes a protein with 10 potential transmembrane domains.* Mol Biochem Parasitol, 1994. **63**(1): p. 143-51.

109. Siegel, T.N., et al., *Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei.* Genes Dev, 2009. **23**(9): p. 1063-76.

110. Kolev, N.G., et al., *The transcriptome of the human pathogen Trypanosoma brucei at single-nucleotide resolution.* PLoS Pathog, 2010. **6**(9): p. e1001090.

111. Clayton, C.E., *Life without transcriptional control? From fly to man and back again.* EMBO J, 2002. **21**(8): p. 1881-8.

112. Gilinger, G. and V. Bellofatto, *Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms.* Nucleic Acids Res, 2001. **29**(7): p. 1556-64.

113. Gunzl, A., et al., *Transcription of the Trypanosoma brucei spliced leader RNA gene is dependent only on the presence of upstream regulatory elements.* Mol Biochem Parasitol, 1997. **85**(1): p. 67-76.

114. Mair, G., et al., *A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA.* RNA, 2000. **6**(2): p. 163-9.

115. Serpeloni, M., et al., *An essential nuclear protein in trypanosomes is a component of mRNA transcription/export pathway.* PLoS One, 2011. **6**(6): p. e20730.

116. Schwede, A., et al., *The role of deadenylation in the degradation of unstable mRNAs in trypanosomes.* Nucleic Acids Res, 2009. **37**(16): p. 5511-28.

117. Vasquez, J.J., et al., *Comparative ribosome profiling reveals extensive translational complexity in different Trypanosoma brucei life cycle stages.* Nucleic Acids Res, 2014. **42**(6): p. 3623-37.

118. Das, A., et al., *Biochemical characterization of Trypanosoma brucei RNA polymerase II.* Mol Biochem Parasitol, 2006. **150**(2): p. 201-10.

119. Devaux, S., et al., *Characterization of RNA polymerase II subunits of Trypanosoma brucei.* Mol Biochem Parasitol, 2006. **148**(1): p. 60-8.

120. Devaux, S., et al., *Diversification of function by different isoforms of conventionally shared RNA polymerase subunits.* Mol Biol Cell, 2007. **18**(4): p. 1293-301.

121. Bayele, H.K., *Trypanosoma brucei: a putative RNA polymerase II promoter.* Exp Parasitol, 2009. **123**(4): p. 313-8.

122. Wright, J.R., T.N. Siegel, and G.A. Cross, *Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in Trypanosoma brucei.* Mol Biochem Parasitol, 2010. **172**(2): p. 141-4.

123. Reynolds, D., et al., *Histone H3 Variant Regulates RNA Polymerase II Transcription Termination and Dual Strand Transcription of siRNA Loci in Trypanosoma brucei.* PLoS Genet, 2016. **12**(1): p. e1005758.

124. Cliffe, L.J., et al., *Two thymidine hydroxylases differentially regulate the formation of glucosylated DNA at regions flanking polymerase II polycistronic transcription units throughout the genome of Trypanosoma brucei.* Nucleic Acids Res, 2010. **38**(12): p. 3923-35.

125. van Leeuwen, F., et al., *Biosynthesis and function of the modified DNA base beta-D-glucosyl-hydroxymethyluracil in Trypanosoma brucei.* Mol Cell Biol, 1998. **18**(10): p. 5643-51.

126. Yu, Z., et al., *The protein that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase.* Nucleic Acids Res, 2007. **35**(7): p. 2107-15.

127. Sekar, A., et al., *Tb927.10.6900 encodes the glucosyltransferase involved in synthesis of base J in Trypanosoma brucei.* Mol Biochem Parasitol, 2014.

128. van Luenen, H.G., et al., *Glucosylated hydroxymethyluracil, DNA base J, prevents transcriptional readthrough in Leishmania.* Cell, 2012. **150**(5): p. 909-21.

129. Reynolds, D., et al., *Regulation of transcription termination by glucosylated hydroxymethyluracil, base J, in Leishmania major and Trypanosoma brucei.* Nucleic Acids Res, 2014.

130. Schimanski, B., et al., *A TFIIB-like protein is indispensable for spliced leader RNA gene transcription in Trypanosoma brucei.* Nucleic Acids Res, 2006. **34**(6): p. 1676-84.

131. Ruan, J.P., et al., *Functional characterization of a Trypanosoma brucei TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes.* Mol Cell Biol, 2004. **24**(21): p. 9610-8.

132. Das, A., et al., *Trypanosomal TBP functions with the multisubunit transcription factor tSNAP to direct spliced-leader RNA gene expression.* Mol Cell Biol, 2005. **25**(16): p. 7314-22.

133. Schimanski, B., T.N. Nguyen, and A. Gunzl, *Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in Trypanosoma brucei.* Mol Cell Biol, 2005. **25**(16): p. 7303-13.

134. Lecordier, L., et al., *Characterization of a TFIIH homologue from Trypanosoma brucei.* Mol Microbiol, 2007. **64**(5): p. 1164-81.

135. Lee, J.H., et al., *Spliced leader RNA gene transcription in Trypanosoma brucei requires transcription factor TFIIH.* Eukaryot Cell, 2007. **6**(4): p. 641-9.

136. Cordingley, J.S., *Nucleotide sequence of the 5S ribosomal RNA gene repeat of Trypanosoma brucei.* Mol Biochem Parasitol, 1985. **17**(3): p. 321-30.

137. Michaeli, S., et al., *The 7SL RNA homologue of Trypanosoma brucei is closely related to mammalian 7SL RNA.* Mol Biochem Parasitol, 1992. **51**(1): p. 55-64.

138. Fantoni, A., A.O. Dare, and C. Tschudi, *RNA polymerase III-mediated transcription of the trypanosome U2 small nuclear RNA gene is controlled by both intragenic and extragenic regulatory elements.* Mol Cell Biol, 1994. **14**(3): p. 2021-8.

139. Nakaar, V., et al., *Upstream tRNA genes are essential for expression of small nuclear and cytoplasmic RNA genes in trypanosomes.* Mol Cell Biol, 1994. **14**(10): p. 6736-42.

140. Daniels, J.P., K. Gull, and B. Wickstead, *Cell biology of the trypanosome genome.* Microbiol Mol Biol Rev, 2010. **74**(4): p. 552-69.

141. Kelly, S., B. Wickstead, and K. Gull, *An in silico analysis of trypanosomatid RNA polymerases: insights into their unusual transcription.* Biochem Soc Trans, 2005. **33**(Pt 6): p. 1435-7.

142. Das, A., M. Banday, and V. Bellofatto, *RNA polymerase transcription machinery in trypanosomes.* Eukaryot Cell, 2008. **7**(3): p. 429-34.

143. Huang, J. and L.H. van der Ploeg, *Maturation of polycistronic pre-mRNA in Trypanosoma brucei: analysis of trans splicing and poly(A) addition at nascent RNA transcripts from the hsp70 locus.* Mol Cell Biol, 1991. **11**(6): p. 3180-90.

144. Fadda, A., et al., *Transcriptome-wide analysis of trypanosome mRNA decay reveals complex degradation kinetics and suggests a role for co-transcriptional degradation in determining mRNA levels.* Mol Microbiol, 2014. **94**(2): p. 307-26.

145. Bangs, J.D., et al., *Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides.* J Biol Chem, 1992. **267**(14): p. 9805-15.

146. Michaeli, S., *Trans-splicing in trypanosomes: machinery and its impact on the parasite transcriptome.* Future Microbiol, 2011. **6**(4): p. 459-74.

147. Huang, J. and L.H. Van der Ploeg, *Requirement of a polypyrimidine tract for trans-splicing in trypanosomes: discriminating the PARP promoter from the immediately adjacent 3' splice acceptor site.* EMBO J, 1991. **10**(12): p. 3877-85.

148. Lopez-Estrano, C., C. Tschudi, and E. Ullu, *Exonic sequences in the 5' untranslated region of alpha-tubulin mRNA modulate trans splicing in Trypanosoma brucei.* Mol Cell Biol, 1998. **18**(8): p. 4620-8.

149. Patzelt, E., K.L. Perry, and N. Agabian, *Mapping of branch sites in trans-spliced pre-mRNAs of Trypanosoma brucei.* Mol Cell Biol, 1989. **9**(10): p. 4291-7.

150. Koch, H., et al., *The polyadenylation complex of Trypanosoma brucei: Characterization of the functional poly(A) polymerase.* RNA Biol, 2016. **13**(2): p. 221-31.

151. Matthews, K.R., C. Tschudi, and E. Ullu, *A common pyrimidine-rich motif governs trans-splicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes.* Genes Dev, 1994. **8**(4): p. 491-501.

152. Siegel, T.N., et al., *Gene expression in Trypanosoma brucei: lessons from high-throughput RNA sequencing.* Trends Parasitol, 2011. **27**(10): p. 434-41.

153. Haanstra, J.R., et al., *Control and regulation of gene expression: quantitative analysis of the expression of phosphoglycerate kinase in bloodstream form Trypanosoma brucei.* J Biol Chem, 2008. **283**(5): p. 2495-507.

154. Hotz, H.R., et al., *Role of 3'-untranslated regions in the regulation of hexose transporter mRNAs in Trypanosoma brucei.* Mol Biochem Parasitol, 1995. **75**(1): p. 1-14.

155. Blattner, J. and C.E. Clayton, *The 3'-untranslated regions from the Trypanosoma brucei phosphoglycerate kinase-encoding genes mediate developmental regulation.* Gene, 1995. **162**(1): p. 153-6.

156. Hotz, H.R., et al., *Mechanisms of developmental regulation in Trypanosoma brucei: a polypyrimidine tract in the 3'-untranslated region of a surface protein mRNA affects RNA abundance and translation.* Nucleic Acids Res, 1997. **25**(15): p. 3017-26.

157. Kolev, N.G., E. Ullu, and C. Tschudi, *The emerging role of RNA-binding proteins in the life cycle of Trypanosoma brucei.* Cell Microbiol, 2014. **16**(4): p. 482-9.

158. Wurst, M., et al., *An RNAi screen of the RRM-domain proteins of Trypanosoma brucei.* Mol Biochem Parasitol, 2009. **163**(1): p. 61-5.

159. Subota, I., et al., *ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation.* Mol Biol Cell, 2011. **22**(22): p. 4205-19.

160. Erben, E., C. Chakraborty, and C. Clayton, *The CAF1-NOT complex of trypanosomes.* Front Genet, 2014. **4**: p. 299.

161. Manful, T., A. Fadda, and C. Clayton, *The role of the 5'-3' exoribonuclease XRNA in transcriptome-wide mRNA degradation.* RNA, 2011. **17**(11): p. 2039-47.

162. Clayton, C.E., *Networks of gene expression regulation in Trypanosoma brucei.* Mol Biochem Parasitol, 2014.

163. Warner, J.R., *The economics of ribosome biosynthesis in yeast.* Trends Biochem Sci, 1999. **24**(11): p. 437-40.

164. Kief, D.R. and J.R. Warner, *Coordinate control of syntheses of ribosomal ribonucleic acid and ribosomal proteins during nutritional shift-up in Saccharomyces cerevisiae.* Mol Cell Biol, 1981. **1**(11): p. 1007-15.

165. Schneider, D.A., *RNA polymerase I activity is regulated at multiple steps in the transcription cycle: recent insights into factors that influence transcription elongation.* Gene, 2012. **493**(2): p. 176-84.

166. Hamperl, S., et al., *Chromatin states at ribosomal DNA loci.* Biochim Biophys Acta, 2013.

167. Hasan, G., M.J. Turner, and J.S. Cordingley, *Ribosomal RNA genes of Trypanosoma brucei: mapping the regions specifying the six small ribosomal RNAs.* Gene, 1984. **27**(1): p. 75-86.

168. Viktorovskaya, O.V. and D.A. Schneider, *Functional divergence of eukaryotic RNA polymerases: unique properties of RNA polymerase I suit its cellular role.* Gene, 2015. **556**(1): p. 19-26.

169. Russell, J. and J.C. Zomerdijk, *The RNA polymerase I transcription machinery.* Biochem Soc Symp, 2006(73): p. 203-16.

170. Peyroche, G., et al., *The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3.* EMBO J, 2000. **19**(20): p. 5473-82.

171. Lalo, D., et al., *RRN11 encodes the third subunit of the complex containing Rrn6p and Rrn7p that is essential for the initiation of rDNA transcription by yeast RNA polymerase I.* J Biol Chem, 1996. **271**(35): p. 21062-7.

172. Keener, J., et al., *Reconstitution of yeast RNA polymerase I transcription in vitro from purified components. TATA-binding protein is not required for basal transcription.* J Biol Chem, 1998. **273**(50): p. 33795-802.

173. Keys, D.A., et al., *Multiprotein transcription factor UAF interacts with the upstream element of the yeast RNA polymerase I promoter and forms a stable preinitiation complex.* Genes Dev, 1996. **10**(7): p. 887-903.

174. Hontz, R.D., et al., *Transcription of multiple yeast ribosomal DNA genes requires targeting of UAF to the promoter by Uaf30.* Mol Cell Biol, 2008. **28**(21): p. 6709-19.

175. Steffan, J.S., et al., *The role of TBP in rDNA transcription by RNA polymerase I in Saccharomyces cerevisiae: TBP is required for upstream activation factor-dependent recruitment of core factor.* Genes Dev, 1996. **10**(20): p. 2551-63.

176. Kuhn, C.D., et al., *Functional architecture of RNA polymerase I.* Cell, 2007. **131**(7): p. 1260-72.

177. El Hage, A., et al., *Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis.* Genes Dev, 2010. **24**(14): p. 1546-58.

178. French, S.L., et al., *Distinguishing the roles of Topoisomerases I and II in relief of transcription-induced torsional stress in yeast rRNA genes.* Mol Cell Biol, 2011. **31**(3): p. 482-94.

179. Schneider, D.A., et al., *RNA polymerase II elongation factors Spt4p and Spt5p play roles in transcription elongation by RNA polymerase I and rRNA processing.* Proc Natl Acad Sci U S A, 2006. **103**(34): p. 12707-12.

180. Zhang, Y., et al., *The RNA polymerase-associated factor 1 complex (Paf1C) directly increases the elongation rate of RNA polymerase I and is required for efficient regulation of rRNA synthesis.* J Biol Chem, 2010. **285**(19): p. 14152-9.

181. Lang, W.H. and R.H. Reeder, *Transcription termination of RNA polymerase I due to a T-rich element interacting with Reb1p.* Proc Natl Acad Sci U S A, 1995. **92**(21): p. 9781-5.

182. Merkl, P., et al., *Binding of the termination factor Nsi1 to its cognate DNA site is sufficient to terminate RNA polymerase I transcription in vitro and to induce termination in vivo.* Mol Cell Biol, 2014.

183. Moorefield, B., E.A. Greene, and R.H. Reeder, *RNA polymerase I transcription factor Rrn3 is functionally conserved between yeast and human.* Proc Natl Acad Sci U S A, 2000. **97**(9): p. 4724-9.

184. Russell, J. and J.C. Zomerdijk, *RNA-polymerase-I-directed rDNA transcription, life and works.* Trends Biochem Sci, 2005. **30**(2): p. 87-96.

185. Jantzen, H.M., et al., *Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins.* Nature, 1990. **344**(6269): p. 830-6.

186. Panov, K.I., et al., *RNA polymerase I-specific subunit CAST/hPAF49 has a role in the activation of transcription by upstream binding factor.* Mol Cell Biol, 2006. **26**(14): p. 5436-48.

187. Stefanovsky, V.Y., et al., *ERK modulates DNA bending and enhancesome structure by phosphorylating HMG1-boxes 1 and 2 of the RNA polymerase I transcription factor UBF.* Biochemistry, 2006. **45**(11): p. 3626-34.

188. Bartsch, I., C. Schoneberg, and I. Grummt, *Evolutionary changes of sequences and factors that direct transcription termination of human and mouse ribsomal genes.* Mol Cell Biol, 1987. **7**(7): p. 2521-9.

189. Jansa, P. and I. Grummt, *Mechanism of transcription termination: PTRF interacts with the largest subunit of RNA polymerase I and dissociates paused transcription complexes from yeast and mouse.* Mol Gen Genet, 1999. **262**(3): p. 508-14.

190. Dammann, R., et al., *Chromatin structures and transcription of rDNA in yeast Saccharomyces cerevisiae.* Nucleic Acids Res, 1993. **21**(10): p. 2331-8.

191. Haaf, T., D.L. Hayman, and M. Schmid, *Quantitative determination of rDNA transcription units in vertebrate cells.* Exp Cell Res, 1991. **193**(1): p. 78-86.

192. Conconi, A., et al., *Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle.* Cell, 1989. **57**(5): p. 753-61.

193. Jones, H.S., et al., *RNA polymerase I in yeast transcribes dynamic nucleosomal rDNA.* Nat Struct Mol Biol, 2007. **14**(2): p. 123-30.

194. Xie, W., et al., *The chromatin remodeling complex NuRD establishes the poised state of rRNA genes characterized by bivalent histone modifications and altered nucleosome positions.* Proc Natl Acad Sci U S A, 2012. **109**(21): p. 8161-6.

195. McStay, B. and I. Grummt, *The epigenetics of rRNA genes: from molecular to chromosome biology.* Annu Rev Cell Dev Biol, 2008. **24**: p. 131-57.

196. Sander, E.E. and I. Grummt, *Oligomerization of the transcription termination factor TTF-I: implications for the structural organization of ribosomal transcription units.* Nucleic Acids Res, 1997. **25**(6): p. 1142-7.

197. Yuan, X., et al., *Activation of RNA polymerase I transcription by cockayne syndrome group B protein and histone methyltransferase G9a.* Mol Cell, 2007. **27**(4): p. 585-95.

198. Shen, M., et al., *The chromatin remodeling factor CSB recruits histone acetyltransferase PCAF to rRNA gene promoters in active state for transcription initiation.* PLoS One, 2013. **8**(5): p. e62668.

199. Stancheva, I., et al., *Chromatin structure and methylation of rat rRNA genes studied by formaldehyde fixation and psoralen cross-linking.* Nucleic Acids Res, 1997. **25**(9): p. 1727-35.

200. Santoro, R., J. Li, and I. Grummt, *The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription.* Nat Genet, 2002. **32**(3): p. 393-6.

201. Zhou, Y., R. Santoro, and I. Grummt, *The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription.* EMBO J, 2002. **21**(17): p. 4632-40.

202. Lawrence, R.J., et al., *A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance.* Mol Cell, 2004. **13**(4): p. 599-609.

203. Li, J., G. Langst, and I. Grummt, *NoRC-dependent nucleosome positioning silences rRNA genes.* EMBO J, 2006. **25**(24): p. 5735-41.

204. Zhou, Y. and I. Grummt, *The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing.* Curr Biol, 2005. **15**(15): p. 1434-8.

205. Mayer, C., et al., *Intergenic transcripts regulate the epigenetic state of rRNA genes.* Mol Cell, 2006. **22**(3): p. 351-61.

206. Manelyte, L., et al., *Chromatin targeting signals, nucleosome positioning mechanism and non-coding RNA-mediated regulation of the chromatin remodeling complex NoRC.* PLoS Genet, 2014. **10**(3): p. e1004157.

207. Santoro, R. and I. Grummt, *Molecular mechanisms mediating methylation-dependent silencing of ribosomal gene transcription.* Mol Cell, 2001. **8**(3): p. 719-25.

208. Schmitz, K.M., et al., *Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes.* Genes Dev, 2010. **24**(20): p. 2264-9.

209. Gagnon-Kugler, T., et al., *Loss of human ribosomal gene CpG methylation enhances cryptic RNA polymerase II transcription and disrupts ribosomal RNA processing.* Mol Cell, 2009. **35**(4): p. 414-25.

210. Majumder, S., et al., *Role of DNA methyltransferases in regulation of human ribosomal RNA gene transcription.* J Biol Chem, 2006. **281**(31): p. 22062-72.

211. Srivastava, R., R. Srivastava, and S.H. Ahn, *The Epigenetic Pathways to Ribosomal DNA Silencing.* Microbiol Mol Biol Rev, 2016. **80**(3): p. 545-63.

212. Allen, H.F., P.A. Wade, and T.G. Kutateladze, *The NuRD architecture.* Cell Mol Life Sci, 2013. **70**(19): p. 3513-24.

213. Ling, T., et al., *CHD4/NuRD maintains demethylation state of rDNA promoters through inhibiting the expression of the rDNA methyltransferase recruiter TIP5.* Biochem Biophys Res Commun, 2013. **437**(1): p. 101-7.

214. Cedar, H. and Y. Bergman, *Linking DNA methylation and histone modification: patterns and paradigms.* Nat Rev Genet, 2009. **10**(5): p. 295-304.

215. Murayama, A., et al., *Epigenetic control of rDNA loci in response to intracellular energy status.* Cell, 2008. **133**(4): p. 627-39.

216. French, S.L., et al., *In exponentially growing Saccharomyces cerevisiae cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes.* Mol Cell Biol, 2003. **23**(5): p. 1558-68.

217. Shou, W., et al., *Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex.* Cell, 1999. **97**(2): p. 233-44.

218. Straight, A.F., et al., *Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity.* Cell, 1999. **97**(2): p. 245-56.

219. Imai, S., et al., *Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase.* Nature, 2000. **403**(6771): p. 795-800.

220. Huang, J., et al., *Inhibition of homologous recombination by a cohesin-associated clamp complex recruited to the rDNA recombination enhancer.* Genes Dev, 2006. **20**(20): p. 2887-901.

221. Wittner, M., et al., *Establishment and maintenance of alternative chromatin states at a multicopy gene locus.* Cell, 2011. **145**(4): p. 543-54.

222. Scott, R.S., K.Y. Truong, and J.M. Vos, *Replication initiation and elongation fork rates within a differentially expressed human multicopy locus in early S phase.* Nucleic Acids Res, 1997. **25**(22): p. 4505-12.

223. Dammann, R., et al., *Transcription in the yeast rRNA gene locus: distribution of the active gene copies and chromatin structure of their flanking regulatory sequences.* Mol Cell Biol, 1995. **15**(10): p. 5294-303.

224. Conconi, A., et al., *Repair-independent chromatin assembly onto active ribosomal genes in yeast after UV irradiation.* Mol Cell Biol, 2005. **25**(22): p. 9773-83.

225. Fahy, D., A. Conconi, and M.J. Smerdon, *Rapid changes in transcription and chromatin structure of ribosomal genes in yeast during growth phase transitions.* Exp Cell Res, 2005. **305**(2): p. 365-73.

226. Walgraffe, D., et al., *Characterization of subunits of the RNA polymerase I complex in Trypanosoma brucei.* Mol Biochem Parasitol, 2005. **139**(2): p. 249-60.

227. Nguyen, T.N., B. Schimanski, and A. Gunzl, *Active RNA polymerase I of Trypanosoma brucei harbors a novel subunit essential for transcription.* Mol Cell Biol, 2007. **27**(17): p. 6254-63.

228. Nguyen, T.N., et al., *Purification of an eight subunit RNA polymerase I complex in Trypanosoma brucei.* Mol Biochem Parasitol, 2006. **149**(1): p. 27-37.

229. Penate, X., et al., *RNA pol II subunit RPB7 is required for RNA pol I-mediated transcription in Trypanosoma brucei.* EMBO Rep, 2009. **10**(3): p. 252-7.

230. Park, S.H., et al., *Transcription by the multifunctional RNA polymerase I in Trypanosoma brucei functions independently of RPB7.* Mol Biochem Parasitol, 2011. **180**(1): p. 35-42.

231. Figueiredo, L.M. and D. Horn, *Trypanosoma Brucei Subtelomeres: Monoallelic Expression and Antigenic Variation.* Subtelomeres, 2014: p. 137-152.

232. Rudenko, G., et al., *A ribosomal DNA promoter replacing the promoter of a telomeric VSG gene expression site can be efficiently switched on and off in T. brucei.* Cell, 1995. **83**(4): p. 547-53.

233. Brandenburg, J., et al., *Multifunctional class I transcription in Trypanosoma brucei depends on a novel protein complex.* EMBO J, 2007. **26**(23): p. 4856-66.

234. Nguyen, T.N., et al., *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, 2012. **11**(12): p. 1573-81.

235. Kirkham, J.K., et al., *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.* Mol Cell Biol, 2015.

236. Nguyen, T.N., et al., *Promoter occupancy of the basal class I transcription factor A differs strongly between active and silent VSG expression sites in Trypanosoma brucei.* Nucleic Acids Res, 2014. **42**(5): p. 3164-76.

237. Alsford, S. and D. Horn, *Elongator protein 3b negatively regulates ribosomal DNA transcription in african trypanosomes.* Mol Cell Biol, 2011. **31**(9): p. 1822-32.

238. Pena, A.C., et al., *Trypanosoma brucei histone H1 inhibits RNA polymerase I transcription and is important for parasite fitness in vivo.* Mol Microbiol, 2014.

239. Povelones, M.L., et al., *Histone H1 Plays a Role in Heterochromatin Formation and VSG Expression Site Silencing in Trypanosoma brucei.* PLoS Pathog, 2012. **8**(11): p. e1003010.

240. Alsford, S. and D. Horn, *Cell-cycle-regulated control of VSG expression site silencing by histones and histone chaperones ASF1A and CAF-1b in Trypanosoma brucei.* Nucleic Acids Res, 2012. **40**(20): p. 10150-60.

241. Siegel, T.N., et al., *Acetylation of histone H4K4 is cell cycle regulated and mediated by HAT3 in Trypanosoma brucei.* Mol Microbiol, 2008. **67**(4): p. 762-71.

242. Kawahara, T., et al., *Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes.* Mol Microbiol, 2008. **69**(4): p. 1054-68.

243. Glover, L. and D. Horn, *Locus-specific control of DNA resection and suppression of subtelomeric VSG recombination by HAT3 in the African trypanosome.* Nucleic Acids Res, 2014. **42**(20): p. 12600-13.

244. Janzen, C.J., et al., *Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in Trypanosoma brucei.* Mol Cell, 2006. **23**(4): p. 497-507.

245. Figueiredo, L.M., C.J. Janzen, and G.A. Cross, *A histone methyltransferase modulates antigenic variation in African trypanosomes.* PLoS Biol, 2008. **6**(7): p. e161.

246. Gassen, A., et al., *DOT1A-dependent H3K76 methylation is required for replication regulation in Trypanosoma brucei.* Nucleic Acids Res, 2012. **40**(20): p. 10302-11.

247. Lopez-Farfan, D., et al., *SUMOylation by the E3 ligase TbSIZ1/PIAS1 positively regulates VSG expression in Trypanosoma brucei.* PLoS Pathog, 2014. **10**(12): p. e1004545.

248. Hughes, K., et al., *A novel ISWI is involved in VSG expression site downregulation in African trypanosomes.* EMBO J, 2007. **26**(9): p. 2400-10.

249. Stanne, T., et al., *Identification of the ISWI Chromatin Remodeling Complex of the Early Branching Eukaryote Trypanosoma brucei.* J Biol Chem, 2015. **290**(45): p. 26954-67.

250. Stanne, T.M., et al., *TbISWI regulates multiple polymerase I (Pol I)-transcribed loci and is present at Pol II transcription boundaries in Trypanosoma brucei.* Eukaryot Cell, 2011. **10**(7): p. 964-76.

251. Narayanan, M.S., et al., *NLP is a novel transcription regulator involved in VSG expression site control in Trypanosoma brucei.* Nucleic Acids Res, 2011. **39**(6): p. 2018-31.

252. Denninger, V., et al., *The FACT subunit TbSpt16 is involved in cell cycle specific control of VSG expression sites in Trypanosoma brucei.* Mol Microbiol, 2010. **78**(2): p. 459-74.

253. Denninger, V. and G. Rudenko, *FACT plays a major role in histone dynamics affecting VSG expression site control in Trypanosoma brucei.* Mol Microbiol, 2014. **94**(4): p. 945-62.

254. Conway, C., et al., *Ku is important for telomere maintenance, but not for differential expression of telomeric VSG genes, in African trypanosomes.* J Biol Chem, 2002. **277**(24): p. 21269-77.

255. Janzen, C.J., et al., *Telomere length regulation and transcriptional silencing in KU80-deficient Trypanosoma brucei.* Nucleic Acids Res, 2004. **32**(22): p. 6575-84.

256. Li, B., A. Espinal, and G.A. Cross, *Trypanosome telomeres are protected by a homologue of mammalian TRF2.* Mol Cell Biol, 2005. **25**(12): p. 5011-21.

257. Dreesen, O. and G.A. Cross, *Consequences of telomere shortening at an active VSG expression site in telomerase-deficient Trypanosoma brucei.* Eukaryot Cell, 2006. **5**(12): p. 2114-9.

258. Yang, X., et al., *RAP1 is essential for silencing telomeric variant surface glycoprotein genes in Trypanosoma brucei.* Cell, 2009. **137**(1): p. 99-109.

259. Hovel-Miner, G.A., et al., *Telomere length affects the frequency and mechanism of antigenic variation in Trypanosoma brucei.* PLoS Pathog, 2012. **8**(8): p. e1002900.

260. Pandya, U.M., R. Sandhu, and B. Li, *Silencing subtelomeric VSGs by Trypanosoma brucei RAP1 at the insect stage involves chromatin structure changes.* Nucleic Acids Res, 2013. **41**(16): p. 7673-82.

261. Jehi, S.E., et al., *Suppression of subtelomeric VSG switching by Trypanosoma brucei TRF requires its TTAGGG repeat-binding activity.* Nucleic Acids Res, 2014. **42**(20): p. 12899-911.

262. Jehi, S.E., F. Wu, and B. Li, *Trypanosoma brucei TIF2 suppresses VSG switching by maintaining subtelomere integrity.* Cell Res, 2014.

263. DuBois, K.N., et al., *NUP-1 Is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions.* PLoS Biol, 2012. **10**(3): p. e1001287.

264. Cestari, I. and K. Stuart, *Inositol phosphate pathway controls transcription of telomeric expression sites in trypanosomes.* Proc Natl Acad Sci U S A, 2015. **112**(21): p. E2803-12.

265. Landeira, D., et al., *Cohesin regulates VSG monoallelic expression in trypanosomes.* J Cell Biol, 2009. **186**(2): p. 243-54.

266. Kim, H.S., et al., *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, 2013. **8**(2): p. e57001.

267. Benmerzouga, I., et al., *Trypanosoma brucei Orc1 is essential for nuclear DNA replication and affects both VSG silencing and VSG switching.* Mol Microbiol, 2013. **87**(1): p. 196-210.

268. Tiengwe, C., et al., *Genome-wide analysis reveals extensive functional interaction between DNA replication initiation and transcription in the genome of Trypanosoma brucei.* Cell Rep, 2012. **2**(1): p. 185-97.

269. Schulz, D., et al., *Bromodomain Proteins Contribute to Maintenance of Bloodstream Form Stage Identity in the African Trypanosome.* PLoS Biol, 2015. **13**(12): p. e1002316.

270. Glover, L., et al., *VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes.* Proc Natl Acad Sci U S A, 2016. **113**(26): p. 7225-30.

271. Roditi, I. and C. Clayton, *An unambiguous nomenclature for the major surface glycoproteins of the procyclic form of Trypanosoma brucei.* Mol Biochem Parasitol, 1999. **103**(1): p. 99-100.

272. Haenni, S., et al., *Bidirectional silencing of RNA polymerase I transcription by a strand switch region in Trypanosoma brucei.* Nucleic Acids Res, 2009. **37**(15): p. 5007-18.

273. Acosta-Serrano, A., et al., *The procyclin repertoire of Trypanosoma brucei. Identification and structural characterization of the Glu-Pro-rich polypeptides.* J Biol Chem, 1999. **274**(42): p. 29763-71.

274. Biebinger, S., et al., *The PARP promoter of Trypanosoma brucei is developmentally regulated in a chromosomal context.* Nucleic Acids Res, 1996. **24**(7): p. 1202-11.

275. Vanhamme, L., et al., *Stimuli of differentiation regulate RNA elongation in the transcription units for the major stage-specific antigens of Trypanosoma brucei.* Nucleic Acids Res, 1995. **23**(11): p. 1862-9.

276. Hotz, H.R., et al., *PARP gene expression: control at many levels.* Mol Biochem Parasitol, 1998. **91**(1): p. 131-43.

277. Vassella, E., et al., *Multiple procyclin isoforms are expressed differentially during the development of insect forms of Trypanosoma brucei.* J Mol Biol, 2001. **312**(4): p. 597-607.

278. Vassella, E., et al., *Expression of a major surface protein of Trypanosoma brucei insect forms is controlled by the activity of mitochondrial enzymes.* Mol Biol Cell, 2004. **15**(9): p. 3986-93.

279. Glover, L., et al., *Antigenic variation in African trypanosomes: the importance of chromosomal and nuclear context in VSG expression control.* Cell Microbiol, 2013. **15**(12): p. 1984-93.

280. Alsford, S. and D. Horn, *Cell-cycle-regulated control of VSG expression site silencing by histones and histone chaperones ASF1A and CAF-1b in Trypanosoma brucei.* Nucleic Acids Res, 2012.

281. Alsford, S., et al., *A sirtuin in the African trypanosome is involved in both DNA repair and telomeric gene silencing but is not required for antigenic variation.* Mol Microbiol, 2007. **63**(3): p. 724-36.

282. Vanhamme, L., et al., *Differential RNA elongation controls the variant surface glycoprotein gene expression sites of Trypanosoma brucei.* Mol Microbiol, 2000. **36**(2): p. 328-40.

283. Kassem, A., E. Pays, and L. Vanhamme, *Transcription is initiated on silent variant surface glycoprotein expression sites despite monoallelic expression in Trypanosoma brucei.* Proc Natl Acad Sci U S A, 2014. **111**(24): p. 8943-8.

284. Rudenko, G., et al., *VSG gene expression site control in insect form Trypanosoma brucei.* EMBO J, 1994. **13**(22): p. 5470-82.

285. Navarro, M. and K. Gull, *A pol I transcriptional body associated with VSG mono-allelic expression in Trypanosoma brucei.* Nature, 2001. **414**(6865): p. 759-63.

286. Fenn, K. and K.R. Matthews, *The cell biology of Trypanosoma brucei differentiation.* Curr Opin Microbiol, 2007. **10**(6): p. 539-46.

287. Ziegelbauer, K., et al., *Synchronous differentiation of Trypanosoma brucei from bloodstream to procyclic forms in vitro.* Eur J Biochem, 1990. **192**(2): p. 373-8.

288. Amiguet-Vercher, A., et al., *Loss of the mono-allelic control of the VSG expression sites during the development of Trypanosoma brucei in the bloodstream.* Mol Microbiol, 2004. **51**(6): p. 1577-88.

289. Navarro, M., G.A. Cross, and E. Wirtz, *Trypanosoma brucei variant surface glycoprotein regulation involves coupled activation/inactivation and chromatin remodeling of expression sites.* EMBO J, 1999. **18**(8): p. 2265-72.

290. Graham, S.V. and J.D. Barry, *Transcriptional regulation of metacyclic variant surface glycoprotein gene expression during the life cycle of Trypanosoma brucei.* Mol Cell Biol, 1995. **15**(11): p. 5945-56.

291. Lenardo, M.J., et al., *Metacyclic variant surface glycoprotein genes of Trypanosoma brucei subsp. rhodesiense are activated in situ, and their expression is transcriptionally regulated.* Mol Cell Biol, 1986. **6**(6): p. 1991-7.

292. Graham, S.V., B. Wymer, and J.D. Barry, *Activity of a trypanosome metacyclic variant surface glycoprotein gene promoter is dependent upon life cycle stage and chromosomal context.* Mol Cell Biol, 1998. **18**(3): p. 1137-46.

293. Goodwin, G.H., C. Sanders, and E.W. Johns, *A new group of chromatin-associated proteins with a high content of acidic and basic amino acids.* Eur J Biochem, 1973. **38**(1): p. 14-9.

294. Stros, M., D. Launholt, and K.D. Grasser, *The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins.* Cell Mol Life Sci, 2007. **64**(19-20): p. 2590-606.

295. Stros, M., *HMGB proteins: interactions with DNA and chromatin.* Biochim Biophys Acta, 2010. **1799**(1-2): p. 101-13.

296. Malarkey, C.S. and M.E. Churchill, *The high mobility group box: the ultimate utility player of a cell.* Trends Biochem Sci, 2012. **37**(12): p. 553-62.

297. Yang, H., et al., *High Mobility Group Box Protein 1 (HMGB1): The Prototypical Endogenous Danger Molecule.* Mol Med, 2015. **21 Suppl 1**: p. S6-S12.

298. Barreiro-Alonso, A., et al., *High Mobility Group B Proteins, Their Partners, and Other Redox Sensors in Ovarian and Prostate Cancer.* Oxid Med Cell Longev, 2016. **2016**: p. 5845061.

299. Ranzato, E., et al., *High mobility group box protein-1 in wound repair.* Cells, 2012. **1**(4): p. 699-710.

300. Reeves, R., *Nuclear functions of the HMG proteins.* Biochim Biophys Acta, 2010. **1799**(1-2): p. 3-14.

301. Ueda, T. and M. Yoshida, *HMGB proteins and transcriptional regulation.* Biochim Biophys Acta, 2010. **1799**(1-2): p. 114-8.

302. Gonzalez-Huici, V., et al., *DNA bending facilitates the error-free DNA damage tolerance pathway and upholds genome integrity.* EMBO J, 2014. **33**(4): p. 327-40.

303. Dai, Y., et al., *Determinants of HMGB proteins required to promote RAG1/2-recombination signal sequence complex assembly and catalysis during V(D)J recombination.* Mol Cell Biol, 2005. **25**(11): p. 4413-25.

304. Bonaldi, T., et al., *The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding.* EMBO J, 2002. **21**(24): p. 6865-73.

305. Hepp, M.I., et al., *Nucleosome remodeling by the SWI/SNF complex is enhanced by yeast High Mobility Group Box (HMGB) proteins.* Biochim Biophys Acta, 2014. **1839**(9): p. 764-72.

306. Agresti, A. and M.E. Bianchi, *HMGB proteins and gene expression.* Curr Opin Genet Dev, 2003. **13**(2): p. 170-8.

307. Ge, H. and R.G. Roeder, *The high mobility group protein HMG1 can reversibly inhibit class II gene transcription by interaction with the TATA-binding protein.* J Biol Chem, 1994. **269**(25): p. 17136-40.

308. Gadal, O., et al., *Hmo1, an HMG-box protein, belongs to the yeast ribosomal DNA transcription system.* EMBO J, 2002. **21**(20): p. 5498-507.

309. Kamau, E., K.T. Bauerle, and A. Grove, *The Saccharomyces cerevisiae high mobility group box protein HMO1 contains two functional DNA binding domains.* J Biol Chem, 2004. **279**(53): p. 55234-40.

310. Higashino, A., et al., *Both HMG boxes in Hmo1 are essential for DNA binding in vitro and in vivo.* Biosci Biotechnol Biochem, 2014: p. 1-10.

311. Bauerle, K.T., E. Kamau, and A. Grove, *Interactions between N- and C-terminal domains of the Saccharomyces cerevisiae high-mobility group protein HMO1 are required for DNA bending.* Biochemistry, 2006. **45**(11): p. 3635-45.

312. Hall, D.B., J.T. Wade, and K. Struhl, *An HMG protein, Hmo1, associates with promoters of many ribosomal protein genes and throughout the rRNA gene locus in Saccharomyces cerevisiae.* Mol Cell Biol, 2006. **26**(9): p. 3672-9.

313. Merz, K., et al., *Actively transcribed rRNA genes in S. cerevisiae are organized in a specialized chromatin associated with the high-mobility group protein Hmo1 and are largely devoid of histone molecules.* Genes Dev, 2008. **22**(9): p. 1190-204.

314. Kasahara, K., et al., *Saccharomyces cerevisiae HMO1 interacts with TFIID and participates in start site selection by RNA polymerase II.* Nucleic Acids Res, 2008. **36**(4): p. 1343-57.

315. Kasahara, K., Y. Ohyama, and T. Kokubo, *Hmo1 directs pre-initiation complex assembly to an appropriate site on its target gene promoters by masking a nucleosome-free region.* Nucleic Acids Res, 2011. **39**(10): p. 4136-50.

316. Berger, A.B., et al., *Hmo1 is required for TOR-dependent regulation of ribosomal protein gene transcription.* Mol Cell Biol, 2007. **27**(22): p. 8015-26.

317. Panday, A., L. Xiao, and A. Grove, *Yeast high mobility group protein HMO1 stabilizes chromatin and is evicted during repair of DNA double strand breaks.* Nucleic Acids Res, 2015. **43**(12): p. 5759-70.

318. Yoshikawa, K., et al., *Comprehensive phenotypic analysis of single-gene deletion and overexpression strains of Saccharomyces cerevisiae.* Yeast, 2011. **28**(5): p. 349-61.

319. O'Mahony, D.J. and L.I. Rothblum, *Identification of two forms of the RNA polymerase I transcription factor UBF.* Proc Natl Acad Sci U S A, 1991. **88**(8): p. 3180-4.

320. Kuhn, A., et al., *Functional differences between the two splice variants of the nucleolar transcription factor UBF: the second HMG box determines specificity of DNA binding and transcriptional activity.* EMBO J, 1994. **13**(2): p. 416-24.

321. Jantzen, H.M., et al., *Multiple domains of the RNA polymerase I activator hUBF interact with the TATA-binding protein complex hSL1 to mediate transcription.* Genes Dev, 1992. **6**(10): p. 1950-63.

322. O'Sullivan, A.C., G.J. Sullivan, and B. McStay, *UBF binding in vivo is not restricted to regulatory sequences within the vertebrate ribosomal DNA repeat.* Mol Cell Biol, 2002. **22**(2): p. 657-68.

323. Panov, K.I., et al., *UBF activates RNA polymerase I transcription by stimulating promoter escape.* EMBO J, 2006. **25**(14): p. 3310-22.

324. Stefanovsky, V., et al., *Growth factor signaling regulates elongation of RNA polymerase I transcription in mammals via UBF phosphorylation and r-chromatin remodeling.* Mol Cell, 2006. **21**(5): p. 629-39.

325. Sanij, E., et al., *UBF levels determine the number of active ribosomal RNA genes in mammals.* J Cell Biol, 2008. **183**(7): p. 1259-74.

326. Chen, D., A.S. Belmont, and S. Huang, *Upstream binding factor association induces large-scale chromatin decondensation.* Proc Natl Acad Sci U S A, 2004. **101**(42): p. 15106-11.

327. Hannan, R.D., et al., *Overexpression of the transcription factor UBF1 is sufficient to increase ribosomal DNA transcription in neonatal cardiomyocytes: implications for cardiac hypertrophy.* Proc Natl Acad Sci U S A, 1996. **93**(16): p. 8750-5.

328. Chen, J. and Y.M. Deng, *Influenza virus antigenic variation, host antibody production and new approach to control epidemics.* Virol J, 2009. **6**: p. 30.

329. Stoenner, H.G., T. Dodd, and C. Larsen, *Antigenic variation of Borrelia hermsii.* J Exp Med, 1982. **156**(5): p. 1297-311.

330. Meier, J.T., M.I. Simon, and A.G. Barbour, *Antigenic variation is associated with DNA rearrangements in a relapsing fever Borrelia.* Cell, 1985. **41**(2): p. 403-9.

331. Plasterk, R.H., M.I. Simon, and A.G. Barbour, *Transposition of structural genes to an expression sequence on a linear plasmid causes antigenic variation in the bacterium Borrelia hermsii.* Nature, 1985. **318**(6043): p. 257-63.

332. Guizetti, J. and A. Scherf, *Silence, activate, poise and switch! Mechanisms of antigenic variation in Plasmodium falciparum.* Cell Microbiol, 2013. **15**(5): p. 718-26.

333. Pays, E., et al., *The molecular arms race between African trypanosomes and humans.* Nat Rev Microbiol, 2014. **12**(8): p. 575-84.

334. Jackson, A.P., et al., *Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species.* Proc Natl Acad Sci U S A, 2012. **109**(9): p. 3416-21.

335. Roth, C., et al., *Antigenic variation in Trypanosoma equiperdum.* Res Microbiol, 1991. **142**(6): p. 725-30.

336. Atwood, J.A., 3rd, et al., *The Trypanosoma cruzi proteome.* Science, 2005. **309**(5733): p. 473-6.

337. Mugnier, M.R., G.A. Cross, and F.N. Papavasiliou, *The in vivo dynamics of antigenic variation in Trypanosoma brucei.* Science, 2015. **347**(6229): p. 1470-3.

338. Chess, A., et al., *Allelic inactivation regulates olfactory receptor gene expression.* Cell, 1994. **78**(5): p. 823-34.

339. Magklara, A., et al., *An epigenetic signature for monoallelic olfactory receptor expression.* Cell, 2011. **145**(4): p. 555-70.

340. Lyons, D.B., et al., *An epigenetic trap stabilizes singular olfactory receptor expression.* Cell, 2013. **154**(2): p. 325-36.

341. Shykind, B.M., et al., *Gene switching and the stability of odorant receptor gene choice.* Cell, 2004. **117**(6): p. 801-15.

342. Dalton, R.P., D.B. Lyons, and S. Lomvardas, *Co-opting the unfolded protein response to elicit olfactory receptor feedback.* Cell, 2013. **155**(2): p. 321-32.

343. Dzikowski, R., M. Frank, and K. Deitsch, *Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production.* PLoS Pathog, 2006. **2**(3): p. e22.

344. Lopez-Rubio, J.J., L. Mancio-Silva, and A. Scherf, *Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites.* Cell Host Microbe, 2009. **5**(2): p. 179-90.

345. Lopez-Rubio, J.J., et al., *5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites.* Mol Microbiol, 2007. **66**(6): p. 1296-305.

346. Volz, J.C., et al., *PfSET10, a Plasmodium falciparum methyltransferase, maintains the active var gene in a poised state during parasite division.* Cell Host Microbe, 2012. **11**(1): p. 7-18.

347. Munoz-Jordan, J.L., K.P. Davies, and G.A. Cross, *Stable expression of mosaic coats of variant surface glycoproteins in Trypanosoma brucei.* Science, 1996. **272**(5269): p. 1795-7.

348. Chaves, I., et al., *Control of variant surface glycoprotein gene-expression sites in Trypanosoma brucei.* EMBO J, 1999. **18**(17): p. 4846-55.

349. Sheader, K., et al., *Variant surface glycoprotein RNA interference triggers a precytokinesis cell cycle arrest in African trypanosomes.* Proc Natl Acad Sci U S A, 2005. **102**(24): p. 8716-21.

350. MacGregor, P., et al., *Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act.* Nat Rev Microbiol, 2012. **10**(6): p. 431-8.

351. Vink, C., G. Rudenko, and H.S. Seifert, *Microbial antigenic variation mediated by homologous DNA recombination.* FEMS Microbiol Rev, 2012. **36**(5): p. 917-48.

352. McCulloch, R. and J.D. Barry, *A role for RAD51 and homologous recombination in Trypanosoma brucei antigenic variation.* Genes Dev, 1999. **13**(21): p. 2875-88.

353. Hartley, C.L. and R. McCulloch, *Trypanosoma brucei BRCA2 acts in antigenic variation and has undergone a recent expansion in BRC repeat number that is important during homologous recombination.* Mol Microbiol, 2008. **68**(5): p. 1237-51.

354. Robinson, N.P., et al., *Predominance of duplicative VSG gene conversion in antigenic variation in African trypanosomes.* Mol Cell Biol, 1999. **19**(9): p. 5839-46.

355. Thon, G., et al., *Trypanosome variable surface glycoproteins: composite genes and order of expression.* Genes Dev, 1990. **4**(8): p. 1374-83.

356. Marcello, L. and J.D. Barry, *Analysis of the VSG gene silent archive in Trypanosoma brucei reveals that mosaic gene expression is prominent in antigenic variation and is favored by archive substructure.* Genome Res, 2007. **17**(9): p. 1344-52.

357. Hall, J.P., H. Wang, and J.D. Barry, *Mosaic VSGs and the scale of Trypanosoma brucei antigenic variation.* PLoS Pathog, 2013. **9**(7): p. e1003502.

358. Glover, L., S. Alsford, and D. Horn, *DNA break site at fragile subtelomeres determines probability and mechanism of antigenic variation in African trypanosomes.* PLoS Pathog, 2013. **9**(3): p. e1003260.

359. Boothroyd, C.E., et al., *A yeast-endonuclease-generated DNA break induces antigenic switching in Trypanosoma brucei.* Nature, 2009. **459**(7244): p. 278-81.

360. Horn, D., *Antigenic variation in African trypanosomes.* Mol Biochem Parasitol, 2014. **195**(2): p. 123-9.

361. Devlin, R., et al., *Mapping replication dynamics in Trypanosoma brucei reveals a link with telomere transcription and antigenic variation.* Elife, 2016. **5**.

362. Morrison, L.J., et al., *Probabilistic order in antigenic variation of Trypanosoma brucei.* Int J Parasitol, 2005. **35**(9): p. 961-72.

363. Lythgoe, K.A., et al., *Parasite-intrinsic factors can explain ordered progression of trypanosome antigenic variation.* Proc Natl Acad Sci U S A, 2007. **104**(19): p. 8095-100.

364. Davies, K.P., V.B. Carruthers, and G.A. Cross, *Manipulation of the vsg co-transposed region increases expression-site switching in Trypanosoma brucei.* Mol Biochem Parasitol, 1997. **86**(2): p. 163-77.

365. Batram, C., et al., *Expression site attenuation mechanistically links antigenic variation and development in Trypanosoma brucei.* Elife, 2014. **3**: p. e02324.

366. Figueiredo, L.M., G.A. Cross, and C.J. Janzen, *Epigenetic regulation in African trypanosomes: a new kid on the block.* Nat Rev Microbiol, 2009. **7**(7): p. 504-13.

367. Horn, D. and R. McCulloch, *Molecular mechanisms underlying the control of antigenic variation in African trypanosomes.* Curr Opin Microbiol, 2010. **13**(6): p. 700-5.

368. Batram, C., et al., *Expression site attenuation mechanistically links antigenic variation and development in Trypanosoma brucei.* Elife (Cambridge), 2014. **3**: p. e02324.

369. Gunzl, A., et al., *Mono-allelic VSG expression by RNA polymerase I in Trypanosoma brucei: Expression site control from both ends?* Gene, 2015. **556**(1): p. 68-73.

370. Johnson, J.G. and G.A. Cross, *Selective cleavage of variant surface glycoproteins from Trypanosoma brucei.* Biochem J, 1979. **178**(3): p. 689-97.

371. Hirumi, H. and K. Hirumi, *Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers.* J Parasitol, 1989. **75**(6): p. 985-9.

372. Schumann Burkard, G., P. Jutzi, and I. Roditi, *Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters.* Mol Biochem Parasitol, 2011. **175**(1): p. 91-4.

373. Ehlers, B., J. Czichos, and P. Overath, *RNA turnover in Trypanosoma brucei.* Mol Cell Biol, 1987. **7**(3): p. 1242-9.

374. Giresi, P.G., et al., *FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin.* Genome Res, 2007. **17**(6): p. 877-85.

375. Worthen, C., B.C. Jensen, and M. Parsons, *Diverse effects on mitochondrial and nuclear functions elicited by drugs and genetic knockdowns in bloodstream stage Trypanosoma brucei.* PLoS Negl Trop Dis, 2010. **4**(5): p. e678.

376. Patrick, K.L., et al., *Distinct and overlapping roles for two Dicer-like proteins in the RNA interference pathways of the ancient eukaryote Trypanosoma brucei.* Proc Natl Acad Sci U S A, 2009. **106**(42): p. 17933-8.

377. Groth, A., et al., *Chromatin challenges during DNA replication and repair.* Cell, 2007. **128**(4): p. 721-33.

378. Li, B., *A newly discovered role of telomeres in an ancient organism.* Nucleus, 2010. **1**(3): p. 260-3.

379. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. **125**(2): p. 315-26.

380. Puri, D., et al., *High-wire act: the poised genome and cellular memory.* FEBS J, 2014.

381. Lim, P.S., et al., *Epigenetic regulation of inducible gene expression in the immune system.* Immunology, 2013. **139**(3): p. 285-93.

382. Krangel, M.S., *Gene segment selection in V(D)J recombination: accessibility and beyond.* Nat Immunol, 2003. **4**(7): p. 624-30.

383. Clos, J., D. Buttgereit, and I. Grummt, *A purified transcription factor (TIF-IB) binds to essential sequences of the mouse rDNA promoter.* Proc Natl Acad Sci U S A, 1986. **83**(3): p. 604-8.

384. Johnson, P.J., J.M. Kooter, and P. Borst, *Inactivation of transcription by UV irradiation of T. brucei provides evidence for a multicistronic transcription unit including a VSG gene.* Cell, 1987. **51**(2): p. 273-81.

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

386. Bastin, P., et al., *A novel epitope tag system to study protein targeting and organelle biogenesis in Trypanosoma brucei.* Mol Biochem Parasitol, 1996. **77**(2): p. 235-9.

387. Birkett, C.R., et al., *Use of monoclonal antibodies to analyse the expression of a multi-tubulin family.* FEBS Lett, 1985. **187**(2): p. 211-8.

388. Aresta-Branco, F., S. Pimenta, and L.M. Figueiredo, *A transcription-independent epigenetic mechanism is associated with antigenic switching in Trypanosoma brucei.* Nucleic Acids Res, 2015.

389. Benmerzouga, I., et al., *Trypanosoma brucei Orc1 is essential for nuclear DNA replication and affects both VSG silencing and VSG switching.* Mol Microbiol, 2012.

390. Kasahara, K., et al., *Assembly of regulatory factors on rRNA and ribosomal protein genes in Saccharomyces cerevisiae.* Mol Cell Biol, 2007. **27**(19): p. 6686-705.

391. Lamont, G.S., R.S. Tucker, and G.A. Cross, *Analysis of antigen switching rates in Trypanosoma brucei.* Parasitology, 1986. **92 (Pt 2)**: p. 355-67.

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

393. Alsford, S., et al., *High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome.* Genome Res, 2011.

CURRICULUM VITAE

NAME	POSITION TITLE
Aresta Branco, Francisco	PhD student

ACADEMIC QUALIFICATIONS:

START MONTH /YEAR	END MONTH/ YEAR	DEGREE *(if applicable)*	INSTITUTION AND LOCATION	TRAINING MENTOR	SCIENTIFIC DISCIPLINE
03/2012	Present	PhD	Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon	Luísa Figueiredo	Parasitology
09/2007	11/2009	MSc	Faculty of Science, University of Lisbon	Rodrigo de Almeida & Susana Marinho	Medical Biochemistry (17 out of 20)
09/2003	07/2007	BSc	Faculty of Science, University of Lisbon	-	Biochemistry (16 out of 20)

SCIENTIFIC ACTIVITY:

START MONTH/ YEAR	END MONTH /YEAR	POSITION TITLE	DEPARTMENT	INSTITUTION AND LOCATION
03/2012	present	PhD Fellowship	Biology of Parasitism Lab (Luísa Figueiredo)	Instituto de Medicina Molecular
09/2011	02/2012	FCT Research Fellowship	Biology of Parasitism Lab (Luísa Figueiredo)	Instituto de Medicina Molecular
11/2010	08/2011	IEFP Fellowship	Biology of Parasitism Lab (Luísa Figueiredo)	Instituto de Medicina Molecular
03/2010	09/2010	FCT Research Fellowship	Biochemistry of Oxidants and Antioxidants Lab (Susana Marinho)	Centro de Química e Bioquímica, Faculty of Science
09/2008	11/2009	MSc Thesis	Molecular Biophysics (Rodrigo de Almeida) & Biochemistry of Oxidants and Antioxidants (Susana Marinho) Labs	Centro de Química e Bioquímica, Faculty of Science

PEER-REVIEWED PUBLICATIONS

Research articles

1. Trindade S & Rijo-Ferreira F, Carvalho T, Pinto-Neves D, Guegan F, **Aresta-Branco F**, Bento F, Young SA, Pinto A, Van Den Abbeele J, Ribeiro RM, Dias S, Smith TK, Figueiredo LM "*Trypanosoma brucei* parasites occupy and functionally adapt to the adipose tissue in mice" 2016 Cell Host & Microbe 19: 837-848
   - *Research Highlight in Nat Rev Microbiol (2016).*
   - *Preview in Cell Host & Microbe (2016).*
   - *Recommended by Faculty 1000 (2016).*
   - *Highlight in HHMI news (2016).*
2. **Aresta-Branco F**, Pimenta S, Figueiredo LM. "A transcription-independent epigenetic mechanism is associated with antigenic switching in *Trypanosoma brucei*" 2016 Nucleic Acids Res. 44:3131-3146.
   - *Recommended by Faculty 1000 (2016).*
3. Pena AC, Pimentel MR, Manso H, Vaz-Drago R, Pinto-Neves D, **Aresta-Branco F**, Rijo-Ferreira F, Guegan F, Pedro Coelho L, Carmo-Fonseca M, Barbosa-Morais NL, Figueiredo LM. "*Trypanosoma brucei* histone H1 inhibits RNA polymerase I transcription and is important for parasite fitness *in vivo*" 2014 Mol. Microbiol. 93: 645–663.
4. **Aresta-Branco F**, Cordeiro A, Marinho HS, Cyrne L, Antunes F, de Almeida RF. "Gel Domains in the Plasma Membrane of *Saccharomyces cerevisiae*: Highly Ordered, Ergosterol-free and Sphingolipid-enriched Lipid Rafts" 2011 J. Biol. Chem, 286: 5043-5054.
   - *Recommended by Faculty 1000 (2011).*

Book chapter

1. Pena AC, **Aresta-Branco F**, Figueiredo LM. "Epigenetic regulation in T.brucei: changing coats is a chance to survive" In. Epigenetics and Human Health; Ed Springer; **submitted**

SCIENTIFIC COMMUNICATIONS

Oral Presentations

1. Kinetoplastid Molecular Cell Biology Meeting, Woods Hole, Massachussets, USA, 2015. "Keeping an open chromatin in the active VSG expression site is associated to uncommitted switching in *Trypanosoma brucei*", **Francisco Aresta-Branco**, Sílvia Pimenta, Luísa M. Figueiredo – **Teaser Talk**
2. IMM PhD Student Annual Meeting, Lisboa, Portugal, 2015. "Keeping an open chromatin in the active VSG expression site is associated to uncommitted switching in *Trypanosoma brucei*", **Francisco Aresta-Branco**, Sílvia Pimenta, Luísa M. Figueiredo. – **Best Oral Presentation Award**
3. ISWOLD 2010, Rehovot, Israel, 2010. "The plasma membrane of yeast cells has highly ordered ergosterol-free sphingolipid-enriched lipid domains", **Francisco Aresta-Branco**, André M. Cordeiro, H. Susana Marinho, Luísa Cyrne, Fernando Antunes, Rodrigo F. M. de Almeida
4. FEBS Advanced Course on "Lipid signalling and disease", Ortona, Itália, 2009. "The plasma membrane of yeast cells contains sterol-independent highly-ordered, sphingolipid-enriched domains", **Francisco Aresta-Branco**, André M. Cordeiro, H. Susana Marinho, Luísa Cyrne, Fernando Antunes, Rodrigo F. M. de Almeida.

Poster Presentations

1. Kinetoplastid Molecular Cell Biology Meeting, Woods Hole, Massachussets, USA, 2015. "Keeping an open chromatin in the active VSG expression site is associated to uncommitted switching in *Trypanosoma brucei*", **Francisco Aresta-Branco**, Sílvia Pimenta, Luísa M. Figueiredo.
2. Kinetoplastid Molecular Cell Biology Meeting, Woods Hole, Massachussets, USA, 2013. "Open chromatin of VSG Expression Site is maintained independently of transcription in *Trypanosoma brucei*", **Francisco Aresta-Branco**, Mafalda Pimentel, Luísa M. Figueiredo.

3. IMM PhD Student Annual Meeting, Lisboa, Portugal, 2012. “Open chromatin conformation is independent of RNA polymerase I transcription in *Trypanosoma brucei*”, **Francisco Aresta-Branco**, Mafalda Pimentel, Luísa M. Figueiredo.
4. 10th EMBL Conference Transcription and Chromatin, Heidelberg, Germany, 2012. “Open chromatin conformation is independent of RNA polymerase I transcription in *Trypanosoma brucei*”, **Francisco Aresta-Branco**, Mafalda Pimentel, Luísa M. Figueiredo.
5. ISWOLD 2010, Rehovot, Israel, 2010. “The plasma membrane of yeast cells has highly ordered ergosterol-free sphingolipid-enriched lipid domains”, **Francisco Aresta-Branco**, André M. Cordeiro, H. Susana Marinho, Luísa Cyrne, Fernando Antunes, Rodrigo F. M. de Almeida.
6. FEBS Advanced Course on “Lipid signalling and disease”, Ortona, Itália, 2009. “The response of yeast plasma membrane to cell wall removal and to ergosterol biosynthesis impairment”, **Francisco Aresta-Branco**, André M. Cordeiro, H. Susana Marinho, Luísa Cyrne, Fernando Antunes, Rodrigo F. M. de Almeida.

WORKSHOPS AND COURSES ATTENDED

1. EMBO Young Investigator PhD Course *Spotlights on current biology*, 2013, Heidelberg, Germany.
2. Use of Large Datasets to Study Parasite Biology Course, 2012, Lisbon, Portugal
3. Launching Your Research Career Course, 2012, Lisbon, Portugal
4. Applied Biostatistics Course, 2012, Lisbon, Portugal
5. EuPathDB Workshop, 2011 Athens, Georgia, USA.

SUPERVISION EXPERIENCE

1. September 2013 – October 2014: Supervision of Idálio Viegas, Msc student with the thesis entitled “N6-methyladenosine: a new modification in *T. brucei* epitranscriptome”. **Grade: 20 out of 20**.
2. August 2011 – present: Supervision of several Summer/volunteer internships of 2 weeks – 1 month of the following undergraduate students:
    - Sílvia Pimenta: Medicine undergraduate that performed two different internships;
    - Afonso Bravo: Biology undergraduate;
    - Ana Margarida Venda: Biology undergraduate.

Published online 15 December 2015

Nucleic Acids Research, 2016, Vol. 44, No. 7 ***3131–3146***
doi: 10.1093/nar/gkv1459

A transcription-independent epigenetic mechanism is associated with antigenic switching in *Trypanosoma brucei*

Francisco Aresta-Branco, Silvia Pimenta and Luisa M. Figueiredo*

Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa 1649-028, Portugal

Received July 14, 2015; Revised October 30, 2015; Accepted November 28, 2015

ABSTRACT

Antigenic variation in *Trypanosoma brucei* relies on periodic switching of variant surface glycoproteins (VSGs), which are transcribed monoallelically by RNA polymerase I from one of about 15 bloodstream expression sites (BES). Chromatin of the actively transcribed BES is depleted of nucleosomes, but it is unclear if this open conformation is a mere consequence of a high rate of transcription, or whether it is maintained by a transcription-independent mechanism. Using an inducible BES-silencing reporter strain, we observed that chromatin of the active BES remains open for at least 24 hours after blocking transcription. This conformation is independent of the cell-cycle stage, but dependent upon TDP1, a high mobility group box protein. For two days after BES silencing, we detected a transient and reversible derepression of several silent BESs within the population, suggesting that cells probe other BESs before commitment to one, which is complete by 48 hours. FACS sorting and subsequent subcloning confirmed that probing cells are switching intermediates capable of returning to the original BES, switch to the probed BES or to a different BES. We propose that regulation of BES chromatin structure is an epigenetic mechanism important for successful antigenic switching.

INTRODUCTION

RNA Polymerase I (Pol I) transcribes ribosomal DNA genes (rDNA), which account for over 60% of total nuclear transcription (1). Most organisms have large tandem arrays of rDNA genes, but only a fraction is transcriptionally active. Consistent with their transcriptional activity, rDNA genes can be found in one of two chromatin states: a compact nucleosome-rich 'closed state' and an accessible nucleosome-depleted chromatin 'open state' (2). In yeast, replication of DNA converts the chromatin of most rDNA genes into the closed state. When replication is complete and transcription is reinitiated, a stochastic fraction of rDNA genes regains the open state, in a process that is dependent on Pol I. Once chromatin has an open state, its status is maintained by a high-mobility box protein, HMO1, independent from Pol I (3).

Trypanosoma brucei, a unicellular parasite responsible for sleeping sickness, is an unusual eukaryote that also uses Pol I to transcribe genes that encode two classes of abundant surface proteins, the variant surface glycoprotein (VSG) and procyclins (4,5). Periodic exchange of the exposed VSG allows the parasites to evade the host immune system, a process known as antigenic variation. Although the *T. brucei* genome has more than 2000 VSG genes and pseudogenes (6), only one is transcribed at any given time. *VSG*s are transcribed from bloodstream expression sites (BESs), specialized polycistronic units in which a Pol I promoter drives transcription towards the telomere. Between the promoter and the telomere there is a variable number of expression site-associated genes (ESAGs) followed by an array of tandem 70-bp repeats that precede the telomere-proximal *VSG* (7). Among the approximately 15 BESs present in the genome, only one is functionally active, ensuring the monoallelic expression that is the heart of antigenic variation.

VSG switching can happen by two main mechanisms, by recombining a new VSG into the active BES or by switching off a BES and activating another one (*in situ* or transcriptional switch) (8). VSG transcriptional switching involves two BESs, one that is silenced and another that is concomitantly activated. Cross-talk among BESs has been proposed to explain the phenotype observed when two BESs were simultaneously selected with drug selectable markers (9). Davies *et al.* also detected that deletion of a non-coding DNA sequence upstream of the active *VSG*, increased switching frequency to a new BES (10). More recently, Batram *et al.* showed that upon overexpression of an exogenous VSG, the active BES is partially attenuated into an intermediate stage, which may allow probing of silent BESs before commitment to one (11).

*To whom correspondence should be addressed. Tel: +351 219 999 512; Fax: +351 217 999 504; Email: lmf@medicina.ulisboa.pt

While in the mammalian-infective bloodstream stage, the chromatin of the active and silent BESs is dramatically different. The actively transcribed BES is nucleosome-depleted (open state), while silent BESs are organized in regularly spaced nucleosomes (12,13). The promoters of both active and silent BESs are bound by the multi-subunit class I transcription factor A (CITFA) (14), although to a different extent (15), which is necessary to regulate BES transcription initiation. Several epigenetic factors have been shown to be necessary to prevent transcription from silent BESs (16); others are necessary to ensure fast switching between BESs (17), but only TDP1 has been shown to be a core component of the active BES. TDP1 is a high mobility group box protein that is present in the chromatin of active BES, and at rRNA genes, and it is necessary for their transcription (18,19). The interplay among these factors and the mechanisms by which they affect *VSG* transcription are essentially unknown.

During the life cycle of *T. brucei*, parasites shift between the mammalian host and the Tsetse fly (20). To survive in such different hosts, parasites undergo significant changes in gene expression, which include replacement of the VSG by the similarly abundant procyclins (21). Silencing of the active BES during differentiation is characterized by a progressive downregulation of transcription along the BES (22), together with a re-localization of the BES to the heterochromatic nuclear periphery (23). It has also been shown that the chromatin of the originally active BES becomes less open to T7 polymerase, suggesting that it acquires a more compact structure during differentiation (24).

The goal of this manuscript was to understand the interplay between transcription and chromatin during antigenic variation. We characterized the early events that take place when the active BES is silenced. We found that trypanosomes have a cell-cycle and transcription-independent mechanism to maintain the open chromatin conformation of the active BES, which is dependent upon TDP1. We also observed that, in the first two days after transcriptional silencing is induced, parasites experience an intermediate stage in which several previously silent BESs are temporarily transcribed at higher levels, suggesting that cells probe different BESs before commitment to a new single BES. Our findings provide evidence that regulating chromatin conformation is tightly associated to antigenic switching, an important virulence mechanism of this pathogen.

MATERIALS AND METHODS

Trypanosome cell-lines and plasmid construction

T. brucei bloodstream form (BSF) parasites (strain Lister 427, antigenic type MiTat 1.2, clone 221a) (25) were cultured in HMI-11 as described in (26). Differentiation studies were performed in PL1A, a cell-line described in (27). All transfections were made with an AMAXA nucleofector (Lonza), program X-001, using the previously optimized homemade Tb-BSF buffer (90 mM sodium phosphate, 5 mM potassium chloride, 0.15 mM calcium chloride, 50 mM HEPES, pH 7.3) (28). GLB1 and its derivative cell-lines were modified from the parental 2T1.ESPiGFP:NPT (29), in which the *NPT* gene was replaced by a PCR product containing a stop codon for the *GFP* gene, an *Aldolase* 3′ UTR, *Luciferase* and *BSR* genes. PCR product was obtained using primers with long tails (provided in Supplementary Table S1) that serve as target recombination regions for *GFP* ORF and the BES sequence downstream of *GFP::NPT* 3' UTR.

GLB1-TDP1::TY1 was generated by transfecting pFAB11, which inserts a TY1 tag at the 3' end of one of the *TDP1* endogenous alleles. pFAB11 contains a *Hygromycin resistance* gene, a 5'-truncated *TDP1*, a TY1 tag at the 3'-end of TDP1 and an *Aldolase* 3'UTR. pFAB11 was linearized with *Sma*I (New England Biolabs), which digests in the middle of *TDP1* ORF. GLB1-TDP1::3xcMyc was generated by transfecting pFAB14 which inserts a triple MYC tag in the 3' end of one of the *TDP1* endogenous alleles. pFAB14 contains a *Hygromycin resistance* gene, a 5'-end truncated *TDP1*, a triple MYC tag in the 3'-end of *TDP1 and* the predicted endogenous 3'UTR of *TDP1* (30). pFAB14 was linearized with *Sma*I (New England Biolabs), which digests in the middle of *TDP1* ORF.

GLB1-R15 was obtained by first generating switchers of GLB1 using the BES silencing inducible assay and subsequently transfecting pFAB17 in switchers. pFAB17 contains a 405 nucleotide sequence upstream BES promoter, the BES promoter, an *RFP* ORF lacking NLS, PEST and stop codons, amplified from pCAGGS, a *NPT* ORF lacking the start codon, an *Actin* 3'UTR and a 475 nucleotide sequence downstream of the BES promoter. *RFP* and *NPT* ORFs formed a fused *RFP::NPT* gene. pFAB17 was digested with *Asc*I and *Nde*I (New England Biolabs) prior to transfection. Cloning of VSG transcripts and sequencing allowed identification of the BES in which *RFP::NPT* had integrated. Transfected switchers were reverted to the original BES, essentially as described in (7) by removing G418 and adding 1 μg/ml of tetracycline to the medium. Afterwards, serial dilutions were performed and 10 μg/ml of Blasticidin-S was added to each dilution to select revertants.

All cloning was performed using the In-Fusion® HD Cloning system (Clontech) following to the manufacturer's instructions. The maps of constructs generated in this work are shown in Supplementary Figure S1.

Differentiation assay

Parasites were collected at a density of 1.0–1.5 × 10^6 cells/ml and centrifuged at 650 g for 10 min at room temperature. Cells were resuspended in DTM medium with 6 mM of *cis*-aconitate (Sigma) at a density of 1.5–2 × 10^6 cells/ml and grown at 27°C without CO_2.

Transcript quantification

Parasites were harvested by centrifugation at 650 g for 10 min at 4°C and immediately resuspended in PureZOL (Bio-Rad) or TRIzol (Life Technologies). RNA was isolated following the manufacturer's instructions and RNA quantity and quality was assessed on a NanoDrop 2000 (Thermo Fisher Scientific). cDNA was generated using a Superscript cDNA Synthesis Kit (Life Technologies), according to manufacturer's protocol. Quantitative PCR (qPCR) was performed using 1× SYBR Green PCR Master Mix (Applied Biosystems). Negative controls lacking reverse transcriptase (RT−) were confirmed by qPCR. Amplification

reactions were performed in duplicates. The ΔΔCt method was used to determine transcript levels relative to normalizing gene.

Tetracycline-inducible BES silencing assay

Parasites were centrifuged at 650 g and washed three times with warm HMI-11. Pellets were resuspended in medium with drugs except Blasticidin-S and tetracycline and density was adjusted to 0.5×10^6 cells/ml. Cells were split in two flasks with and without Blasticidin-S and tetracycline (Tet+ and Tet-, respectively).

Survival frequency and commitment assay

To determine the percentage of cells that survive the BES-silencing assay, after washing away drugs, cells were diluted to a density of 10 cells/ml and plated in two 96-well plates with or without tetracycline (Tet+ and Tet-, respectively). Seven days after plating, surviving clones were counted in Tet- and Tet+ plates and its ratio yielded the survival frequency. For commitment assay, cell density of Tet- cultures was determined at 8, 24, 48 and 72 h after BES silencing and subsequently diluted to a density of 10 or 50 cells/ml containing 1 μg/ml of tetracycline. Dilutions were performed with the same drugs and complemented with 1 μg/ml of tetracycline. Diluted cells were then plated in two 96-well plates. All plates containing tetracycline were replenished with fresh tetracycline three days after plating to maintain excess concentration. Six days later, around 20 surviving wells were passaged to new 96-well plates and were analyzed by a FACS High Throughput Sampler (BD Biosciences) to score for expression of GFP.

Cell-cycle profile after BES silencing

At each time point after inducing BES silencing, 2 million cells were centrifuged for 10 min at 1300 g, 4°C and washed once with ice-cold PBS. Cells were resuspended in PBS with 2 mM EDTA and slowly fixed with 2.5 ml of absolute ethanol. After fixing for at least one hour, cells were washed once and resuspended in 1 ml PBS/EDTA. Cells were incubated with 10 μg RNase A and 1 μg of Propidium Iodide for 30 min at 37°C and further analyzed by flow cytometry for DNA content.

Luciferase assay

1.5×10^6 cells were harvested for 5 min at 2800 g, 4°C and washed once with 1 ml cold TDB. Pellets were resuspended in Lysis Buffer (Biotium) and protocol was followed according to the manufacturer's instructions. Luminescence was measured by a microplate reader (Tecan).

FAIRE and Chromatin Immunoprecipitation

Formaldehyde-assisted isolation of regulatory elements (FAIRE) was performed as described previously (31).

H3 and TDP1 ChIPs were carried out in $3–5 \times 10^7$ cells essentially as described elsewhere (12), but with several modifications. Cells fixed for 20 min in a 1:11 dilution of a formaldehyde solution (50 mM HEPES–KOH pH7.5, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA and 11% formaldehyde) in HMI-9 medium. DNA was sonicated in a Bioruptor UCD-200 (Diagenode) for 10 min total (30 seconds, on and off cycles). Lysate was incubated with Dynabeads Protein G (Life Technologies) combined with 10 μg of rabbit anti-H3 antibody (kind gift from Christian Janzen) or mouse anti-TY1 antibody (32) overnight at 4°C. Samples were treated with 80 μg of RNase A for 2 h at 37°C and then with 80 μg of Proteinase K for 2 h at 55°C before being purified using a Gel Extraction Kit (Qiagen). Immunoprecipitated material was quantified by qPCR. Amplification reactions were performed in duplicates.

RPA31 ChIP was carried out in 10^8 cells as described by the Arthur Günzl's laboratory (33), but immunoprecipitation was performed with Protein G Sepharose 4 Fast Flow (GE Healthcare) with a rat anti-RPA31 antibody (kind gift from Arthur Gunzl).

FACS cell sorting

To sort cells based on cell-cycle stage, 10^8 cells were centrifuged for 10 min at 650 g, resuspended in 20 ml of HMI-11 and cross-linked with 1.1% formaldehyde (Sigma) for 10 min. Cross-linking was stopped with 0.125 M of Glycine (Sigma) and cells were washed once with PBS and resuspended in 1 ml PBS/2 mM EDTA. Cells were permeabilized for 5 min with 40 μM digitonin, washed twice with PBS and resuspended in 1 ml PBS/2 mM EDTA. Staining was performed with 1 μl FxCycle Violet Stain (Molecular Probes, Thermo Fisher Scientific) for 30 min protected from light. Cells were sorted according to their DNA content using a FACS Aria sorter (BD Biosciences) and collected in PBS.

To sort cells based on GFP/RFP expression, tetracycline was removed from medium and 24 h later, 2×10^6 cells were collected, washed in TDB and analysed by FACS for GFP and RFP expression. 3000 cells were sorted according to the GFP/RFP expression. The same number of cells was also sorted from a non-induced culture condition (Tet+). 200 (from Tet+ condition) or 1000 cells (from Tet- conditions) were plated in two 96-well plates with or without tetracycline, in the presence of 100 U/ml of PenStrep (Life Technologies). Six days later, surviving wells were passaged to new 96-well plates and were analyzed by a FACS High Throughput Sampler (BD Biosciences) to score for expression of GFP and RFP.

dsRNA production

dsRNA against TDP1 was made using the MEGAscript RNAi Kit (Ambion, Thermo Fisher Scientific) and protocol was followed according to the manufacturer's instructions. The sequence used to deplete TDP1 mRNA is identical to the one published in (18). 20 μg of dsRNA or the same buffer volume (for the mock control) were transfected into 100 million cells, using the X-001 program, and the homemade Tb-BSF buffer in an AMAXA nucleofector. The primers used for amplification of DNA template include a T7 promoter for *in vitro* transcription and are presented in Supplementary Table S1.

Western blotting

2×10^6 cells were lysed using Laemmli sample buffer and treated with 200 U/ml Benzonase (Sigma). After SDS-PAGE separation, proteins were transferred to a nitrocellulose membrane using an iBlot Dry Blotting System (Invitrogen). Primary antibodies used in this work were diluted as follows: 1:1000 for mouse anti-cMyc (clone 9E10, Antibody and Bioresource Core Facility, The Rockefeller University), 1:5000 for rabbit anti-*T. brucei* H2A antibody (custom made), 1:1000 for mouse anti-β-tubulin KMX antibody (34) (gift from Keith Gull). Picture acquisition was made by a Chemidoc XRS+ (Bio-Rad). Quantification was performed using the Image Lab software (Bio-Rad).

VSG staining of live cells

0.5×10^6 cells were harvested for 5 min at 2800 g, 4°C. Cells were resuspended in 50 μl of cold HMI-11 in which Alexa Fluor 647 anti-VSG13 conjugated antibody had been previously diluted (1:5000). After 15 min of incubation at 4°C with gentle shaking, cells were washed three times in cold TDB, resuspended in cold TDB, and immediately analyzed on a FACS Fortessa (Becton Dickinson Biosciences). Data were processed with FlowJo software (FlowJo, LLC).

RESULTS

Transcription and chromatin dynamics during differentiation

When a bloodstream parasite undergoes transcriptional VSG switching or when it differentiates into procyclic forms, the active BES must be silenced and its chromatin closed (12,13,24). In both processes, the interplay between transcription and chromatin is unknown. What is the order of these events: does chromatin condensation lead to a transcriptional silencing, or does transcription stop first and chromatin is condensed later? We began to tackle this problem by characterizing what happens during differentiation (Figure 1), and then in BES switching (Figure 2).

To study the interplay between transcriptional silencing and chromatin condensation in the active BES during differentiation, we induced cell differentiation *in vitro* by adding the chemical trigger *cis*-aconitate and lowering the temperature from 37°C to 27°C. We measured levels of mRNA at different time points, using a cell-line in which BES1 is actively transcribed (Figure 1A). As expected, *VSG2* transcript levels rapidly decreased, reaching 2% of the bloodstream form levels, 24 h after inducing differentiation (Figure 1B). Ribosomal DNA (rDNA) *18S* transcripts remained unchanged, as expected (22), and transcript levels of silent VSG9 were slightly reduced. As previously observed, transcript levels of RNA Polymerase II (Pol II)-transcribed genes, *β-tubulin* and *GAPDH*, also decreased (22,35,36), the second reflecting the reduced metabolic dependence of glycolysis in procyclic forms.

To test if this transcriptional silencing of the active BES was a consequence of chromatin conformational changes, we assessed nucleosome occupancy during differentiation by FAIRE (37) (Figure 1C). This technique allows the purification and quantification of DNA with low protein content from what we call a more 'open' chromatin conformation (12,31). We observed that, throughout the first 24 h of differentiation, chromatin of silent *VSG9*, *β-tubulin* and *GAPDH* kept the same FAIRE-enrichment, suggesting that the chromatin remained equally condensed. Importantly, at BES1, although chromatin condensed slightly (2.5-fold at *VSG2*), it remained 12-fold more open than a silent *VSG*.

These results show that, during differentiation, the drop in mRNA levels of the active BES starts earlier and it is more pronounced than the condensation of chromatin (Figure 1D), suggesting that chromatin does not close immediately after transcription is halted.

Transcription and chromatin dynamics during BES switching

Next we hypothesized that, during BES switching, chromatin condensation may also be delayed relative to transcription silencing. We postulated that, if open chromatin of the active BES was exclusively the consequence of high transcription, induction of transcription silencing should lead to an immediate condensation of chromatin. To test this hypothesis, we used a BES-inducible silencing system to block the transcription at the active BES and we followed transcription and chromatin dynamics. Because BES switching happens at very low frequency, we adapted a previously established reporter strain in which BES1 has a tetracycline operator sequence at the promoter followed by a GFP reporter (29). When tetracycline is removed from the culture medium, the heterologous tetracycline repressor is free and binds to the tetracycline operator, thus sterically blocking Pol I transcription. We introduced a *luciferase* gene downstream of the *GFP* gene since this is a more sensitive transcriptional reporter (Figure 2A). This cell-line was named GLB1, for *GFP*, *luciferase* and *BSR* in BES1.

During the first 8 h after tetracycline was removed from the medium (Tet-), GLB1 cells grew at the same rate as the control (Tet+), after which cell growth lagged until 48–72 h (Figure 2B). This lag phase was characterized by abnormal cell morphology and motility (data not shown) and cell-cycle arrest in G_2/M, with a considerable accumulation of cells with polyploidy or abnormal DNA content (Figure 2C). Quantification of cell death using propidium iodide, which only stains dead cells or in late apoptosis (38) (Supplementary Figure S2), showed a significant number of stained cells from 24 to 72 h after tetracycline removal, reaching a peak of 7.5% at 48 h. Consistent with this significant number of dead cells for several days, clonogenic assay confirmed that only 8% of the initial population of cells survive the silencing assay (Figure 2D). After 96 h, surviving cells took over the culture and grew at normal rate of around 8 h per population doubling. As expected, all surviving clones no longer expressed GFP, indicating they had successfully switched to a new BES (Figure 2E).

To test the efficiency of the steric blockade of Pol I upon tetracycline removal, we checked whether Pol I was evicted from BES1 chromatin fiber by conducting ChIP against the RPA31 subunit of Pol I (Supplementary Figure S3). As expected, before BES silencing, Pol I was present in the *18S* rDNA gene and in the active BES genes (*Luciferase* and *VSG2*) and essentially absent from Pol II loci (*β-tubulin* and *GAPDH*). 8 hr after BES silencing (tetracycline removal), the levels of Pol I in the active BES decreased to similar background levels detected in the silent *VSG9, β-tubulin*

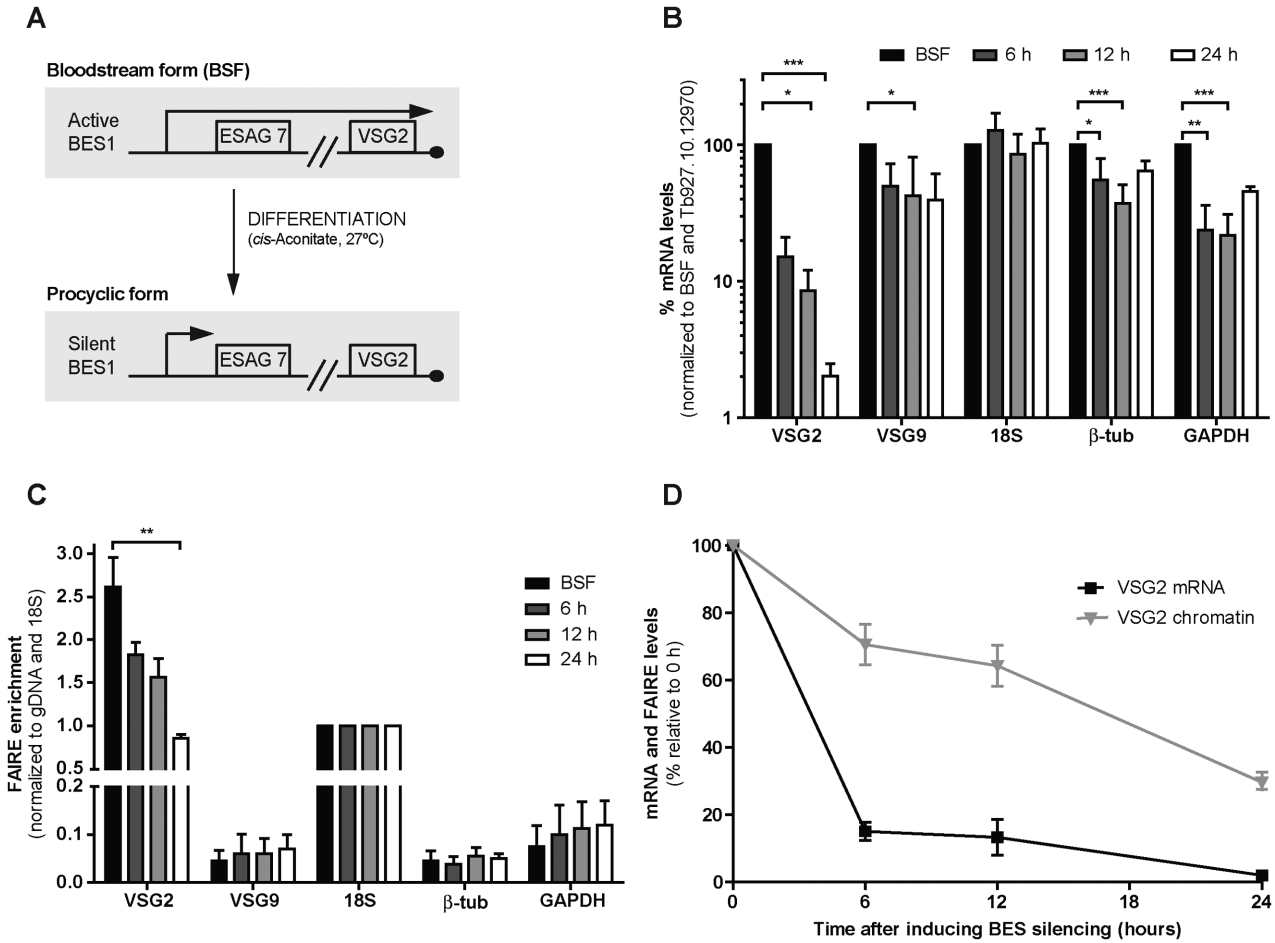


Figure 1. BES transcriptional silencing precedes chromatin condensation during differentiation. (**A**) A bloodstream form cell-line that expressed *VSG2* (BES1) was differentiated to procyclic forms by adding *cis*-aconitate to the medium and changing temperature from 37°C to 27°C. During differentiation, BES1 is silenced. Procyclic forms do not express VSG at the surface. Quantification of mRNA levels and (**B**) chromatin conformation (**C**) 6, 12 and 24 h after induction of differentiation. (**B**) Transcript levels were measured by qPCR and normalized to bloodstream form (BSF) levels and *Tb927.10.12970* (56), a gene previously shown to maintain constant transcript levels during differentiation. Four to six independent experiments were analyzed. (**C**) DNA purified from FAIRE was quantified by qPCR and normalized to *18S* as its transcript levels also remained constant throughout differentiation (23) and FAIRE signal was more intense than *Tb927.10.12970*. Three to five independent experiments were analyzed. (**D**) Comparison between transcript levels and FAIRE enrichment for *VSG2* gene. Values were extracted from analysis in (**B**) and (**C**). Statistical significance was determined by one-way ANOVA with Bonferroni post-test comparison.. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

and *GAPDH*. Hence, Pol I is indeed efficiently removed from chromatin fiber upon BES silencing.

To characterize the dynamic of chromatin structure once transcription is halted, we focused on the first 8 h because this is the period in which cells grow well and present a normal morphology (Figure 2B, Supplementary Figure S2). As a proxy of transcription, we followed luciferase activity 2, 5 and 8 h after removing tetracycline (Figure 3A). Luciferase activity showed an exponential decrease to ∼20% of the initial activity at 5 hr and to only 5% 8 h after tetracycline removal. Luciferase activity was confirmed by quantifying mRNA transcript levels 8 hr after inducing BES silencing (Figure 3B). *18S* rDNA and *β-tubulin* did not suffer transcriptional changes, confirming that BES silencing only affected expression sites and it did not cause any other major indirect changes in the rest of the genome. *Luciferase* and *BSR* transcripts decreased 90% of the initial levels, while active *VSG2* decreased less (32%) probably due to its stability (half-life around 4.5 h) (35). Concomitantly, we observed a 4–8-fold transcriptional up-regulation of several silent VSGs, which is likely a result from either newly activated BESs or derepressed silent BESs.

To determine the chromatin structure of the inducibly silenced BES, we performed FAIRE (Figure 3C) and histone H3 chromatin immunoprecipitation (ChIP) (Figure 3D). By FAIRE, we observed that, both at 5 and 8 h after BES silencing, the chromatin at the active BES remained highly enriched in the aqueous phase, indicating an open conformation. During this period, FAIRE-enrichment of silent BESs (*VSG9* and *VSG13* genes) remained unchanged. Although at 8 h the chromatin of the active *VSG2* presented a 2-fold decrease in FAIRE-enrichment, it was still 123-fold higher than FAIRE-enrichment of the same gene six days post-silencing, when this expression site was completely silenced. Six days post-silencing, genes from previously silent BESs (*VSG9* and *VSG13*) showed an increase of three-fold

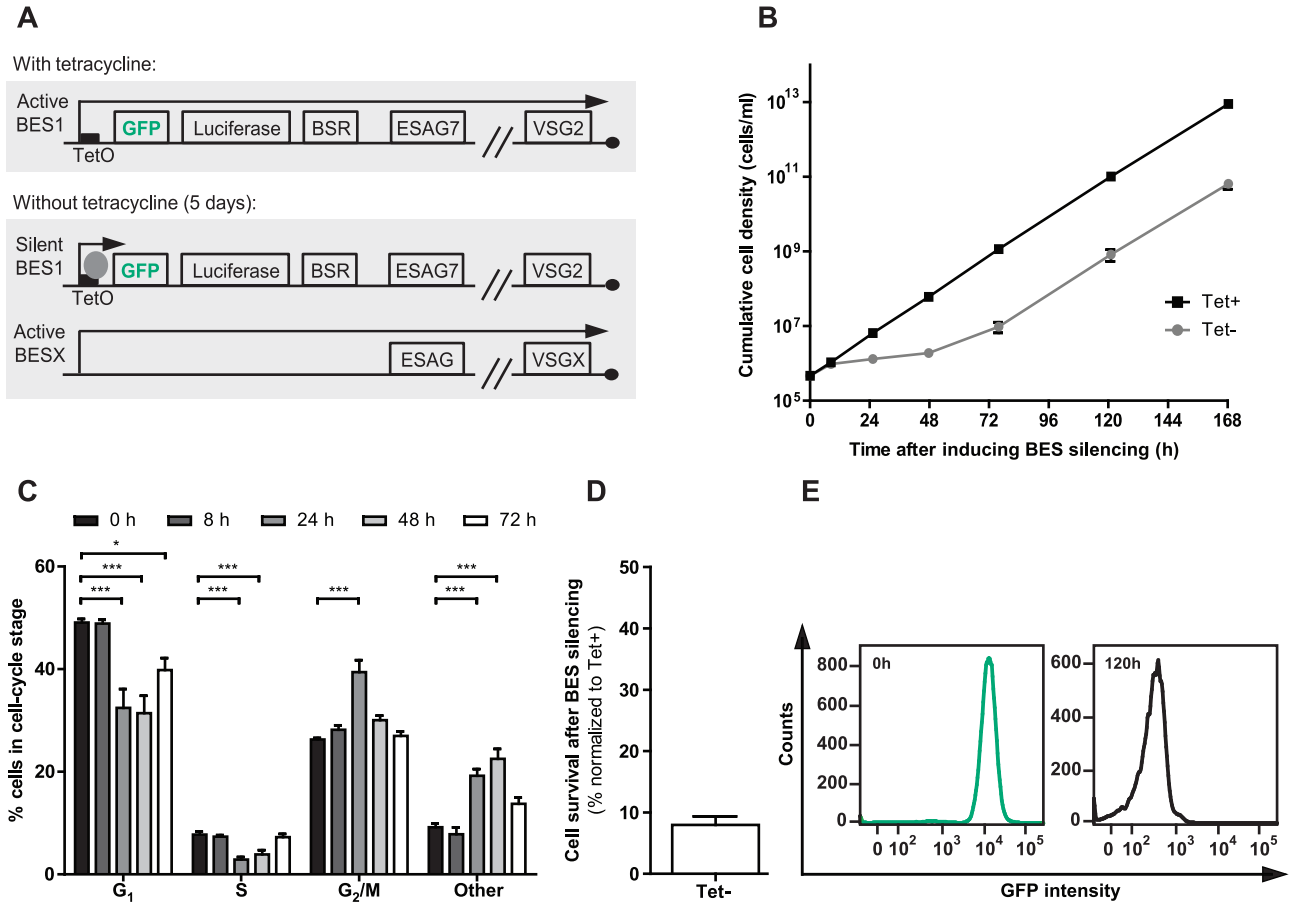


Figure 2. BES silencing causes growth delay and G_2/M cell-cycle arrest but only after more than 8 h of induction. (**A**) In the bloodstream reporter cell-line, GLB1, removing tetracycline from the medium induces BES1 silencing and subsequent activation of another BES. **Upper panel** shows that in the presence of tetracycline, BES1 is actively transcribed. BES1 contains a tet operator sequence (TetO, black rectangle), *GFP*, *luciferase* and *BSR* genes downstream of the promoter. **Lower panel** shows the outcome of tetracycline removal: BES1 becomes silent (because tetracycline repressor (grey circle) binds TetO, sterically blocking Pol I transcription) and a new BES is activated. *BSR*, Blasticidin-S Resistance. (**B**) Growth curves of cells in the presence (Tet+, black curve) or after removal (Tet-, grey curve) of tetracycline. Four independent experiments were analyzed. (**C**) Cell-cycle profile of GLB1 at different time-points after removal of tetracycline. 'Other' represents cells with abnormal DNA content. Four independent experiments were analyzed. Statistical significance was determined by a two-way ANOVA with Bonferroni post-test comparison. $*P < 0.05$; $***P < 0.001$. (**D**) Percentage of GLB1 cells that survive the BES silencing induction was determined by a clonogenic assay and normalized to Tet+ cells. Four independent experiments were analyzed. (**E**) Flow cytometry analysis of GFP expression of cells at 0 h (**left panel**) and 120 h (**right panel**) after tetracycline removal.

in FAIRE enrichment. This increase was expected not to be maximal (up to around 100) because FAIRE was performed on a mixed population of switchers. As expected, *18S* rDNA and *β-tubulin* did not show major chromatin alterations at any time. These results were mirrored by histone H3 ChIP: at 8 h, chromatin of originally active BES1 (*BSR* and *VSG2*) was still heavily depleted of histone H3, which is consistent with a high FAIRE-enrichment. *VSG2* had slightly more histone H3, but the levels were still much lower than those detected at silent *VSG9* or *VSG13*. As expected for a healthy parasite population, ChIP of other control genes was not affected during the first 8 h post-silencing induction, including at the *18S* rDNA.

Overall, our results show that, during BES switching, chromatin condensation lags significantly behind transcriptional silencing, suggesting that *T. brucei* has a mechanism of maintaining chromatin open when transcription has been halted.

Chromatin conformation is cell-cycle independent

In yeast, the ratio between open and closed rDNA genes changes throughout the cell-cycle: entrance into S phase leads to repression of most rDNA genes and their chromatin becomes more compact, while transcription re-initiation in G_2 re-opens chromatin (3). As BESs are transcribed by Pol I, we hypothesized that chromatin of active BES may also close as a function of the cell-cycle. To test this hypothesis, we stained GLB1 fixed cells with FxCycle Violet DNA stain and we FACS-sorted them, according to the DNA content, into G_1, S and G_2/M subpopulations (Figure 4A). The chromatin conformation of these subpopulations was subsequently assessed by FAIRE (Figure 4B). The first observation was that the FAIRE-enrichment of sorted subpopulations revealed patterns very similar to unsorted cells (*BSR*, *luciferase* and *18S* rDNA 50–100; *VSG9*, *β-tubulin* and *GAPDH* around 1) (Figure 3C), suggesting that the sorting procedure did not affect chromatin confor-

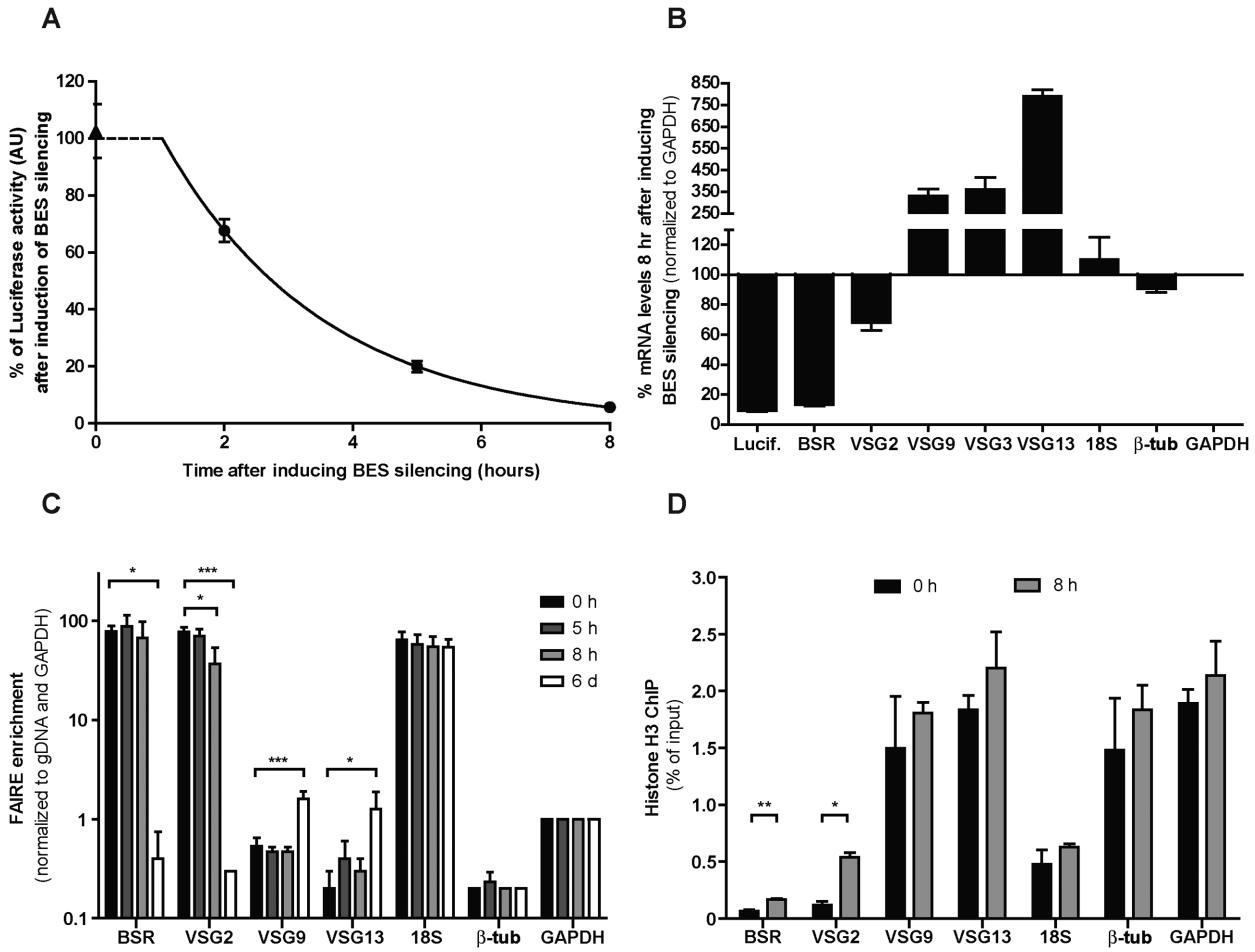


Figure 3. BES chromatin retains an open conformation despite its transcription being reduced 90%. (**A**) % of luciferase activity of Tet- relative to Tet+ cells after 0, 2, 5 and 8 h of tetracycline removal. Curve represents the best decay fit for time-points 2, 5 and 8 h. Five to seven independent experiments were analyzed for each time-point. (**B**) % of mRNA levels after 8 h of tetracycline removal relative to Tet+ cells, measured by qPCR and normalized to *GAPDH* transcripts. Five independent experiments were analyzed. (**C**) Chromatin conformation was measured by FAIRE at 0, 5, 8 h and 6 days after tetracycline removal. DNA isolated by FAIRE was quantified by qPCR and normalized to gDNA copy number and *GAPDH*. Three independent experiments were analyzed. Statistical significance was determined by one-way ANOVA with Bonferroni post-test comparison. (**D**) Nucleosome occupancy was determined by histone H3 ChIP at 0 and 8 h after tetracycline removal. Immunoprecipitated DNA was compared to the total input material. Three independent experiments were analyzed. Statistical significance was determined by a paired *t*-test against time-point 0 h. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$.

mation of these genomic loci. Second, for each gene, the FAIRE-enrichment was constant overall for the three stages of the cell-cycle, suggesting that chromatin conformation is essentially insensitive to cell-cycle. We cannot exclude the possibility that very rapid and transient chromatin changes in conformation take place, which could not be captured at the time-resolution used here.

Unlike rDNA genes in yeast, we observed no major changes in the chromatin conformation of the active BES throughout the cell-cycle, which suggests that the mechanism that keeps BES chromatin open in the absence of transcription open is very likely cell-cycle independent.

TDP1 maintains chromatin open in the absence of transcription

In budding yeast, a Pol I transcriptional restriction for four hours under G_1 arrest does not affect the conformation of rDNA chromatin, which is maintained by HMO1, a high mobility group box (HMGB) protein (3). In. *T. brucei*, TDP1, a high-mobility group box protein, facilitates transcription of Pol I transcribed genes (18). Here, we tested if TDP1 was necessary to maintain chromatin open in the absence of Pol I transcription. For this, we tagged an endogenous allele of *TDP1* with a TY1 epitope and we performed ChIP. As previously reported (18), we found TDP1 highly enriched in the active BES (*Luciferase* and *VSG2* genes) and *18S* rDNA (0 h) (Figure 5A). After 8 h of transcriptional silencing, the chromatin of the promoter-proximal region of the BES contained the same amount of TDP1, whereas the telomeric region had two-fold less TDP1. Nonetheless, these levels were still higher (7–9-fold) than in silent *VSG*s. As TDP1 inversely correlates with histone H3 (18), this result is in accordance with the H3 distribution after 8 h of BES silencing (Figure 3C and D), and it shows that TDP1 is still largely present in the active BES when its transcription is reduced by at least 90% (Figure 3A). The surviving switchers (six days post-silencing induction) presented, as

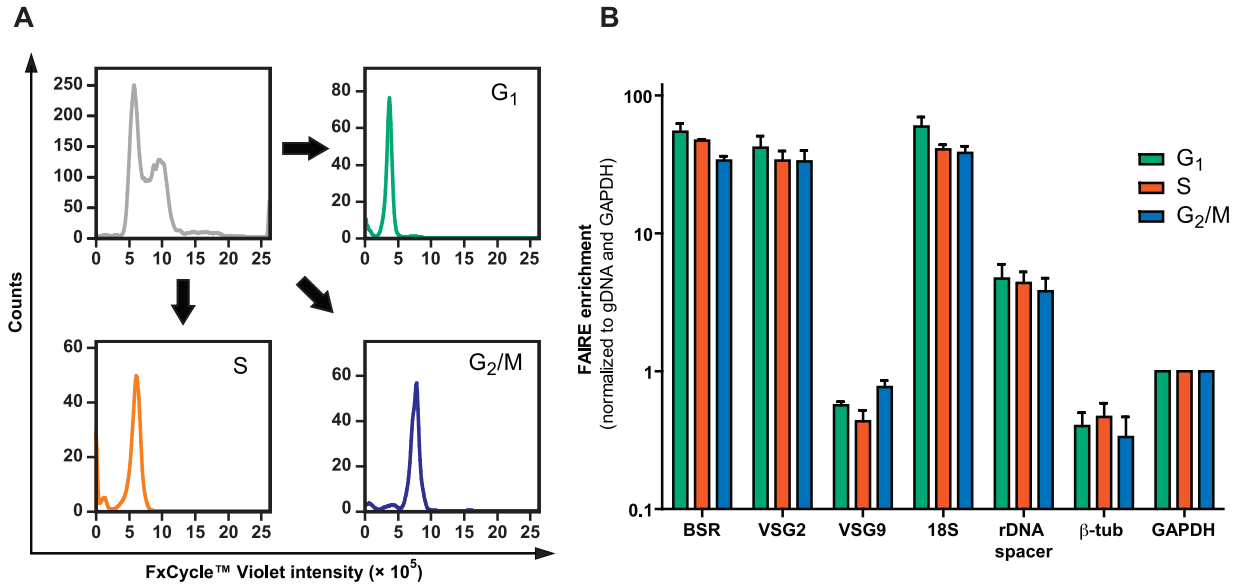


Figure 4. Chromatin conformation of BESs is cell-cycle independent. (**A**) Nuclear DNA of GLB1 cells was stained by FxCycle Violet and sorted by flow cytometry in G_1 (red), S (orange) and G_2/M (blue) cell-cycle stages. Panel in grey represents original population, colored panels represent analysis of post-sorted populations. (**B**) Chromatin conformation of the three cell-cycle populations was measured by FAIRE. DNA isolated by FAIRE was quantified by qPCR and normalized to gDNA copy number and *GAPDH*. Three independent experiments were analyzed. Statistical significance was determined by a 1-way ANOVA with Bonferroni post-test comparison.

expected, very low levels of TDP1 in the previously active BES.

To test if TDP1 is necessary to maintain chromatin open when the active BES is silenced, we depleted TDP1 by transiently transfecting an anti-TDP1 dsRNA (18). This method has been previously used to knock-down α-tubulin transcripts (39). Transfection with buffer (mock control), showed no changes in TDP1 levels over time (Figure 5B, Supplementary Figure S4A). However, transfection of anti-TDP1 dsRNA lead to a reduction of TDP1 protein levels to ∼18% after 24 h and remained low until 48 h, after which TDP1 levels increased (Figure 5B). Consistent with a previous report (18), we observed that 24 h of TDP1 depletion resulted in a ∼43% reduction of luciferase activity, confirming the role of TDP1 as a transcriptional facilitator of the active BES (Supplementary Figure S4B). To test if TDP1 is necessary to maintain open chromatin status in the absence of transcription, we transfected TDP1 dsRNA or buffer into the BES inducible reporter strain and, after 19 h, we transcriptionally silenced the active BES by removing tetracycline from the medium. FAIRE was used to characterize chromatin status 5 h post-silencing (which corresponds to 24 h post TDP1 depletion) in four conditions: when BES silencing was induced or not, and in the presence or absence of TDP1 (Figure 5C). As expected, chromatin changes were not detected for control genes: Pol II-transcribed genes, nor in silent *VSG9*. Chromatin of *18S* rDNA, although transcribed by Pol I, was not affected either, which is consistent with observations in yeast, in which HMO1-null mutants do not lead to condensation of rDNA chromatin (3). Relative to non-transfected control, mock control cells did not display changes in chromatin in any loci, including BES1, indicating that transfection *per se* did not significantly affect chromatin conformation (Figure 5C).

When TDP1 was depleted but BES1 remained active (TDP1 KD, Tet+), chromatin remained open in the promoter region and began to close at the telomere (*VSG2* is 2.6-fold more closed than mock transfection), which is consistent with the observations of Narayanan *et al.* upon depletion of TDP1 by RNA interference (18). However, when TDP1 was depleted and BES1 was silenced (TDP1 KD, Tet-), chromatin of *VSG2* closed even further and chromatin of *luciferase* also closed slightly (Figure 5C). Next we compared the change of FAIRE-enrichment after and before silencing was induced (Tet- versus Tet+), in both conditions: mock transfection (presence of TDP1) and upon TDP1 knock-down (Figure 5D). For most genes in both conditions, this fold-change of FAIRE-enrichment is around 1, indicating that chromatin after 5 h of silencing is not dramatically different and absence of TDP1 does not cause global changes in chromatin during this period. However, in the *VSG2* gene we detected a significant reduction in the fold-change of FAIRE-enrichment between mock and TDP1 depleted conditions. These results indicate that TDP1 is necessary to keep an open chromatin conformation when transcription of BES is halted.

We conclude that TDP1 is a key player of the chromatin of active BES, especially at the telomeric end. Narayanan *et al.* has previously shown that TDP1 maintains chromatin nucleosome-depleted and facilitates transcription. Here we show that TDP1 is crucial to keep an open chromatin in a BES that has been recently silenced (especially at the telomeric end of BES).

BES probing precedes commitment to switching

Next we investigated the advantages of having a mechanism that keeps chromatin open when transcription has been silenced. We hypothesized that this may be a mechanism that

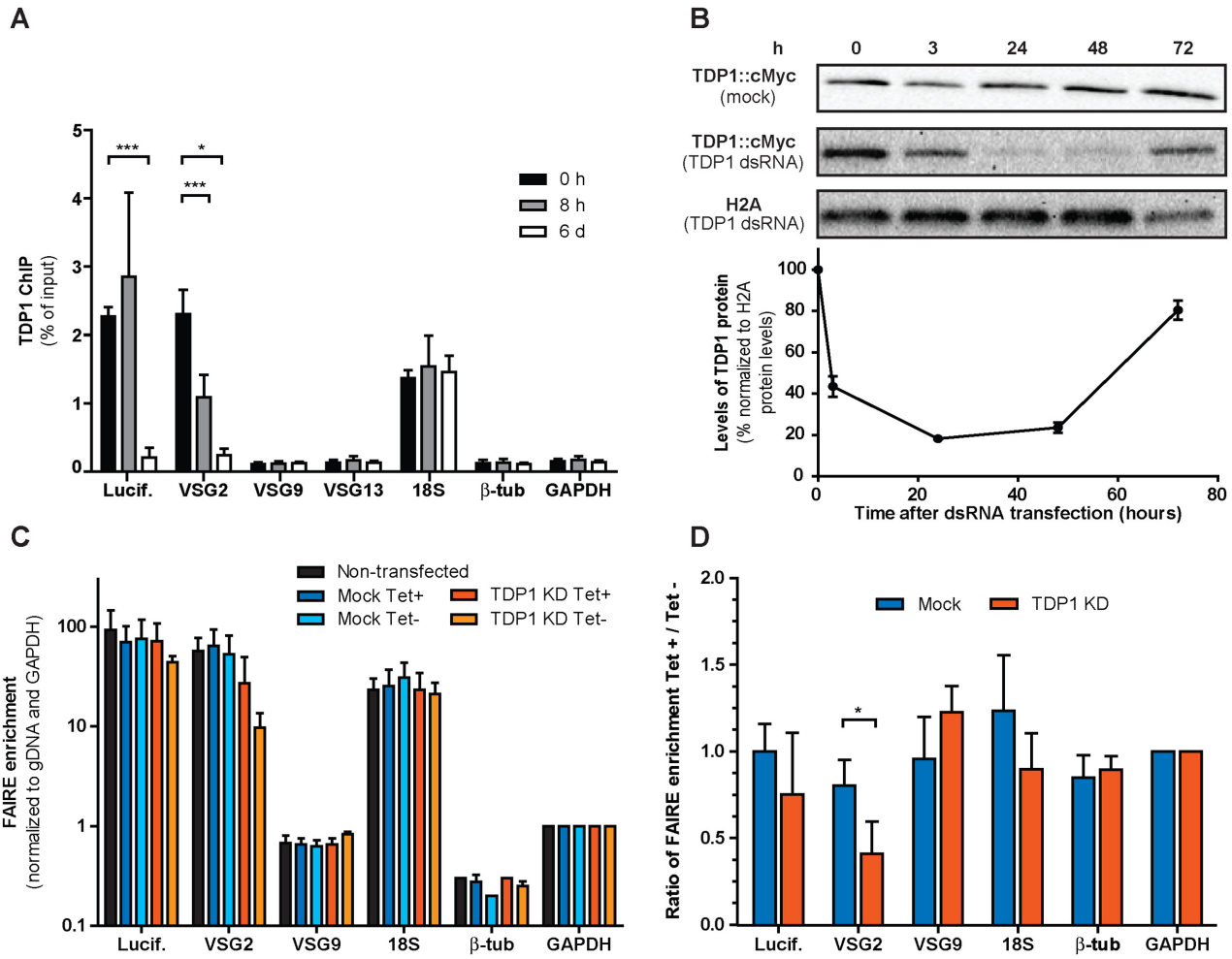


Figure 5. When active BES is silenced, TDP1 keeps its chromatin open. (**A**) TDP1 ChIP at 0, 8 h and 6 days after tetracycline removal in GLB1-TDP1::TY1, a cell-line in which one endogenous allele of TDP1 is fused with a TY1 tag. Immunoprecipitated DNA was compared to the total input material and normalized to *18S* DNA. Statistical significance was determined by a one-way ANOVA with Bonferroni post-test comparison. Three independent experiments were analyzed. (**B**) Western blotting analysis of TDP1 protein after 3, 24, 48 and 72 h of transfection with buffer (mock) or anti-TDP1 dsRNA in GLB1-TDP1::3xcMyc, a cell-line in which one endogenous allele of TDP1 is fused with a triple c-MYC tag that is more sensitive for western blot. Time-point 0 h indicates mock control cells transfected only with buffer. Each lane corresponds to lysates from 2×10^6 cells. Quantification of TDP1 signal is indicated in the **lower panel**. TDP1 protein levels were normalized for H2A protein levels and mock control. Four independent experiments were analyzed. (**C**) Chromatin conformation of GLB1-TDP1::3xcMyc cells after 5 h of BES silencing (24 h of TDP1 depletion) was measured by FAIRE. DNA isolated by FAIRE was quantified by qPCR and normalized to gDNA copy number and to *GAPDH*. Four independent experiments were analyzed. Statistical significance was determined by an unpaired t-test comparing Tet- to Tet+ in each condition (Mock or TDP1 KD). (**D**) Ratio of FAIRE enrichment (calculated from data in panel C) between Tet- and Tet+ for Mock and TDP1 KD conditions. Statistical significance was determined by an unpaired *t*-test. $*P < 0.05$; $***P < 0.001$.

allows parasites to probe different silent BESs before committing to a new BES. To test this hypothesis, first, we determined how long cells take to commit to a second BES. For that, we repeated the BES silencing assay but we re-added tetracycline to the cells 8, 24 or 48 h post-silencing. Tetracycline relieves the transcriptional block of Tet repressor protein and thus allows cells that re-activate BES1 to survive (Figure 6A). Cells were cloned by limiting dilution and GFP-intensity of each clone was measured by FACS. GFP-positive clones indicate that cells re-activated BES1, while GFP-negative clones indicate cells that switched to a new BES. We observed that, none of the wells were GFP-negative at 8 h, suggesting that all surviving cells could potentially reactivate BES1. Instead, adding tetracycline at 24 h resulted in 49% of clones no longer expressing GFP, while at 48 h this number was 96%. These results show that most surviving cells are already committed to a new BES two days after silencing was induced. These results also show that during the first 8 h of switching, most cells are not committed to a VSG switch and can revert to transcribe the original BES.

We postulated that if commitment of most cells happens between 8 and 48 h, during this period we may be able to detect cells transiently probing new BESs at intermediate levels. Such cells were detected by Chaves *et al.* (9). To test if transcription of silent BESs increases before commitment, we constructed another reporter strain, GLB1-R15, in which an *RFP::NPT* fusion gene was inserted downstream of the silent BES15 in GLB1 cell-line (Figure 6B). This reporter allowed us to test at single-cell level whether

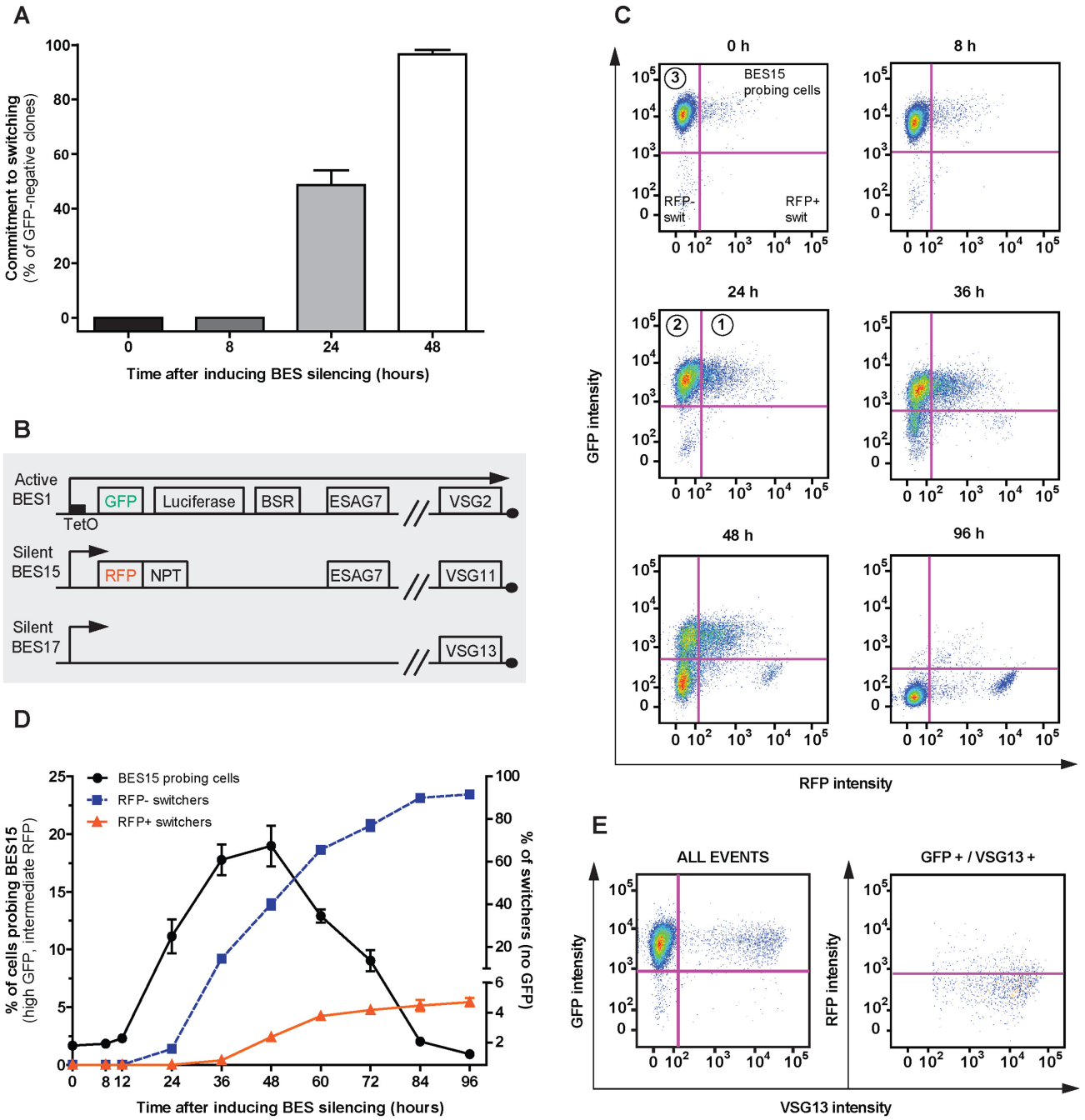


Figure 6. Cells transcriptionally probe silent BESs for up to two days, when most cells are committed to switching. (**A**) Commitment assay. 8, 24 or 48 h after inducing BES silencing, tetracycline was added back to the medium and cells were cloned. Six days later FACS was used to assess if clones were GFP-positive, indicating re-expression of original BES1. A minimum of 95 total individual wells was analyzed for each time-point between five individual experiments. (**B**) The cell-line GLB1-R15 is a derivative of GLB1, in which the fused gene *RFP::NPT* was introduced downstream the promoter of a silent BES. *NPT*, Neomycin Phosphotransferase. (**C**) Representative examples of FACS plots at several time-points post-BES silencing showing GFP and RFP expression (see Supplementary Figure S5 for complete set of time points). (**D**) Proportion of probing cells was assessed by measuring number of cells expressing RFP at intermediate levels and still present high levels of GFP (black line, left Y axis); switchers were defined as GFP-negative cells and either RFP- or RFP+ (blue and orange lines, respectively, right Y axis) after tetracycline removal. Three independent experiments were analyzed. Circled 1,2 and 3 labels indicate the sorted populations described in Figure 7. (**E**) **Left panel** - Representative example of a FACS plot at 24 h post-BES silencing showing VSG13 and GFP expression. **Right panel** - Representative example of a FACS plot showing VSG13 and RFP expression in the cells present on the top right gate in **left panel**. Four independent experiments were analyzed.

silent BES15 was being transiently more transcribed during switching (Figure 6C and D, Supplementary Figure S5). Before silencing was induced, a small number of cells (1–2%) expressed low levels of RFP (11–14-fold higher intensity than background levels), consistent with previous studies showing that silent BES are transcribed at low rate (40,41). 12 h after silencing was induced, we observed an increase in the proportion of total cells expressing RFP (2–3%), suggesting that more cells are transcribing BES15. The number of cells probing BES15 increased with time up until 36–48 h, in which around 20% of the cells showed elevated levels of RFP (Figure 6D).

At 36 h, two new and distinct GFP-negative populations were detected: one expressed high levels of RFP (0.7%) (∼200-fold higher intensity than RFP background levels) and the other was RFP-negative (14.5%), suggesting that these cells are switchers that silenced BES1 and activated BES15 (orange curve) or another BES (blue curve), respectively (Figure 6D). In this mixed population of cells, the switcher subpopulations became more predominant with time, while the number of cells probing silent BES15 gradually decreased. At 96 h, switchers were almost the sole populations in culture (around 5% expressed RFP and 92% did not).

Is probing restricted to the promoter region or does it span an entire BES? Because there is no antibody against the VSG11 of BES15, we used an anti-VSG13 antibody to test if we could detect VSG13 (from BES17) at the cell surface of silencing-induced cells. Silencing was induced in GLB1-R15 and, at 24 h, cells were stained with anti-VSG13 (BES17) and analyzed by FACS (Figure 6E). We observed that ∼5% of cells that expressed GFP also expressed heterogeneous levels of VSG13. Of these, ∼20% simultaneously expressed intermediate levels of RFP and VSG13, indicating that cells can simultaneously probe two BESs, BES15 (RFP) and BES17 (VSG13). These results also show that probing is not restricted to the promoter region and, at least in some cells, the entire BES is upregulated all the way until the telomeric end of BES17.

Taken together, we conclude that parasites transiently increase the transcription levels of silent BESs for around 2 days, when commitment to a new BES is almost complete; probing spans the entire BES, until the VSG gene and, albeit at a low frequency, two BESs can be simultaneously probed in individual cells.

BES probing is a reversible intermediate switching step

Next we investigated whether cells that partially upregulated RFP (probing BES15) are true switching intermediates. We chose to characterize the 24 h time-point because its FACS profile is less heterogeneous (Figure 6C and Supplementary Figure S5) and there are fewer dead cells (Supplementary Figure S2). We induced BES silencing for 24 h, FACS-sorted GFP+/RFP+ cells, and placed them in culture in limiting dilutions in the presence or absence of tetracycline (Population 1, indicated in 24 hr FACS plot Figure 6C). We also FACS-sorted and plated GFP+/RFP- cells (Population 2) because, although BES15 was not upregulated, other BESs may be upregulated. As a control, a population of cells in which silencing was not induced (Population 3) was also sorted (Figure 7A). Cells that attempt to reactivate BES1 die in the absence of tetracycline and survive in the presence of this drug. Therefore, the number of surviving clones six days after sorting reflects the efficiency of BES switching of each sorted population at 24 h post-silencing. Overall, in the absence of tetracycline, 3–5% of the clones survived the silencing assay, which is consistent with 8% of surviving clones measured in the clonogenic assay of the whole population (Figure 2D) and confirms that, with this inducible system, although we can dramatically increase the switching frequency from around 1 in a million cells to 1 in 10 cells, 9 in 10 cells still fail to switch to a new BES, even in the populations 1 and 2. When tetracycline was added to the culture medium after sorting, the number of surviving clones doubled (7–9%), which was expected because at 24 hr half of the population is still capable of returning to the originally active BES1 (Figure 6A).

Seven days post BES silencing, surviving clones were characterized by FACS for their expression of GFP and RFP, which we used as reporter of BES activity: GFP+/RFP- profile indicates clones that probably did not switch and reactivated BES1; GFP-/RFP+ indicates switchers to BES15; and GFP-/RFP- indicates switchers to another BES. No clones were obtained that simultaneously expressed GFP and RFP. For both populations 1 and 2, all post-sorting clones obtained in the absence of tetracycline no longer expressed GFP and thus represent switchers. Interestingly, from population 1, 57% (12 clones on average of three experiments) of switchers expressed RFP, while 43% (nine clones) did not, indicating that cells can probe one BES at 24 h post-silencing and eventually switch to another BES (Figure 7B). From population 2, all switcher clones were RFP-negative (13 clones), indicating that switchers activated a BES other than BES15 (Figure 7B). These results suggest that if a BES is not probed at 24 h, apparently it is not fully activated later.

When sorted populations were plated in the presence of tetracycline we obtained some clones that expressed GFP and not RFP, as expected and consistent with the fact that at 24 h there are still many cells that are not committed for switching and reactivate BES1 (Figure 7A). Post-sorting clones from population 1 showed similar proportions of the three different types of GFP/RFP expression profiles, indicating a large plasticity of this population before commitment (Figure 7B). In contrast, most post-sorting clones from population 2 reactivated BES1 (77%, 24 clones) and none activated BES15 (Figure 7B), suggesting once again that if a BES is not probed at 24 h, it seems not to be activated later.

Clones obtained after sorting population 3, in which silencing was never induced, resulted as expected in 100% of clones expressing GFP and not RFP (Figure 7B), consistent with BES1 remaining active. This shows that the stress associated with sorting did not induce any unexpected changes in BES expression.

Taken together, our results indicate that cells in a probing state are not a dead end product resulting from the silencing inducible system used in this work. Importantly, these data show that probing cells can revert to the original BES1, or switch and commit to the probed BES, or switch and commit to another BES. Probing is therefore a reversible step

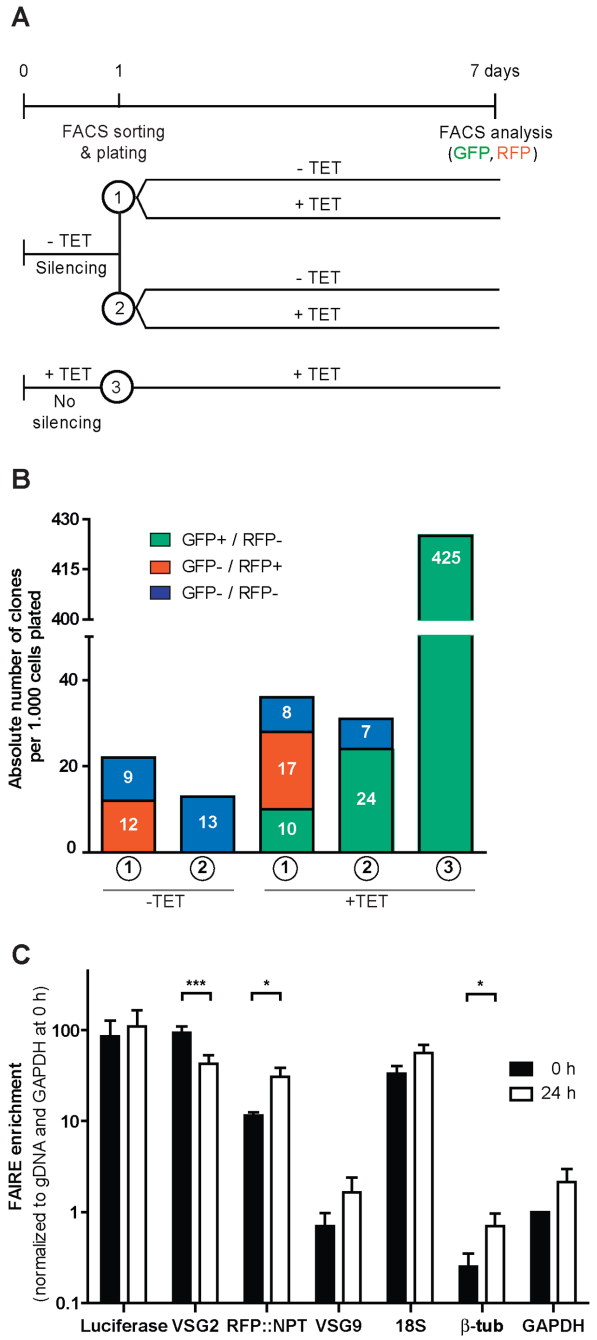


Figure 7. Probing cells are switching intermediates that can choose different fates. (**A**) Experimental design of FACS sorting and subsequent phenotype characterization. Tetracycline was removed from the medium of GLB1-R15 cell-line. One day later, two populations 1 and 2 (indicated in Figure 6C) were sorted and plated by limiting dilution. Population 1 consists of cells that express GFP and RFP, while population 2 consists of cells that express GFP, but not RFP. As control, a third population of GFP+/RFP- cells was sorted from a culture kept under tetracycline pressure. Clones obtained after six days were characterized in terms of expression of GFP and RFP by FACS. (**B**) GFP and RFP expression of surviving clones was assessed by FACS. Numbers in white show mean number of clones with each specific phenotype from three individual experiments. **(C)** Chromatin conformation of GLB1-R15 cells after 24 h of BES silencing was measured by FAIRE. DNA was quantified by qPCR and normalized to gDNA copy number and to *GAPDH* at 0 h. Statistical significance was determined by a paired *t*-test against time-point 0 h. $*P < 0.05$; $***P < 0.001$.

during BES switching, allowing cells to reactivate the originally active BES or activate previously silent BESs.

Given that at 24 h, the population consists mainly of cells capable of switching or reverting to original BES (with only around 3% of dead cells, Supplementary Figure S2), we decided to check whether chromatin of the originally active BES1 remains accessible during this period in this cell-line (GLB1-R15). For this, silencing was induced and cells were collected for FAIRE at 0 and 24 h post-silencing (Figure 6E). At 24 h, chromatin in *VSG2* became significantly more compact than at 0 h, which is consistent with the trend previously observed by FAIRE in GLB1 cell-line at 8 h (Figure 3C), but the conformation is still 25-fold more open than that of a silent *VSG9*. Consistent with probing phenotype, we detected a significant increase in chromatin accessibility of the *RFP::NPT* gene. All other tested genes showed a slight increase in accessibility, which may be due to the fact that, at 24 h, the culture has some cells arrested in G_2/M or with an abnormal DNA content (Figure 2C), or already dead (Supplementary Figure S2).

Taken together, our data reveal a novel association between the alterations of chromatin structure at the active BES and BES switching. Chromatin is kept essentially open for at least 24 h, while silent BESs are reversibly probed before committing to a new BES.

DISCUSSION

In this study, we showed that the chromatin structure of the active BES is cell-cycle independent and TDP1 keeps its open conformation when transcription is halted, especially in the telomeric region. We propose that these properties are critical for the dynamics of switching between BESs because it gives time for cells to reversibly probe multiple silent BESs before committing to a new BES.

Transcription and chromatin dynamics in the active BES

The initial trigger that leads to VSG switching remains a mystery. However, several studies in which transcription of the active BES was somewhat interrupted or diminished, either due to loss of the VSG upstream sequence (CTR) (10) or replacement of a BES promoter by the T7 promoter (24), resulted in more frequent *in situ* switching, suggesting one of the earliest events during BES switching is the silencing of the active BES. In this work, we first checked if halting transcription is also one of the earliest steps during differentiation from bloodstream to procyclic forms, a process in which the active BES needs to be silenced while procyclin genes are upregulated. We show that, 24 h after inducing differentiation, transcription of the active BES is highly reduced to 2% (Figure 1B), while the chromatin of this locus remains essentially open (Figure 1C). We conclude that during differentiation, transcription silencing precedes chromatin changes, and thus it is likely that a similar mechanism happens during BES switching in bloodstream forms.

In the tetracycline-inducible system developed by the Horn lab, when the tetracycline repressor binds the tetracycline operator, it sterically blocks Pol I transcription, which results in silencing of the initially active BES and activation of a new one (Figure 2). In this work, we confirmed

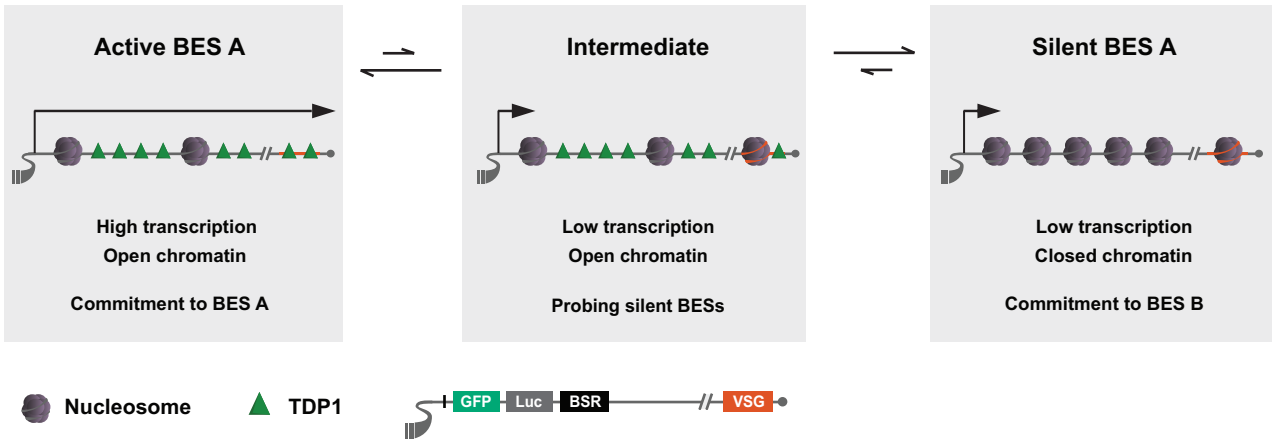


Figure 8. Model for VSG expression site switching: transcriptional probing silent BESs before commitment is associated to a temporary maintenance of open chromatin by TDP1. Chromatin of the actively transcribed BES possesses an open chromatin enriched in TDP1. Upon transcriptional silencing of the active BES, cells undergo an intermediate state characterized by stabilization of the active BES open chromatin by TDP1 while probing silent BESs. Commitment to switch or not switch seems to take around two days. This decision can either be returning to the initial active BES (which, in our reporter cell-line, because BES1 is blocked by tetracycline repressor, it would result in cell death) or switching to a new BES. In this later case, the chromatin of the originally active BES loses TDP1 and becomes nucleosome-enriched.

that inducible BES silencing promotes a rapid decrease in transcript levels of genes from initially active BES (at least 90% drop in the first 8 h, as accessed by transcript levels of the unstable *luciferase* reporter) and an eviction of Pol I from chromatin fiber, which confirms that transcription of active BES is efficiently and rapidly stopped (Figure 3A, B and Supplementary Figure S3). During this period, chromatin remained in an open state despite transcription being halted (Figure 3C and D). By ChIP, we observed a slight enrichment of histone H3 in the *VSG2* gene and a decrease in TDP1, suggesting a more compact chromatin mainly at the telomeric end of BES.

Our study also shows that chromatin conformation of the active BES is cell-cycle independent. Using FAIRE, we detected no significant differences in chromatin conformation of active BES in cells in G_1, S and G_2/M. Given that, when parasites divide, most daughter cells use the 'mother's' BES and only very few switch to a new BES, it is tempting to speculate that keeping an open chromatin structure serves as an epigenetic marker that, after cell division, signals which BES should be used by the daughter cells. Perhaps TDP1 interacts with the cohesion complex, which is necessary after S phase to maintain the two sister chromatids of the active BES associated to the single Expression Site Body while waiting for chromosome segregation during mitosis (42). This hypothesis deserves further investigation in the future.

TDP1 maintains open chromatin when BES has been silenced

High-mobility group box (HMGB) proteins are essential nuclear components in chromatin structure, transcriptional activity and DNA damage repair (43–45). In yeast, HMO1 is a HMGB protein that is associated with Pol I transcription machinery (46), which can competitively displace histone H1 (47) and also maintain chromatin in an open conformation (3). In *T. brucei*, TDP1 is an essential HMGB protein that is highly enriched in Pol I loci (18,19). Like yeast HMO1, TDP1 is a facilitator of Pol I transcription and it is necessary to keep chromatin nucleosome-depleted in a steady-state situation. Here, we show that when BES silencing was induced in the absence of TDP1, chromatin became significantly more compact, indicating that TDP1 is necessary to maintain an open chromatin structure in the absence of transcription. This role is consistent with what has been observed in yeast, in which HMO1 maintains Pol I chromatin open when transcription is halted (3).

In eukaryotes, deposition of canonical histones is normally replication-dependent (48). Although our data shows that maintenance of BES chromatin structure is cell-cycle independent (Figure 4), we predict that during the rare event of BES switching, chromatin remodelling occurs in S phase of the cell-cycle. Thus, during replication, eviction of TDP1 from a silencing BES may be an important step to complete a BES switching. Regulating the timing of TDP1 eviction may be a means to determine for how long a cell stays in a 'probing' stage, before commitment.

It is interesting to observe that, in two different circumstances (BES transcriptional silencing in wild-type or TDP1 depleted conditions), the chromatin of the BES telomeric end is more rapidly closed than chromatin close to the BES promoter (which can be up to 50 kb upstream from telomere). Telomeres possess a highly controlled chromatin structure and, in trypanosomes, they are very important for antigenic variation (49). It is possible that, upon transcriptional silencing, loading of nucleosomes onto chromatin happens faster at the telomeric region. Tiengwe *et al.* found that although most origins of replication are found in core regions of the chromosome, VSG genes also harbor origins of replication, which could drive inward replication and deposition of nucleosomes (50). Spreading of such chromatin condensation may happen evenly towards the BES promoter or it may happen in steps if the BES has insulator elements, such as the 70-bp repeat array for example, which could prevent spreading of nucleosome loading into the upstream ESAGs.

Probing and commitment to a new BES

Chaves *et al.* proposed that, during a switching event, a cell pre-activates a silent BES before a natural intermediate rapidly and transiently express two VSGs and ultimately commits for switching or reverts to the original BES (9). A more recent study, in which BES transcriptional attenuation was observed as a consequence of ectopically overexpressing a silent VSG, also proposed that this attenuation could give time for parasites to probe silent BES before switching to a fully competent BES (11). The nature of what cells are probing is unknown. It is possible that cells are testing the order of VSG switching, or if ESAGs are functional, or if previously untested or mosaic VSGs are functional.

Our results are in agreement with these models of 'probing before decision' and we propose that regulation of chromatin structure may be the underlying epigenetic mechanism. We show here that trypanosomes test silent BES before a choice is made and this is associated to keeping chromatin of the active BES open (Figure 8). Our commitment assay showed that, 8 hr after silencing, no cells are committed to switching although silent VSG transcripts are already detected (Figure 3B). Half of the cells are committed 24 h post-silencing, while most of them are committed at 48 h. From 12 to 48 h, more cells probe silent BESs, but only a fraction of the cells expresses higher levels of RFP, which indicates that not all BESs are being probed at the same time in each cell. Nevertheless, we found that at least two BESs can be probed simultaneously by the same cell (Figure 6E). At 24 h, a small fraction of cells becomes GFP-negative, consistent with committed switching and this population becomes more predominant with time. These results show that probing of one or more BESs is a reversible process that lasts around two days, during which cells either switch to the probed BES, switch to a different BES or revert to BES1. Reverting to original BES1 is probably facilitated by the fact that BES1 retains an open chromatin structure for at least 24 h (Figure 7C).

In some aspects, our model resembles the formation of poised chromatin. Poised chromatin at Pol II transcribed genes is characterized by the presence of bivalent marks: the co-existence of activating and repressive histone marks within the same domains (51). This epigenetic status has been associated to genes that are transcribed at low levels but need to be in a prepared state for developmental fates of rapid activation or repression, such as differentiation of embryonic stem cells or activation of T-cells (52,53). A poised chromatin state has also been found at *var* genes in *Plasmodium falciparum*, the causative agent of malaria. Switching between *var* genes is essential for antigenic variation (54). It is possible that *T. brucei* uses a TDP1-dependent mechanism as a means of temporarily keeping the active BES in a poised state, ready for being activated or repressed. In the future, it will be interesting to test whether this mechanism of keeping chromatin temporarily open also facilitates VSG switching by recombination. In lymphocytes, for example, it has been already shown that the chromatin that is accessible for V(D)J recombination typically displays elevated acetylation of histones H3 and H4, which is a hallmark of a more accessible chromatin (55).

In this study, we showed that the chromatin structure is kept open by TDP1 probably to facilitate probing, before cells commit to a new BES. It is intriguing that, during differentiation, chromatin of the originally active BES is also hold open for some time while transcription has already been reduced. This suggests that differentiating cells may also undergo an intermediate stage, which may be reversible. This is consistent with findings by Batram *et al.*, in which attenuation of the active BES results in an intermediate stage that can reversibly progress in two directions: differentiation to procyclic or returning to proliferation in the mammalian form. The implications of chromatin dynamics in differentiation should be further investigated, and preferably in a pleomorphic strain, whose differentiation process is more similar to what happens *in vivo*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Christian Janzen and David Horn for the kind gifts of histone H3 antibody and 2T1.ESPiGFP:NPT cell-line, respectively. The authors would also like to thank Ruy M. Ribeiro and Sandra Trindade for help with statistical analysis, Daniel Neves for drawing the model in Figure 8, Arthur Gunzl for the RPA31 antibody and sharing protocols and George A.M. Cross for critically reading the manuscript.
Author contributions: SP helped in the construction of GLB1. FAB and LMF designed the experiments and wrote the manuscript. FAB performed the experiments.

FUNDING

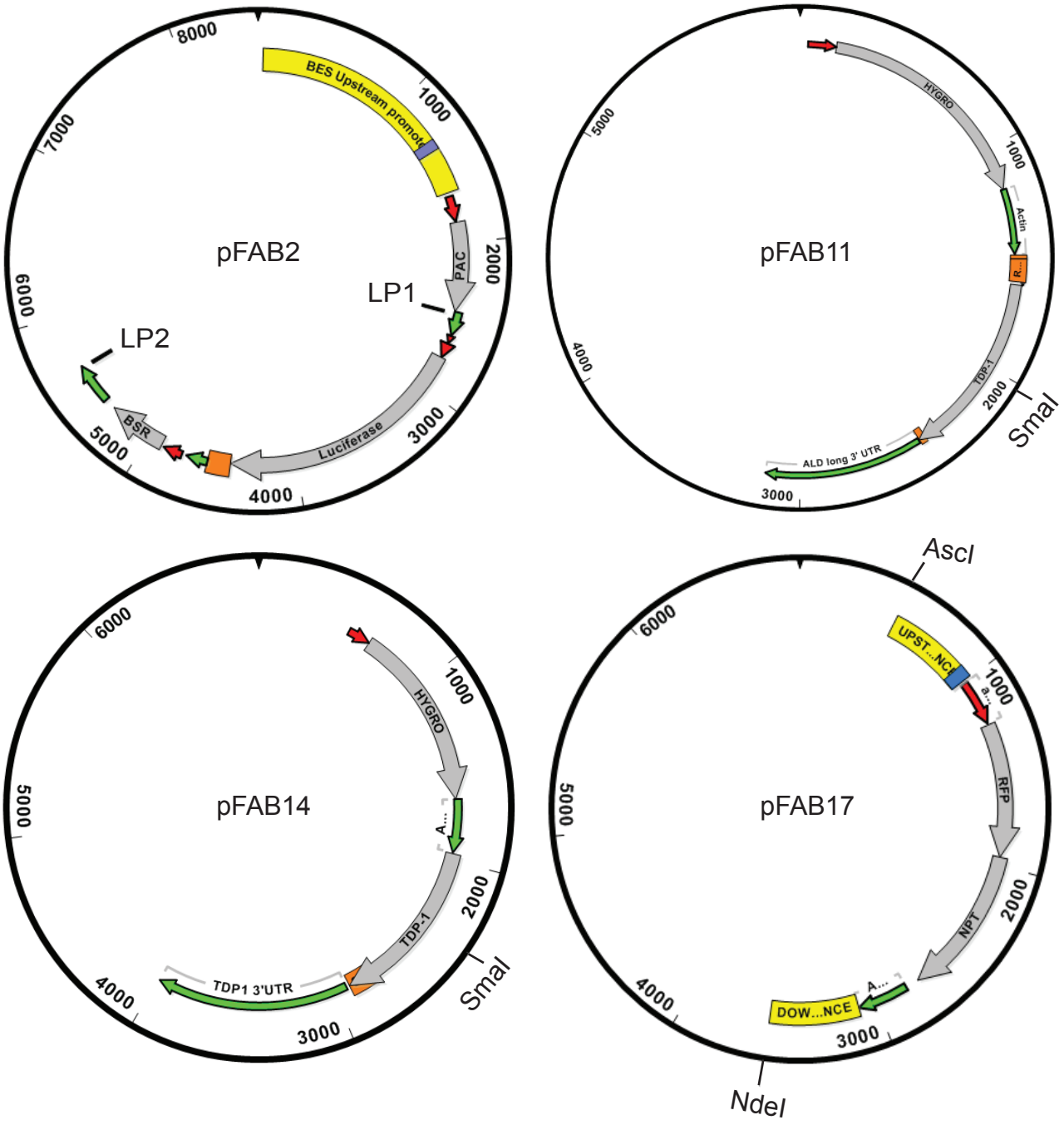
Fundação para a Ciência e Tecnologia [PTDC/SAU-MIC/113225/2009] (in part); European Molecular Biology Organization Installation grant [Project 2151]; Howard Hughes Medical Institute International Early Career Scientist Program [55007419]; Fundação para a Ciência e Tecnologia (to L.M.F.); Fellowship [SFRH/BD/80718/2011 to F.A.B.].
Conflict of interest statement. None declared.

REFERENCES

1. Warner,J.R. (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.*, **24**, 437–440.
2. McStay,B. and Grummt,I. (2008) The epigenetics of rRNA genes: from molecular to chromosome biology. *Annu. Rev. Cell Dev. Biol.*, **24**, 131–157.
3. Wittner,M., Hamperl,S., Stockl,U., Seufert,W., Tschochner,H., Milkereit,P. and Griesenbeck,J. (2011) Establishment and maintenance of alternative chromatin states at a multicopy gene locus. *Cell*, **145**, 543–554.
4. Figueiredo,L.M., Cross,G.A. and Janzen,C.J. (2009) Epigenetic regulation in African trypanosomes: a new kid on the block. *Nat. Rev. Microbiol.*, **7**, 504–513.
5. Gunzl,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.

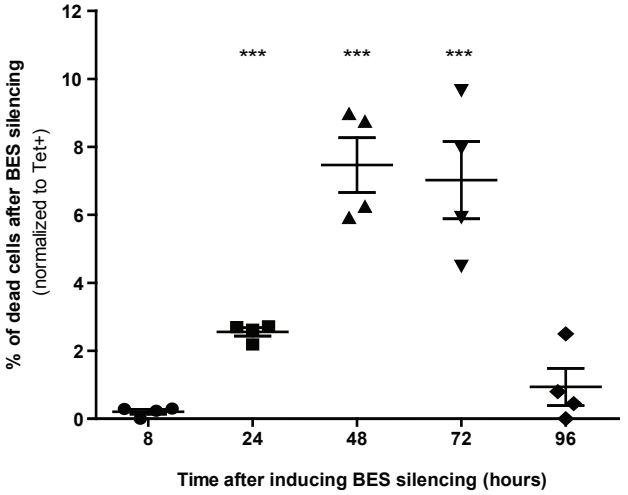
6. Cross,G.A., Kim,H.S. and Wickstead,B. (2014) Capturing the variant surface glycoprotein repertoire (the VSGnome) of *Trypanosoma brucei* Lister 427. *Mol. Biochem. Parasitol.*, **195**, 59–73.
7. Hertz-Fowler,C., Figueiredo,L.M., Quail,M.A., Becker,M., Jackson,A., Bason,N., Brooks,K., Churcher,C., Fahkro,S., Goodhead,I. *et al.* (2008) Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS One*, **3**, e3527.
8. Horn,D. and McCulloch,R. (2010) Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. *Curr. Opin. Microbiol.*, **13**, 700–705.
9. Chaves,I., Rudenko,G., Dirks-Mulder,A., Cross,M. and Borst,P. (1999) Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *EMBO J.*, **18**, 4846–4855.
10. Davies,K.P., Carruthers,V.B. and Cross,G.A. (1997) Manipulation of the vsg co-transposed region increases expression-site switching in *Trypanosoma brucei*. *Mol .Biochem. Parasitol.*, **86**, 163–177.
11. 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 (Cambridge)*, **3**, e02324.
12. 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.
13. Stanne,T.M. and Rudenko,G. (2010) Active VSG expression sites in *Trypanosoma brucei* are depleted of nucleosomes. *Eukaryot. Cell*, **9**, 136–147.
14. Brandenburg,J., Schimanski,B., Nogoceke,E., Nguyen,T.N., Padovan,J.C., Chait,B.T., Cross,G.A. and Gunzl,A. (2007) Multifunctional class I transcription in *Trypanosoma brucei* depends on a novel protein complex. *EMBO J.*, **26**, 4856–4866.
15. Nguyen,T.N., Muller,L.S., Park,S.H., Siegel,T.N. and Gunzl,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 Res.*, **42**, 3164–3176.
16. Gunzl,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.
17. Figueiredo,L.M., Janzen,C.J. and Cross,G.A. (2008) A histone methyltransferase modulates antigenic variation in African trypanosomes. *PLoS Biol.*, **6**, e161.
18. 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.
19. Erondu,N.E. and Donelson,J.E. (1992) Differential Expression of 2 Messenger-Rnas from a Single Gene Encoding an Hmg1-Like DNA-Binding Protein of African Trypanosomes. *Mol. Biochem. Parasitol.*, **51**, 111–118.
20. Fenn,K. and Matthews,K.R. (2007) The cell biology of *Trypanosoma brucei* differentiation. *Curr. Opin. Microbiol.*, **10**, 539–546.
21. Ziegelbauer,K., Quinten,M., Schwarz,H., Pearson,T.W. and Overath,P. (1990) Synchronous differentiation of *Trypanosoma brucei* from bloodstream to procyclic forms *in vitro*. *Eur. J. Biochem.*, **192**, 373–378.
22. Amiguet-Vercher,A., Perez-Morga,D., Pays,A., Poelvoorde,P., Xong,H., Tebabi,P., Vanhamme,L. and Pays,E. (2004) Loss of the mono-allelic control of the VSG expression sites during the development of *Trypanosoma brucei* in the bloodstream. *Mol. Microbiol.*, **51**, 1577–1588.
23. Landeira,D. and Navarro,M. (2007) Nuclear repositioning of the VSG promoter during developmental silencing in *Trypanosoma brucei*. *J. Cell Biol.*, **176**, 133–139.
24. Navarro,M., Cross,G.A. and Wirtz,E. (1999) *Trypanosoma brucei* variant surface glycoprotein regulation involves coupled activation/inactivation and chromatin remodeling of expression sites. *EMBO J.*, **18**, 2265–2272.
25. Cross,G.A. (1975) Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of *Trypanosoma brucei*. *Parasitology*, **71**, 393–417.
26. 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. *J. Parasitol.*, **75**, 985–989.
27. 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.
28. Schumann Burkard,G., Jutzi,P. and Roditi,I. (2011) Genome-wide RNAi screens in bloodstream form trypanosomes identify drug transporters. *Mol. Biochem. Parasitol.*, **175**, 91–94.
29. 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.
30. 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.
31. 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*. *Mol. Microbiol.*, **93**, 645–663.
32. Bastin,P., Bagherzadeh,Z., Matthews,K.R. and Gull,K. (1996) A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, **77**, 235–239.
33. Park,S.H., Nguyen,T.N., Kirkham,J.K., Lee,J.H. and Gunzl,A. (2011) Transcription by the multifunctional RNA polymerase I in *Trypanosoma brucei* functions independently of RPB7. *Mol. Biochem. Parasitol.*, **180**, 35–42.
34. Birkett,C.R., Foster,K.E., Johnson,L. and Gull,K. (1985) Use of monoclonal antibodies to analyse the expression of a multi-tubulin family. *FEBS Lett.*, **187**, 211–218.
35. Ehlers,B., Czichos,J. and Overath,P. (1987) RNA turnover in *Trypanosoma brucei*. *Mol. Cell. Biol.*, **7**, 1242–1249.
36. Queiroz,R., Benz,C., Fellenberg,K., Hoheisel,J.D. and Clayton,C. (2009) Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. *BMC Genomics*, **10**, 495.
37. Giresi,P.G., Kim,J., McDaniell,R.M., Iyer,V.R. and Lieb,J.D. (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res.*, **17**, 877–885.
38. Worthen,C., Jensen,B.C. and Parsons,M. (2010) Diverse effects on mitochondrial and nuclear functions elicited by drugs and genetic knockdowns in bloodstream stage *Trypanosoma brucei*. *PLoS Negl. Trop. Dis.*, **4**, e678.
39. Patrick,K.L., Shi,H., Kolev,N.G., Ersfeld,K., Tschudi,C. and Ullu,E. (2009) Distinct and overlapping roles for two Dicer-like proteins in the RNA interference pathways of the ancient eukaryote *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 17933–17938.
40. Vanhamme,L., Poelvoorde,P., Pays,A., Tebabi,P., Van Xong,H. and Pays,E. (2000) Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. *Mol. Microbiol.*, **36**, 328–340.
41. Kassem,A., Pays,E. and Vanhamme,L. (2014) Transcription is initiated on silent variant surface glycoprotein expression sites despite monoallelic expression in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 8943–8948.
42. Landeira,D., Bart,J.M., Van Tyne,D. and Navarro,M. (2009) Cohesin regulates VSG monoallelic expression in trypanosomes. *J. Cell Biol.*, **186**, 243–254.
43. Reeves,R. (2010) Nuclear functions of the HMG proteins. *Biochim. Biophys. Acta*, **1799**, 3–14.
44. Ueda,T. and Yoshida,M. (2010) HMGB proteins and transcriptional regulation. *Biochim. Biophys. Acta*, **1799**, 114–118.
45. Gonzalez-Huici,V., Szakal,B., Urulangodi,M., Psakhye,I., Castellucci,F., Menolfi,D., Rajakumara,E., Fumasoni,M., Bermejo,R., Jentsch,S. *et al.* (2014) DNA bending facilitates the error-free DNA damage tolerance pathway and upholds genome integrity. *EMBO J.*, **33**, 327–340.
46. Gadal,O., Labarre,S., Boschiero,C. and Thuriaux,P. (2002) Hmo1, an HMG-box protein, belongs to the yeast ribosomal DNA transcription system. *EMBO J.*, **21**, 5498–5507.
47. Stros,M. (2010) HMGB proteins: interactions with DNA and chromatin. *Biochim. Biophys. Acta*, **1799**, 101–113.

48. Groth,A., Rocha,W., Verreault,A. and Almouzni,G. (2007) Chromatin challenges during DNA replication and repair. *Cell*, **128**, 721–733.
49. Li,B. (2010) A newly discovered role of telomeres in an ancient organism. *Nucleus*, **1**, 260–263.
50. 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.
51. Bernstein,B.E., Mikkelsen,T.S., Xie,X., Kamal,M., Huebert,D.J., Cuff,J., Fry,B., Meissner,A., Wernig,M., Plath,K. *et al.* (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, **125**, 315–326.
52. Puri,D., Gala,H., Mishra,R. and Dhawan,J. (2014) High-wire act: the poised genome and cellular memory. *FEBS J.*, **282**, 1675–1691.
53. Lim,P.S., Li,J., Holloway,A.F. and Rao,S. (2013) Epigenetic regulation of inducible gene expression in the immune system. *Immunology*, **139**, 285–293.
54. Volz,J.C., Bartfai,R., Petter,M., Langer,C., Josling,G.A., Tsuboi,T., Schwach,F., Baum,J., Rayner,J.C., Stunnenberg,H.G. *et al.* (2012) PfSET10, a *Plasmodium falciparum* methyltransferase, maintains the active *var* gene in a poised state during parasite division. *Cell Host Microbe*, **11**, 7–18.
55. Krangel,M.S. (2003) Gene segment selection in V(D)J recombination: accessibility and beyond. *Nat. Immunol.*, **4**, 624–630.
56. Kabani,S., Fenn,K., Ross,A., Ivens,A., Smith,T.K., Ghazal,P. and Matthews,K. (2009) Genome-wide expression profiling of *in vivo*-derived bloodstream parasite stages and dynamic analysis of mRNA alterations during synchronous differentiation in *Trypanosoma brucei*. *BMC Genomics*, **10**, 427.



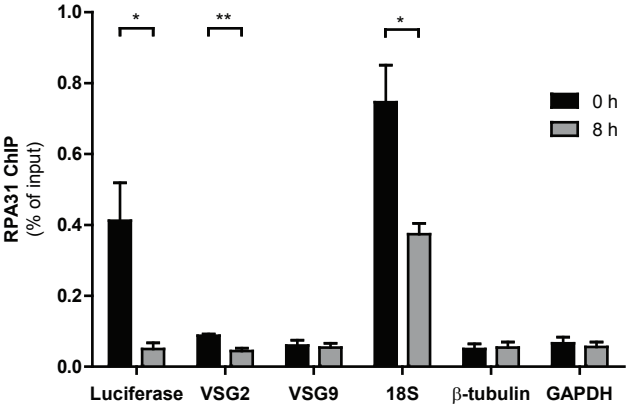
Supplementary Figure S1. Maps of constructs used in this work.

Constructs are described in Materials and Methods. Gene ORFs are represented in gray, 5'UTRs in red, 3'UTRs in green, upstream and downstream BES promoter sequences in yellow and BES promoter in blue. LP1 and LP2 represent the primers with long tails designed to generate GLB1 with the sequence listed in Supplementary Table 1. Restriction sites used prior to transfection are indicated.



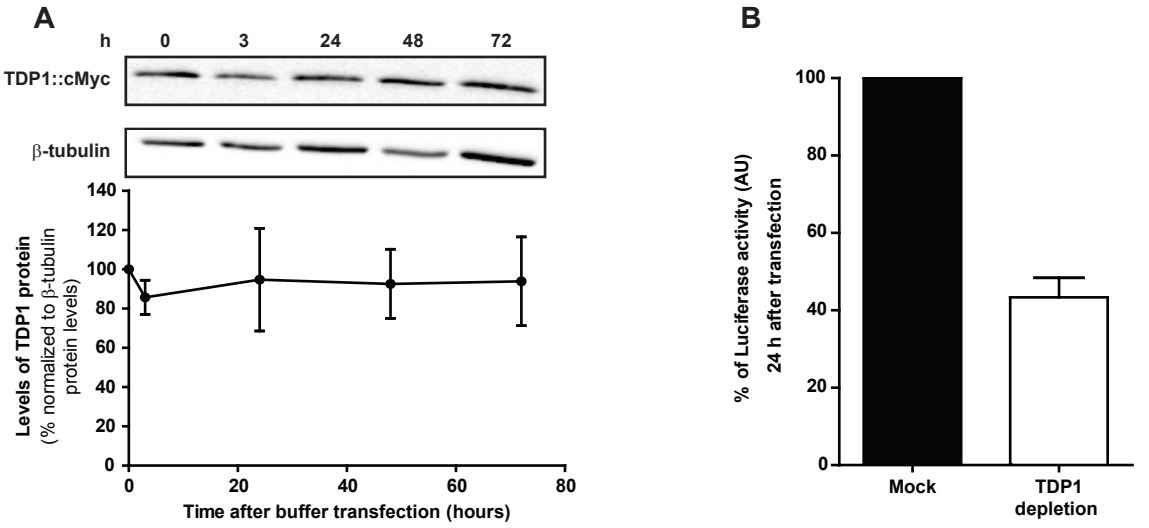
Supplementary Figure S2. Cell death during BES silencing assay.

BES silencing was induced and cells were collected 8, 24, 48, 72 and 96 h later. Induced cells were incubated with 5 μg/ml of propidium iodide (PI) prior to FACS analysis. Dead cells or in late apoptosis are stained by propidium iodide. The number of PI positive cells in 'Tet-' culture was normalized to the number of PI positive cells in 'Tet+' culture. Four independent experiments were analyzed. Statistical significance was determined by a t-test against a hypothetical mean value of 0, corresponding to no cell death. ***: $p < 0.001$.



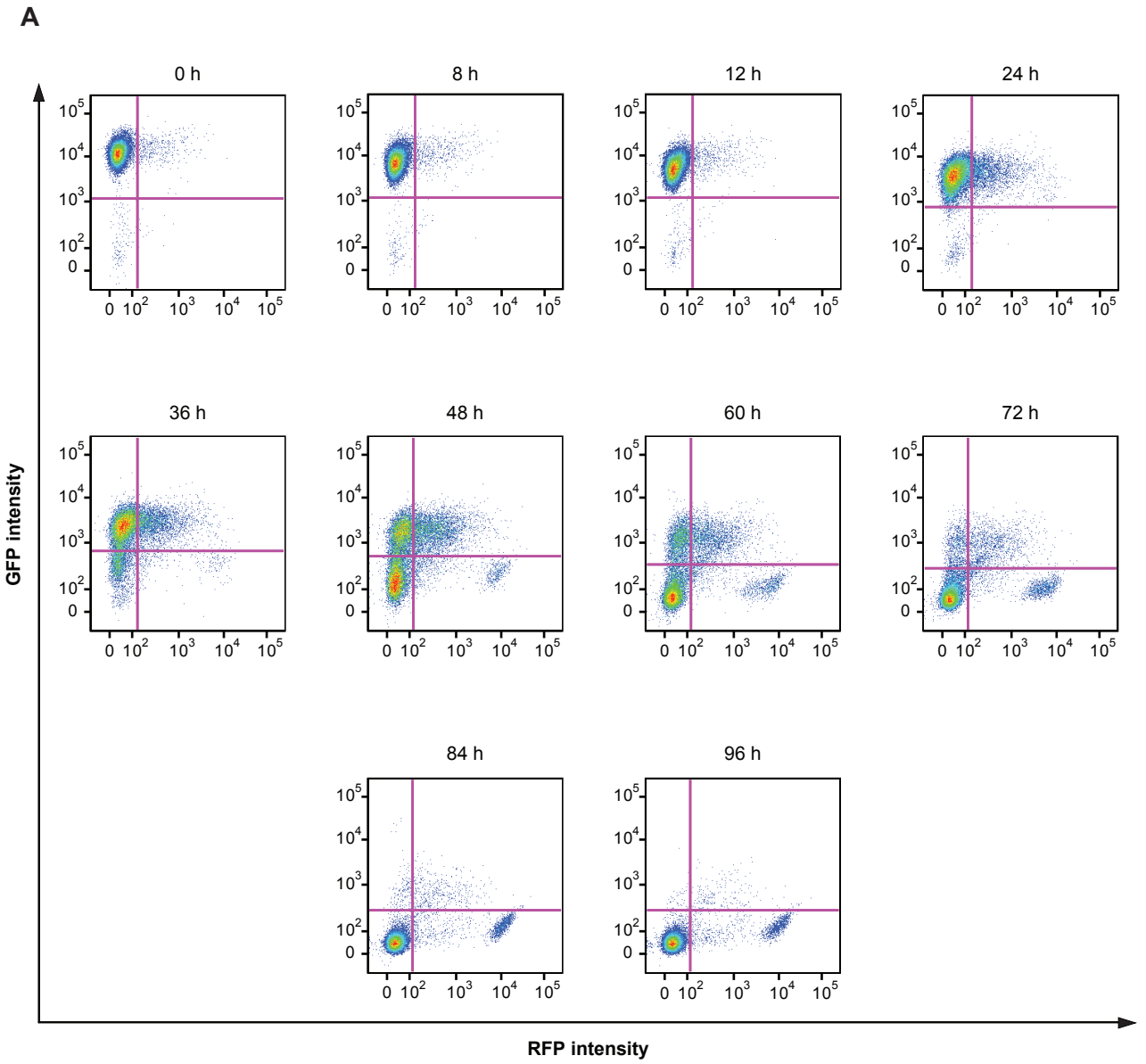
Supplementary Figure S3. RNA Pol I is absent from active BES 8 h after inducing BES silencing.

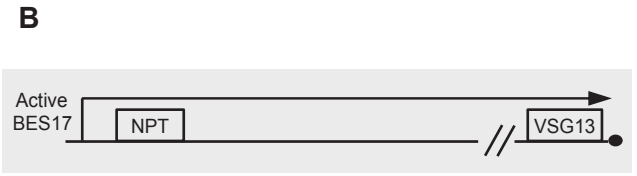
RNA Pol I occupancy was determined by RPA31 subunit ChIP at 0 and 8 h after tetracycline removal. Immunoprecipitated DNA was compared to the total input material. Five independent experiments were analyzed. Statistical significance was determined by a paired t-test against time-point 0 h. *:$p < 0.05$; **: $p < 0.01$.

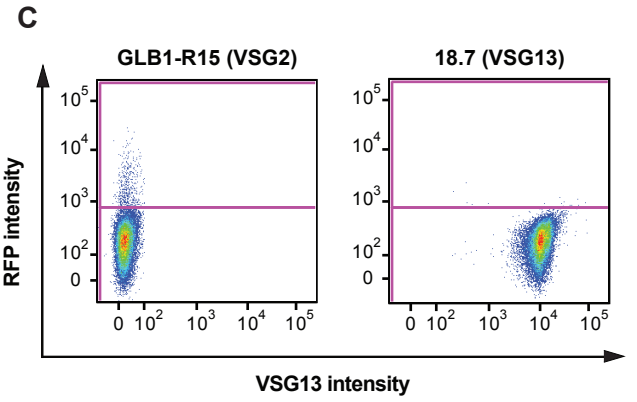


Supplementary Figure S4. TDP1 facilitates BES transcription.

(A) Western blotting analysis of TDP1 protein after 3, 24, 48 and 72 h of transfection with buffer (mock) in GLB1-TDP1::3xcMyc. Time-point 0 h indicates cells which were not transfected. Each lane corresponds to lysates from 2×10^6 cells. Quantification of TDP1 signal is indicated in the lower panel. TDP1 protein levels were normalized for β-tubulin protein levels and to not transfected cells. Four independent experiments were analyzed. **(B)** Luciferase activity was measured 24 h after TDP1 depletion and compared to the mock control in GLB1-TDP1::3xcMyc. Five independent experiments were analyzed.







Supplementary Figure S5. GFP and RFP gene expression during BES silencing.

(A) Representative density plots of GFP and RFP intensities of GLB1-R15 after induction of BES silencing. Cells were analyzed by FACS from 0 to 96 h. X-axis represents RFP intensity while Y-axis represents GFP intensity. As indicated in **Figure 6B**, *GFP* is present in originally active BES1 and *RFP* is present in the originally silent BES15. **(B)** The cell-line 18.7, which expresses VSG13. *NPT*, Neomycin Phosphotransferase. **(C)** Representative density plots of RFP and VSG13 intensities of GLB1-R15 (expresses VSG2, **left panel**) and 18.7 (expresses VSG13, **right panel**). X-axis represents VSG13 intensity while Y-axis represents RFP intensity. Top gate represents RFP positive cells and bottom gate represents RFP negative cells. This gating was applied in **Figure 6E**, **right panel**.

Molecular Microbiology (2014) 93(4), 645–663 ■
doi:10.1111/mmi.12677
First published online 14 July 2014

Trypanosoma brucei histone H1 inhibits RNA polymerase I transcription and is important for parasite fitness *in vivo*

Ana C. Pena,[1] Mafalda R. Pimentel,[1] Helena Manso,[1] Rita Vaz-Drago,[1] Daniel Pinto-Neves,[1] Francisco Aresta-Branco,[1] Filipa Rijo-Ferreira,[1] Fabien Guegan,[1] Luis Pedro Coelho,[1†] Maria Carmo-Fonseca,[1] Nuno L. Barbosa-Morais[1,2] and Luisa M. Figueiredo[1*]

[1]*Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Av. Prof. Egas Moniz, Edifício Egas Moniz, 1649-028 Lisboa, Portugal.*
[2]*Nuffield Department of Obstetrics and Gynaecology, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK.*

Summary

***Trypanosoma brucei* is a unicellular parasite that causes sleeping sickness in humans. Most of its transcription is constitutive and driven by RNA polymerase II. RNA polymerase I (Pol I) transcribes not only ribosomal RNA genes, but also protein-encoding genes, including variant surface glycoproteins (*VSGs*) and procyclins. In *T. brucei*, histone H1 (H1) is required for *VSG* silencing and chromatin condensation. However, whether H1 has a genome-wide role in transcription is unknown. Here, using RNA sequencing we show that H1 depletion changes the expression of a specific cohort of genes. Interestingly, the predominant effect is partial loss of silencing of Pol I loci, such as *VSG* and procyclin genes. Labelling of nascent transcripts with 4-thiouridine showed that H1 depletion does not alter the level of labelled Pol II transcripts. In contrast, the levels of 4sU-labelled Pol I transcripts were increased by two- to sixfold, suggesting that H1 preferentially blocks transcription at Pol I loci. Finally, we observed that parasites depleted of H1 grow almost normally in culture but they have a reduced fitness in mice, suggesting that H1 is important for host–pathogen interactions.**

Accepted 15 June, 2014. *For correspondence. E-mail lmf@fm.ul.pt; Tel. (+351) 217 999 512; Fax (+351) 217 999 504. †Present address: European Molecular Biology Laboratory, Meyerhofstraße 1, 69117 Heidelberg, Germany.

Introduction

Trypanosoma brucei is a protozoan parasite responsible for sleeping sickness in humans (African Trypanosomiasis). Gene expression in *T. brucei* has several peculiarities, namely genes being transcribed polycistronically and in a constitutive manner. As a result, most of the genome is believed to be continuously transcribed and gene regulation to happen mainly post-transcriptionally (Clayton and Shapira, 2007). Another unusual feature of transcription in trypanosomes is that RNA Polymerase I (Pol I) is not solely dedicated to the transcription of ribosomal RNA (rRNA) genes, but also transcribes loci that encode for abundant surface proteins, such as variant surface glycoprotein (VSGs) and procyclins. Interestingly, so far, Pol I is the only RNA polymerase in *T. brucei* for which regulation at the transcription level was demonstrated (Rudenko *et al.*, 1989; Gunzl *et al.*, 2003).

VSGs are expressed when parasites reside in the bloodstream of the mammalian host. They are required for the parasite to escape the mammalian immune system by a mechanism known as antigenic variation. *VSGs* are expressed from a specialized sub-telomeric locus called a bloodstream expression site (BES), which also harbours other expression-site-associated genes (*ESAGs*). In the genome there are ~ 15 BESs (Hertz-Fowler *et al.*, 2008), only one of which is transcriptionally active at a time. Outside BESs, *T. brucei* also encodes a silent archive of up to around 2000 *VSG* genes and pseudogenes (Marcello and Barry, 2007). Procyclins are a much smaller gene family, composed of three *EP* genes and one *GPEET* gene (reviewed in Pays *et al.*, 2004). They are located in non-telomeric polycistronic units, which also contain procyclin-associated genes (*PAGs*). In bloodstream forms (BSFs), procyclins are kept silent by transcriptional and post-transcriptional mechanisms. These genes are activated in a coordinated manner when parasites differentiate in the gut of the tsetse transmitting fly (reviewed in Navarro *et al.*, 2007).

Chromatin plays an essential role in allowing or restricting the access to DNA of the machineries involved in gene transcription, replication and DNA repair. Nucleosomes are the basic units of chromatin and consist of ~ 145–147 bp of DNA wrapped around an octamer core of histones

that consists of two copies of each of the histones H2A, H2B, H3 and H4 (Sanicola *et al*., 1990; Luger *et al*., 1997). An additional histone, histone H1 (H1) or linker histone, binds to the DNA that links nucleosomes, as well as DNA entering and exiting the nucleosome core (Noll and Kornberg, 1977; Allan *et al*., 1980). The binding of H1 stabilizes nucleosomes and promotes the folding of chromatin (Robinson and Rhodes, 2006). Histone H1 is the least conserved of all histones (Izzo *et al*., 2008). In mouse and humans, histone H1 is encoded by multiple non-allelic genes, which form several protein variants (11 variants in humans) that can differ in their temporal and cell-type expression profiles, chromatin-binding affinities, sub-nuclear localization and post-transcriptional modifications (Izzo *et al*., 2008). In contrast, *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Tetrahymena thermophila* each have a single H1 gene (Kasinsky *et al*., 2001). In *T. brucei*, a highly divergent unicellular eukaryote (Simpson and Roger, 2004; Dacks *et al*., 2008), histone H1 is encoded by a multigene family.

In vitro studies have extensively shown the properties of H1 in chromatin condensation and transcriptional repression (Thoma *et al*., 1979; Shimamura *et al*., 1989; Laybourn and Kadonaga, 1991; Bednar *et al*., 1998). *In vivo* studies have confirmed that H1 is a global regulator of chromatin architecture (Fan *et al*., 2005; Masina *et al*., 2007; Hashimoto *et al*., 2010). However, these studies also revealed that H1 is not a global transcription repressor, but rather regulates the transcription of specific sets of genes by a yet unknown mechanism (Shen and Gorovsky, 1996; Hellauer *et al*., 2001; Fan *et al*., 2005). H1 is also involved in other biological processes, including inhibition of DNA repair, telomere maintenance in yeast and mammalian cells (Downs *et al*., 2003; Murga *et al*., 2007), silencing of retrotransposons at the rDNA locus in *D. melanogaster* (Vujatovic *et al*., 2012) and differentiation and virulence in *Leishmania major* (Smirlis *et al*., 2006). Despite the conservation of some functions, it remains unclear why H1 is dispensable for survival or growth of unicellular eukaryotes such as *S. cerevisiae* and *T. thermophila* (Shen *et al*., 1995; Shen and Gorovsky, 1996; Patterton *et al*., 1998; Fan *et al*., 2003), but is absolutely essential for mammalian embryonic development (Fan *et al*., 2003).

In *T. brucei*, H1 seems to be dispensable for growth in culture, but it is important to maintain *VSG* genes silent and to inhibit *VSG* switching during antigenic variation (Povelones *et al*., 2012). Like in other eukaryotes, H1 promotes chromatin condensation *in vitro* and *in vivo* (Burri *et al*., 1994; 1995). In this study we show that *T. brucei* H1 has a specific role in gene expression, by predominantly repressing Pol I transcription, namely of *VSG* and procyclin loci. Using labelling of nascent transcripts with 4-thiouridine (4sU), we show that H1 appears to act as a transcriptional inhibitor of *VSG* expression sites and procyclin loci. We further show that H1 inhibits DNA repair of methyl methanesulphonate (MMS)-induced lesions, indicating a role for H1 beyond transcriptional regulation. Finally, we show that although H1 seems to be dispensable for growth of parasites in culture, it is required for the rapid progression of an infection in mice, suggesting that regulation by this histone is important for the host–pathogen interactions.

Results

Depletion of H1 causes no considerable changes in parasite growth in culture

The current version of *T. brucei* genome (Berriman *et al*., 2005) indicates that there are five unique H1 genes on chromosome 11. As previously suggested (Grüter and Betschart, 2001), these genes are organized in two clusters belonging to a single polycistronic unit and are interspersed by 5 non-histone genes. We showed by PCR and Southern blot that this gene organization is correct in the *T. brucei* Lister 427 strain, although other H1 alleles may be missing from the genome database (Fig. S1). Since each gene is predicted to encode a unique protein, we assigned them paralogue numbers from H1.1 to H1.5 according to the most recent nomenclature for histone variants (Talbert *et al*., 2012) (Fig. 1A). Alignment of the five H1 protein sequences reveals that the first seven amino acids (a.a.) define three different types of N-terminal sequences (Fig. 1B). We classified them accordingly and named them MAKTT (H1.1), MAKASA (H1.2, H1.3 and H1.5) and MNNTT (H1.4). The sequence of the 3′UTRs is almost identical between H1 classes (96% on average), whereas the 5′UTRs are less conserved and specific for each of the three classes of H1 genes (Fig. S2).

To address the function of H1 in *T. brucei*, we generated RNA interference (RNAi) cell-lines that allowed simultaneous inducible depletion of all classes of H1. Because H1 genes are 89% identical, we expected that siRNAs generated from a dsRNA of a MAKASA gene (Alibu *et al*., 2005) would successfully knock-down all classes of H1 transcripts. Indeed, after inducing RNAi in two independent clones (C1 and C2), the mRNA levels of each of the three H1 classes (Fig. S3) were diminished to 39% on average during 6 days indicating that the knock-down was effective for all H1 classes in both clones (Fig. 1C, $P < 0.001$; Fig. S4A, $P < 0.001$). No significant differences in H1 depletion levels were observed between day 2, 4 and 6 of RNAi induction. The non-induced clones also showed a reduction of H1 transcripts (a decrease of 30% of the parental cell-line PL1S) (Fig. S4B), which is likely due to the previously described leaky expression of the RNAi

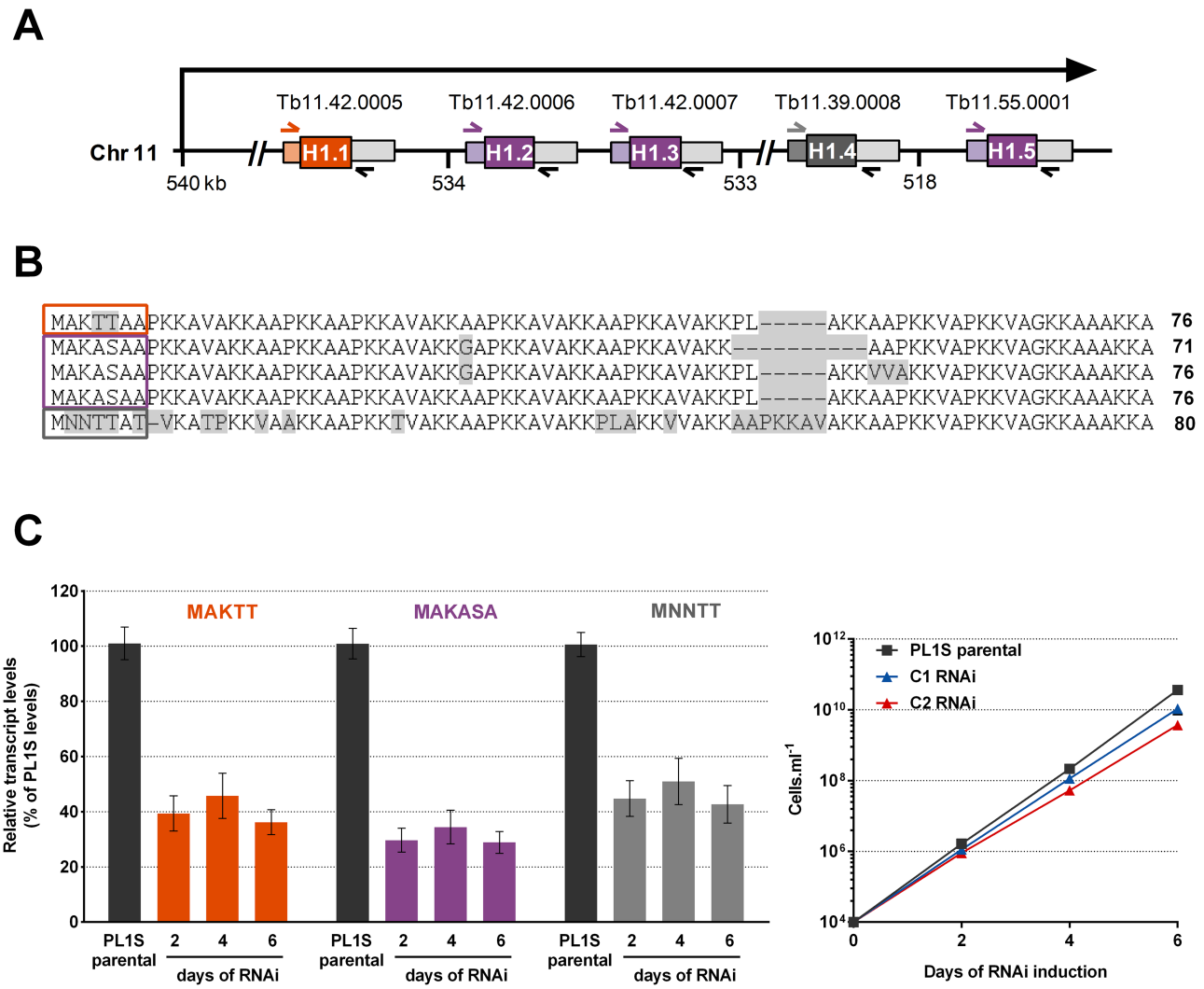


Fig. 1. Depletion of H1 leads to a subtle growth defect in culture.
A. Diagram shows that in *T. brucei* H1 gene family is encoded by five genes (H1.1-H1.5) in two clusters ~ 14.5 kb apart of the same polycistronic unit in chromosome 11. Arrow indicates the transcription start site and direction of transcription. Primers used for qPCR are indicated by colour-code: orange – primer specific for MAKTT 5′UTR; purple – primer specific for MAKASA 5′UTR; grey – primer specific for MNNTT 5′UTR; black – primers that anneal with the 3′UTRs, which are conserved across all H1 genes. Coding sequences are coloured according to the 5′UTR and N-terminal sequence (see panel B).
B. Alignment of the predicted amino acid (a.a.) sequences of H1 variants. Divergent a.a. are shadowed in grey. H1 variants were grouped into three classes according to the N-terminal sequences: MAKTT (orange), MAKASA (purple) and MNNTT (grey).
C. Efficiency of H1 depletion for each class during RNAi induction for 2, 4 and 6 days as an average of the two histone H1 RNAi clones relative to the parental cell-line PL1S. Transcript levels were measured by qPCR. Three independent experiments were analysed.
D. Growth in culture of BSFs of *T. brucei* upon H1 RNAi depletion during 2, 4 and 6 days. Five to nine independent experiments were analysed. Results are shown as mean ± SEM.

cassette (Alsford *et al.*, 2005). In induced clones we observed that depletion of H1 resulted in a minor growth defect *in vitro*. The doubling times were slightly longer in C1 (7:20 h) and C2 (7:40 h) compared with the parental cell-line PL1S (6:30 h) ($P < 0.01$) (Fig. 1D). This is consistent with previous observations (Povelones *et al.*, 2012). Although these mutants are not a complete knockout of the histone H1 gene family, our data suggest that H1 is not essential for growth of *T. brucei* in culture.

H1 compacts chromatin at different levels across the genome

It has been previously demonstrated, using *in vitro* systems, that *T. brucei* H1 can condense chromatin (Burri *et al.*, 1994; 1995). More recently, Povelones *et al.* used electron-microscopy analysis and micrococcal sensitivity assays to show that H1 plays a role in heterochromatin formation *in vivo*. Here we investigated the role of H1 in

chromatin condensation using an alternative approach: FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) of H1-depleted clones subjected to RNAi induction for 6 days. By performing a phenol-chloroform extraction on a cross-linked and sheared chromatin sample, FAIRE fractionates DNA that is preferentially less tightly associated to proteins (such as histones) (Giresi *et al.*, 2007). Quantitative PCR (qPCR) was used to quantify such enrichment and a plasmid spike with an ampicillin-resistance gene (*AmpR*) was used as a normalizer for DNA input. Strikingly, silent BES promoter regions (97 bp downstream of transcription start site) were among the loci in which chromatin opened the most (Fig. 2). In both H1-depleted clones (C1 and C2), chromatin of the silent BES promoter regions opened on average 10-fold compared to the parental cell-line PL1S ($P < 0.001$) (Fig. 2B; Fig. S5A). The chromatin also opened at the luciferase gene (*LUC*), which is 1.2 kb downstream of a silent BES promoter, although in a slightly smaller degree (sevenfold relatively to PL1S; $P < 0.001$). In most other loci, loss of H1 resulted in a 1.8- to 4-fold increase in FAIRE enrichment (Fig. 2B), suggesting that chromatin becomes globally more accessible. These include silent *VSG*s from bloodstream and metacyclic expression sites (*VSG2*, *VSG3*, *VSG18*, *MVSG*; $P < 0.01$), the procyclin promoter region and procyclin *EP2* gene, typically transcribed in procyclic stages ($P < 0.01$), transcription start sites of RNA polymerase II (Pol II) polycistronic units containing either β-tubulin (*β-tub*-PR; $P < 0.01$) or structural maintenance of chromosome 3 gene (*SMC3*-PR; $P < 0.01$) and a number of Pol II-transcribed genes (very abundant *β-tub* and at lower levels *ISP*, *SMC3* and *PAG3*). Three loci showed a non-significant tendency for a more open chromatin conformation: the BES1 actively transcribed genes (blasticidin resistance gene, *BSD*^{R}, and *VSG9*) and the *18S* rDNA, all of which had been previously shown to have a very open, perhaps close to maximal, chromatin conformation (Figueiredo and Cross, 2010; Stanne and Rudenko, 2010).

The FAIRE results were further confirmed by chromatin immunoprecipitation (ChIP) of histone H3 (as a read out of nucleosome density). As previously shown (Figueiredo and Cross, 2010), FAIRE and H3 ChIP are highly consistent techniques and, globally, they reflect almost a mirror image from each other: for most loci, when chromatin becomes more open, we detect an increase in FAIRE-enrichment and a decrease in H3 ChIP for both clones (Fig. 2B–C and Fig. S5). A comparison between the *P*-values of FAIRE and H3 ChIP experiments further confirms the consistency of the two techniques (Table S1). We observe, by ChIP, that most genes lost around 30% of nucleosomes. As expected, ChIP revealed that silent BES promoters are the loci that, upon loss of histone H1, lost more histone H3 (close to 60%). Interestingly, although FAIRE did not detect a very dramatic change in chromatin condensation of the procyclin promoter region, H3 ChIP detected a reduction of nucleosomes more pronounced than average (48%). Also FAIRE did not detect a significant change in chromatin condensation of genes in active BES, but ChIP could detect a significant but modest loss of histone H3 at the promoter region but not at the active *VSG* gene. As expected, a gene encoded by the mitochondrial genome (cytochrome *c* oxidase subunit III, *COIII*, whose genome is not organized around nucleosomes), showed only background levels of histone H3 (Fig. S5B, $P > 0.05$). Overall, these results confirmed that when H1 is depleted there is a global loss of histone H3 across multiple loci of the genome, except for silent BES promoters, which opened 10-fold more (60% loss of histone H3) and procyclin promoters to a lesser extent. Moreover, our data greatly strengthen FAIRE as a robust method to study global changes in chromatin condensation.

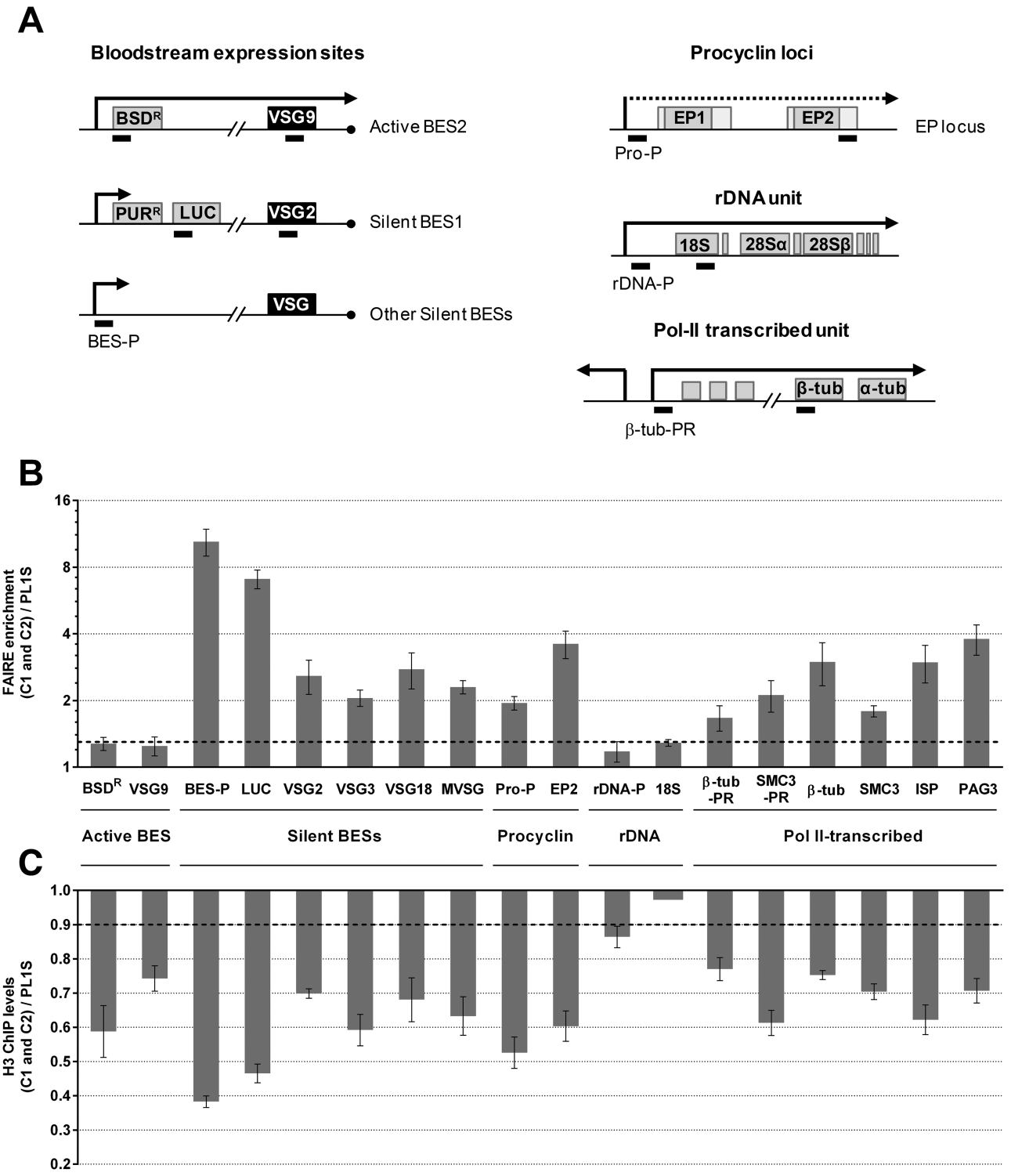
H1 regulates expression of Pol I-transcribed genes

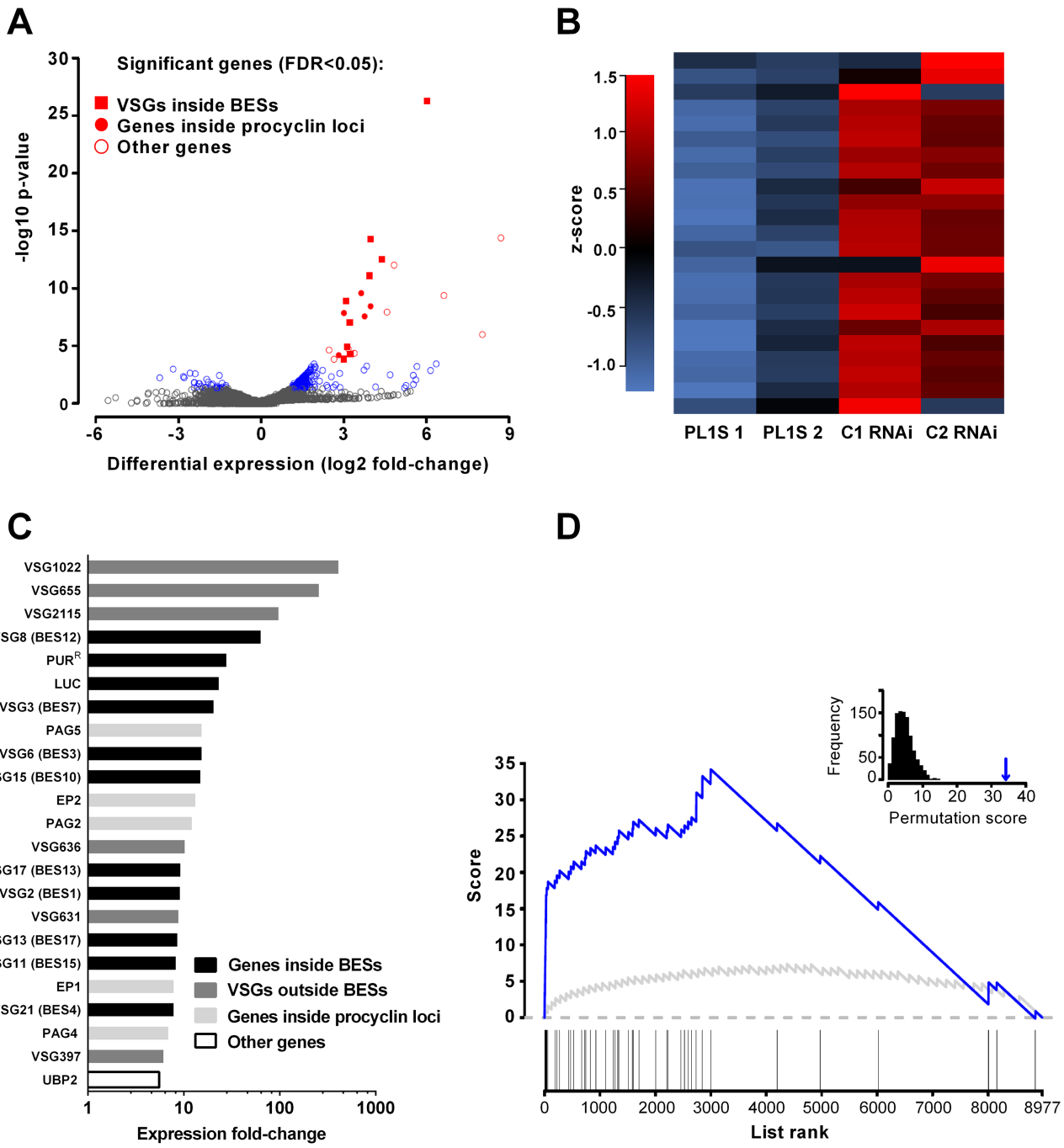
Given the global role of H1 in chromatin condensation, we next tested how changes in chromatin structure affected expression genome-wide. We used RNA-Seq to compare

Fig. 2. H1 compacts chromatin at different levels across genome.
A. Diagram indicates the amplicons (bars) amplified by qPCR for loci in BESs, procyclin loci, rDNA loci and a Pol II-transcribed polycistronic unit. Procyclin loci are partially transcribed in BSFs, as represented by a dashed arrow. Primers for EP2 procyclin are located in the 3′UTR (light grey).
B and C. Chromatin opening was measured by FAIRE (B) and nucleosome occupancy was determined by histone H3 chromatin immunoprecipitation (ChIP) (C) at several gene loci in parental cell-line PL1S and histone H1 RNAi clones 6 days after induction. DNA isolated by FAIRE was quantified by qPCR and an enrichment corresponds to a more open chromatin. An *AmpR* gene contained in a plasmid spike was used as a normalizer for DNA input. For H3 ChIP, DNA was quantified by qPCR, compared to the total input material and normalized to *18S* rDNA. Results are expressed as fold-change relatively to parental cell-line PL1S. Note that FAIRE enrichment is represented in a logarithmic scale. Data for individual clones prior to normalization to the parental PL1S are shown in Fig. S5. Significant fold-change ($P < 0.05$) is shown above (B) and below (C) dashed line in FAIRE and ChIP plots respectively. Two to three independent experiments were analysed. Results are shown in mean ± SEM. BSD^R, blasticidin resistance gene; VSG, variant surface glycoprotein gene; BES-P, promoter region of silent BESs; PUR^R, puromycin resistance gene; LUC, luciferase gene; MVSG, metacyclic VSG gene; Pro-P, procyclin promoter region; EP2, procyclin EP2 gene; rDNA-PR, rDNA promoter region; 18S, ribosomal *18S* gene; β-tub, β-tubulin gene; SMC3, structural maintenance of chromosome 3 gene; β-tub-PR, promoter region of β-tub polycistronic unit; SMC3-PR, promoter region of SMC3 polycistronic unit; ISP, inhibitor of serine peptidase gene; PAG3, procyclin associated gene 3.

the expression profile of H1-depleted clones (at day 4 of RNAi) and parental cell-line PL1S (Fig. 3, Fig. S6, Table S2). H1 depletion resulted in significant changes (false discovery rate (FDR) adjusted *P*-value < 0.05) in the expression of 26 out of 8996 expressed genes (Fig. 3A, Table S2). Interestingly, all were upregulated and all but one gene are transcribed by Pol I (Fig. 3A–C).

Most of these genes (18) are BES-associated genes (*VSG*s and *ESAG*s) or genes located in procyclin loci (EP procyclin genes and *PAG*s), which are normally silent in





this stage of the parasite life cycle. In H1-depleted mutants, BES-associated *VSG*s are on average 18-fold more expressed, with a maximal derepression within this subset close to 65-fold (*VSG8*) (Fig. 3C, Table S2). Other eight *VSG*s annotated as outside a BES were also upregulated. Because we do not know the genomic location of these *VSG*s, we cannot evaluate whether there is a nearby promoter or polycistronic unit that could read through these *VSG*s. The transcript levels of 592 expressed *VSG*s did not change significantly, which is not surprising because in general *VSG* genes are located in tandem arrays that lack a functional promoter (Marcello and Barry, 2007). Most *VSG*s are not expressed and thus were not included in our analysis. H1 depletion also lead to an increased expression of the puromycin resistance (*PUR*^{R}) and the luciferase genes (Fig. 3C), both located immediately downstream of a silent BES promoter (BES1) (Fig. 2A), thus indicating that derepression

Fig. 3. H1 inhibits expression of RNA Pol I-transcribed genes. Genes with altered expression upon H1 depletion (4 days of RNAi) were identified by genome-wide RNA sequencing in BSFs of *T. brucei*.
A. Volcano plot analysis of differentially expressed genes between RNAi clones C1 and C2 relative to two biological replicates of the parental cell-line PL1S shows that H1 regulates a small set of genes, mainly located in procyclin loci (red filled circles) and bloodstream expression sites (BESs) (red filled squares). Each point represents a gene, with log2 fold-change (RNAi/PL1S) of gene expression across samples plotted on the *x*-axis and the corresponding statistical significance (–log10 *P*-value) plotted on the *y*-axis. Genes statistically significant (FDR-adjusted $P < 0.05$) are indicated in red; in grey are non-significant genes; in blue are those genes which would be significant if a less conservative statistical method had been used (FDR-adjusted $P > 0.05$ and $P < 0.05$).
B. Heatmap showing the *z*-scores calculated for the normalized DESeq expression levels of differentially expressed genes (FDR-adjusted $P < 0.05$) obtained by genome-wide RNA sequencing in BSFs of *T. brucei*. Genes are ordered by decreasing fold-change of expression, as in panel C.
C. Bar-plot showing the fold-change of expression of annotated differentially expressed genes in RNAi clones C1 and C2 relative to the parental cell-line PL1S. For *VSG* genes that are inside a BES, the respective BES number in shown in brackets.
D. Gene set enrichment analysis for genes located in BESs and procyclin loci. Blue line indicates the enrichment score distribution across genes ranked by decreasing statistical evidence of differential expression. Grey line represents 95th percentile of the enrichment scores obtained by 1000 random permutations of the gene ranks. The histogram at the top-right corner of the figure represents the distribution of maximum scores obtained by random permutation, with the arrow indicating the experimental enrichment score (i.e. the 'peak' in the main plot). The diagram at the bottom shows where the members of this gene set (i.e. genes located in BESs and procyclin loci) appear in the ranked list of genes. An FDR-adjusted $P < 0.001$ was obtained for this gene set, as described in *Experimental procedures*. Genes with undetermined fold-change due to too low unique read counts were excluded from the analysis.

occurred throughout the entire BES locus. However, as shown by the luciferase assay (Fig 4C), the derepression of silent BESs, is still far from the full activation seen in active BES, indicating that silent BESs are only partially derepressed. Only one of the *ESAG*s was found differentially expressed (*ESAG2* from BES12) most probably because the repetitive nature of *ESAG*s prevented unambiguous read alignment and were thus not included in the analysis. Due to too low unique read counts, the fold-change of *ESAG2* remained undetermined.

Although *PAG*s are located in multiple copies in the genome, in Pol II transcription units (Kim *et al*., 2013) or within the procyclin loci (Haenni *et al*., 2006), the differentially expressed members of *PAG2*, *PAG4* and *PAG5* are specifically located in these latter loci. For example, although there are two *PAG2* genes, one in the procyclin locus and the other in a polycistronic unit on the same chromosome 10, the only gene that is significantly upregulated when H1 is depleted is the one within the procyclin locus.

The results above strongly suggest that H1 predominantly inhibits expression of Pol I genes. We used gene set enrichment analysis (GSEA) (Subramanian *et al*., 2005) to investigate if specific gene sets are enriched or depleted among differentially expressed genes. We found that genes located in BESs and procyclin loci are very significantly enriched in our list of differentially expressed genes (FDR-adjusted $P < 0.001$), as illustrated by the large difference between the detected enrichment score distribution and the 95th percentile of the enrichment scores obtained by random permutations (blue and grey lines, respectively, in Fig. 3D) This enrichment is also observed if the two sets of genes are analysed independently (Fig. S7). This analysis further indicates that H1 plays an important role in the expression of Pol I-dependent genes.

RNA-Seq results were confirmed at the RNA level by qPCR (Fig. 4A–B; Fig. S8) and at the protein level by luciferase activity assay (Fig. 4C). First we checked the correlation between qPCR and RNA-Seq fold-changes of gene expression of 34 genes selected so that we would cover a wide range of expression levels. We observed a very high Pearson's correlation ($r = 0.90$) between the two types of experimental data, suggesting that qPCR and RNA-Seq are highly consistent (Fig. 4A). To further confirm the observation that Pol I, but not Pol II genes become derepressed when H1 is depleted, we measured by qPCR the transcript levels in the two RNAi clones induced for 2, 4 and 6 days (Fig. 4B, Fig. S9). These were normalized to *18S* transcript levels because these are highly expressed genes, depleted of H1 (Povelones *et al*., 2012) and with chromatin largely insensitive to H1 depletion (Fig. 2B). We confirmed that H1 depletion did not affect any of the tested Pol II transcribed genes nor the actively transcribed *VSG9*. As detected by RNA-Seq, *VSG*s in silent BESs (*VSG2* and *VSG3*) were significantly derepressed ($P < 0.001$) (Fig. 4B), although the level of derepression was slightly smaller by qPCR (5- to 12-fold) than RNA-Seq (9- to 21-fold). We also confirmed that procyclin *EP2* was fivefold derepressed. Another *VSG* in a silent BES (*VSG18*) and a *VSG* from a metacyclic expression site (*MVSG*) were found to be four- to fivefold derepressed by qPCR ($P < 0.001$) but were not detected as being differentially expressed in the RNA-Seq data. This inconsistency might be due to the fact that silent *VSG*s and *MVSG*s are expressed at a very low level in bloodstream forms (Donelson, 2003) and for this type of genes, statistical tests for differential expression based on RNA-Seq data are not very powerful.

Finally, BES derepression was further confirmed at the protein level by measuring the activity of luciferase in the

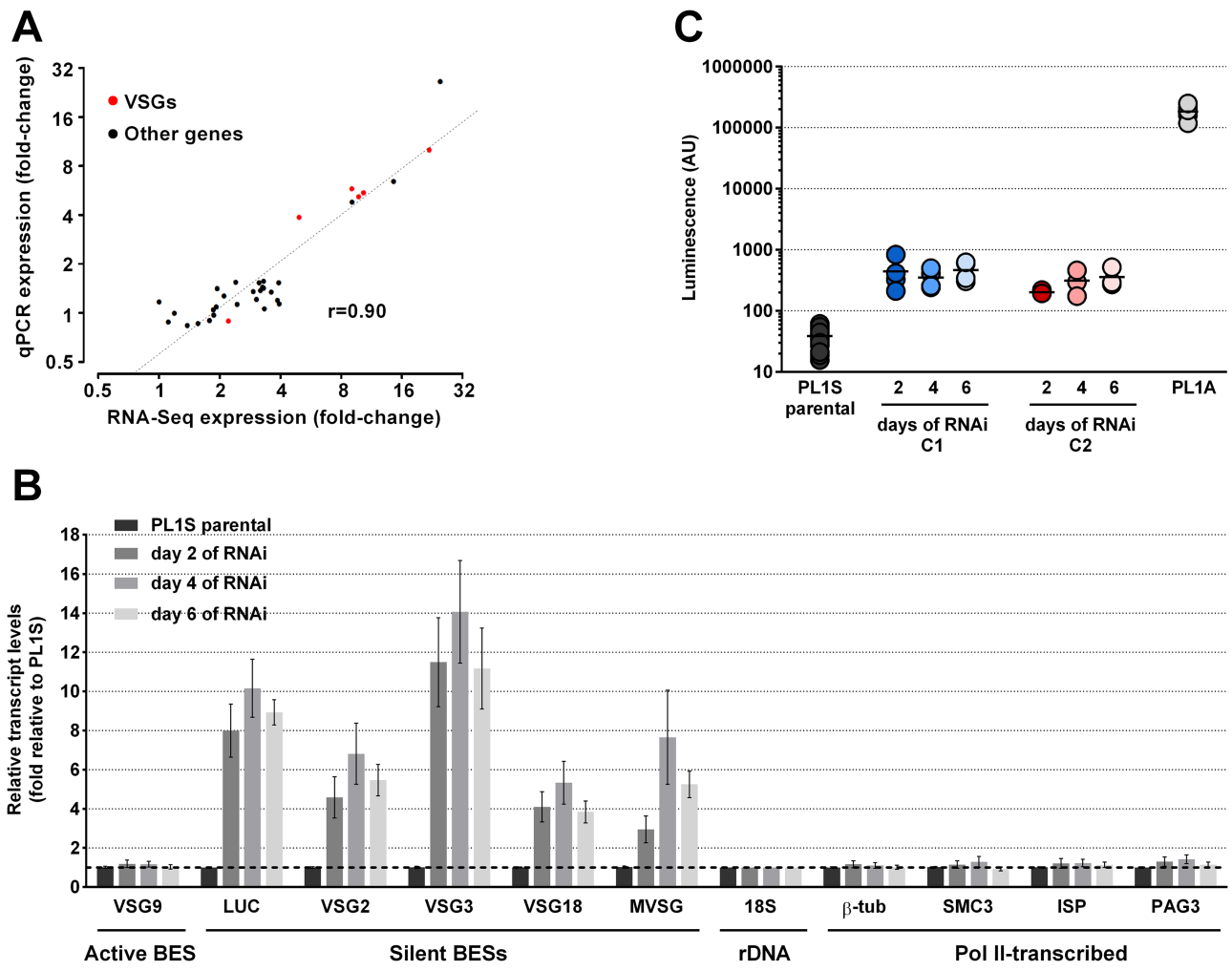


Fig. 4. RNA-Seq data are validated by qPCR and luciferase assays.
A. Scatter plot of fold-changes of gene expression in RNAi clones C1 and C2 relative to the parental cell-line PL1S. Thirty-four genes with mean expression levels ranging from 12 to 512,586 RPKMs by RNA-Seq were randomly selected and quantified by qPCR. Pearson's correlation coefficients (*r*) are indicated.
B. Quantification of mRNA expression after 2, 4 and 6 days of RNAi induction. Transcript levels are plotted as fold-change of RNAi clones C1 and C2 relatively to the parental cell-line PL1S. Transcript levels were measured by qPCR. Three independent experiments were analysed. Results are shown as mean ± SEM.
C. Derepression of silent BESs was assessed by quantification of luminescence throughout 2, 4 and 6 days of RNAi induction. A luciferase reporter gene located downstream of the promoter of silent BES1 (see diagram Fig. 2A) allows the monitoring of its transcriptional activity by measuring the luminescence emitted (AU, arbitrary units). Three to four independent experiments were analysed. PL1A refers to a cell-line that is isogenic of PL1S, but in which BES1 is active and therefore luciferase gene is transcribed at maximal levels.

entire population. We observed an increase in luciferase activity of 7- to 10-fold ($P < 0.001$; Fig. 4C), thus confirming RNA-Seq and qPCR data that silent BES1 was derepressed ($P < 0.001$; Fig. 4B). Overall, qPCR and luciferase activity assays confirmed/validated the RNA-Seq data.

We conclude that, like in genome-wide studies in mammalian cells and yeast (Hellauer *et al.*, 2001; Fan *et al.*, 2005), *T. brucei* H1 is not a global transcription repressor. It rather regulates a particular subset of genes, by repressing expression of Pol I-dependent genes. It is clear that the most significant subgroup of genes that is derepressed are genes from BESs or procyclin loci, confirming that H1 is essential for keeping these specific loci fully silent.

H1 inhibits transcription from silent BESs and procyclin sites

We next sought to understand how H1 depletion leads to an increase of steady-state mRNA levels of the silent *VSG*s and procyclins. Because H1 depletion leads to a more relaxed chromatin at Pol I promoter regions (especially at the silent BES promoters), we hypothesized that the increase in mRNA levels resulted from an increase in

transcription rates at these loci. To test this hypothesis we used metabolic labelling of nascent transcripts with 4-thiouridine (4sU). This method allows the quantification of newly synthesized RNA in live cells, therefore the direct measurement of RNA transcription rates with minimal interference on gene expression and cell viability (Dolken *et al*., 2008). After incubating *T. brucei* BSF parasites with 4sU, total RNA was extracted, thiol-specific biotinylated and subsequently purified on streptavidin-coated magnetic beads. qPCR was then used to quantify the newly transcribed RNAs purified.

We analysed the kinetics of 4sU incorporation into RNA in *T. brucei* during 2, 5, and 10 min of labelling. Since newly transcribed RNA contains higher amounts of unprocessed, primary transcripts, we determined the levels of nascent transcripts by quantifying the ratio of intergenic relative to coding sequence expression for *18S* and β-tubulin transcripts (Fig. 5A). As expected, we observed a rapid increase in the levels of 4sU-tagged RNA with increased duration of labelling (Fig. 5B). Based on this result and on previous reports (Rabani *et al*., 2011; Windhager *et al*., 2012), we chose 10 min as an appropriate 4sU-labelling time to estimate transcription rates. We next compared the levels of nascent transcripts of individual genes in a metabolically labelled sample relative to the total RNA. Although the yield of purification was generally low (0.015–0.025%), for all measured loci we observed that the labelled samples had 7- to 20-fold more nascent transcripts than non-labelled samples (Fig. 5C), indicating that the levels of labelled transcripts are significantly enriched over the background.

To test if the 4sU-labelled RNAs are a product of transcription, we repeated the labelling experiment in cells that were previously treated for 5 min with actinomycin D (ActD), a well-known transcription inhibitor. We observed that, upon transcription inhibition, the levels of 4sU-labelled RNAs recovered by MACS were dramatically reduced to the same low levels detected in the NO 4sU samples (Fig. 5D). Altogether these results further confirmed that 4sU-labelled RNAs are a reflection of the transcription happening inside live parasites.

4sU metabolic labelling was next used to test whether H1 depletion results in a higher transcription rate from BESs and procyclin promoters. For this, RNAi was first induced for 2 days in clones C1 and C2 and, at the end of this period, RNAi clones and parental PL1S parasites were incubated with 4sU. After 10 min the 4sU-labelling reaction was stopped and total RNA was extracted. The relative abundance of intergenic transcripts for *18S* and β-tubulin genes showed that purification of newly transcribed RNA was equally efficient in all cell-lines (Fig. 5E). Analysis of newly transcribed RNA revealed that H1 depletion indeed increases the transcription rate from silent BES promoters, resulting in higher abundance of precursor transcripts from the BES promoter region *BES-P* ($P < 0.001$) and the downstream luciferase gene ($P < 0.001$). When H1 is depleted, transcription rate is also higher from procyclin promoters, since we detected significantly more nascent transcripts from the procyclin promoter region ($P < 0.01$) and at the EP2 procyclin gene ($P < 0.001$) (Fig. 5F). These results show that when H1 is depleted, there seems to be a higher transcription rate of silent BESs and procyclin loci, most likely due to an increase of transcription initiation from the promoter. A Pol II gene (β-tubulin) showed no change in transcription rate, confirming the qPCR analysis of steady-state RNA levels (Fig. 4B). Curiously, the BES promoter regions are the loci in which the levels of nascent transcripts have increased the most (sixfold), which is consistent with the fact that the most dramatic changes in chromatin condensation observed upon H1 depletion were at the silent BES promoters. This may suggest a more significant role for H1 at silent BESs. We also observed a slight increase in transcription rate at the BSD^R gene located downstream of the active BES promoter, consistent with the decrease in H3 detected by ChIP. Taken together, these results not only support our RNA-Seq data but provide additional mechanistic insight, showing that histone H1 is an effective regulator of transcription at the Pol I-transcribed BESs and procyclin loci.

Depletion of H1 increases resistance to MMS-induced DNA damage

The fact that H1 is required for global chromatin compaction, but it has a very limited role in global transcription regulation, led us to hypothesize that H1 may play other functions in the cell. Povelones *et al*. have recently shown that H1 inhibits recombination of telomeric *VSG* genes, but not of a reporter gene *URA3* (Povelones *et al*., 2012). In yeast, H1 suppresses homologous recombination (HR) thus inhibiting DNA repair and promoting genome stability (Downs *et al*., 2003). To test if DNA repair is affected in *T. brucei*, we treated H1-depleted clones with a DNA-damaging agent, methyl methanesulphonate (MMS), and measured the subsequent cell survival. MMS is an alkylating agent that stalls replication forks, which are repaired by homologous recombination (Lundin *et al*., 2005). Parasites in log-phase growth were exposed to MMS for 2 days, after which cell viability was measured with Alamar Blue®. RNAi induction started 2 days before MMS treatment and continued throughout MMS exposure. We observed that both H1-depleted clones were clearly more resistant to MMS-damage than the parental cell-line PL1S (Fig. 6A). IC50 determined from dose–response curve was on average 2.76 parts per million (ppm) for PL1S and significantly higher for C1 and C2, 4.45 and 4.59 ppm respectively ($P < 0.01$) (Fig. 6B). This result indicates that DNA repair is more effective when H1 levels are reduced.

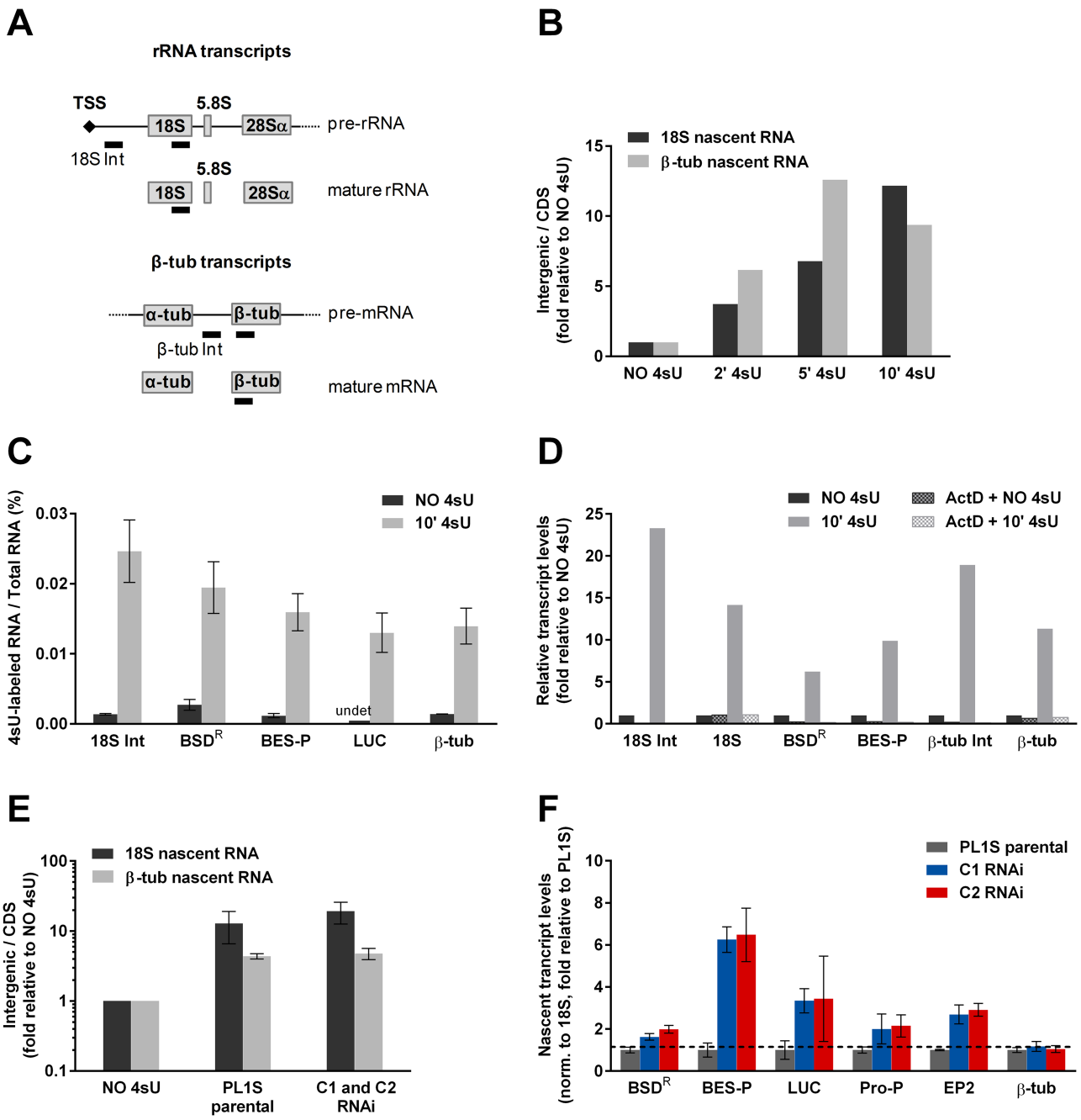


Fig. 5. H1 depletion results in increased transcription rate in silent BESs and procyclin loci. Nascent transcripts were labelled with 4sU for 10 min, biotinylated and streptavidin-purified. RNAs were quantified by qPCR, compared with non-labelled sample and normalized to *18S* rRNA transcripts.

A. Diagram indicates the amplicons (bars) amplified by qPCR for precursor and mature rRNA and β-tubulin transcripts.

B. Kinetics of 4sU-labelling of newly transcribed RNA measured by the ratio of intergenic/CDS transcripts.

C. Percentage of 4sU-labelled RNA relative to Total RNA for several loci with 10 min of 4sU-labelling.

D. Relative transcript levels in 4sU-labelled RNA in the absence or presence of a 5 min treatment with actinomycin D, which inhibits transcription, prior to 4sU-labelling for 10 min.

E. Enrichment in nascent transcripts in PL1S and H1-depleted mutants with 10 min of 4sU-labelling.

F. Fold increase in nascent transcripts in H1-depleted mutants in comparison with PL1S parental cell-line. Significant fold-change ($P < 0.05$) was calculated using an empirical Bayes approach and is shown above dashed line. Two independent experiments were performed; $n = 3–4$ for each cell-line.

Results are shown as mean ± SEM. Primers used to amplify 18S Int were the same used to amplify downstream promoter region of rDNA (rRNA-P). 18S Int, 18S intergenic region; β-tub Int, β-tubulin intergenic region; TSS, transcription start site.

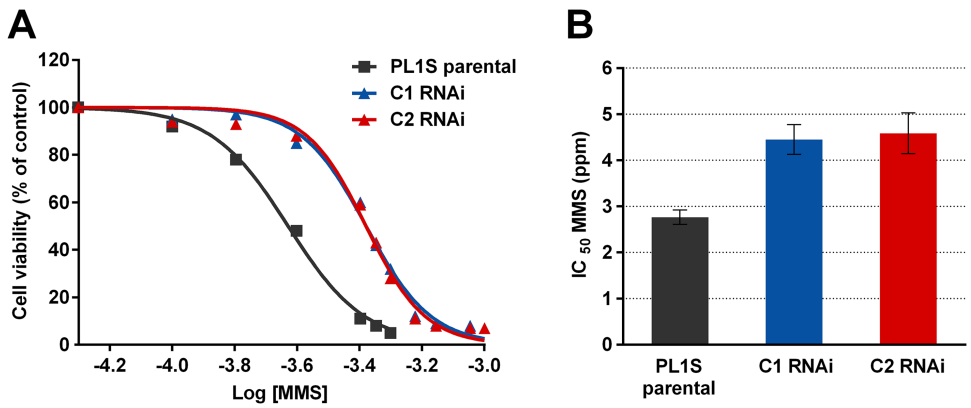


Fig. 6. H1 depletion increases resistance to MMS-induced DNA damage.
A. Dose–response curves of BSF of *T. brucei* to exposure to MMS, in parental and H1-depleted clones. Curves were obtained by fitting a non-linear regression with variable slope to the data. Cell viability was assessed by AlamarBlue® staining at the end of the treatment.
B. IC50 values calculated from the regressions. MMS concentration is expressed in parts per million (ppm) of volume. Seven to five independent experiments were analysed. Results are expressed in mean ± SEM.

This could be because H1 limits access to DNA repair machinery and, as a result, suppresses homologous recombination.

H1 is important for parasite fitness in vivo

Given the involvement of *T. brucei* H1 in DNA repair, *VSG* transcription and global chromatin condensation, it is surprising how H1-depleted parasites can grow so well in culture. Next, we investigated the impact of H1 depletion during an infection in mice. We inoculated C57BL/6 mice with 100 parasites of the parental cell-line, PL1S, or H1-depleted clones. In the latter, RNAi was pre-induced for 2 days with tetracycline. As previously published, mice also received doxycycline, a tetracycline analogue, in water 2 days before infection and during entire infection (Lecordier *et al.*, 2005). Doxycycline had no effect on PL1S-infected mice (Fig. S10A). Interestingly, mice infected with H1-depleted parasites survived longer ($P < 0.01$) (Fig. 7A). Indeed, whereas 55% of the PL1S-infected mice died in the first 9 days of infection, the majority of C1 or C2-infected mice survived until day 13 and 15, respectively. Such prolonged life-span was associated with a delay in parasite appearance in the blood (Fig. 7B; Fig. S11). In fact, whereas parental parasites reached detectable parasitaemia within 4–7 days of infection, H1-depleted clones took 4–15 days. This is much later than what would be expected from the growth delay measured *in vitro* (detectable parasitaemia should be reached at day 4–5), which indicates that host factors must have contributed to this diminished parasitaemia. It is interesting to note that the least virulent clone (C2) was the one with the most efficient H1 depletion ($P < 0.05$; Fig. 7A–C), suggestive of a dose-dependent response. Consistent with the fact that in 'non-induced' RNAi clones H1 transcripts are in fact 20% lower than in parental strain (Fig. S4B), in mice we observed that these clones caused a slight increase in life-span ($P < 0.01$) (Fig. S10B–C). These results clearly indicate that although H1-depleted cells have hardly any growth defect *in vitro*, H1 is essential for parasite fitness *in vivo*.

In around 50% of the mice infected with the parental cell-line PL1S and 40–75% of the mice infected with H1-depleted clones, two peaks of parasitaemia were detected (Fig. 7B; Fig. S11). In order to test if parasites detected in the second peak of H1-depleted clones were RNAi revertants, i.e. parasites that have lost the capacity of performing RNAi, we measured by qPCR the H1 transcript levels at 9–14 days post-infection. We observed that H1 transcripts were still reduced to around 40% and 12%, confirming that parasites were still depleted of H1 (Fig. 7C).

Typically in a *T. brucei* infection, parasites in the second peak of parasitaemia express a *VSG* different from the one expressed in the initially injected parasites. In order to confirm that VSG switching took place, we collected C1 and C2 parasites from the blood on days 9–14 post-infection and tested whether they were still expressing the original *VSG9* or if this *VSG* had become silent. By qPCR we confirmed that in H1-depleted parasites *VSG9* was no longer transcribed confirming that these mutant parasites had switched *VSG*s (Fig. 7 D). Overall, these results show that when H1 is depleted, parasite fitness becomes compromised but parasites are still capable of switching their *VSG*s.

Discussion

H1 condenses chromatin globally but has a more pronounced effect at silent Pol I transcription units

The ability of histone H1 to drive chromatin condensation has been demonstrated both *in vitro* (Thoma *et al.*, 1979;

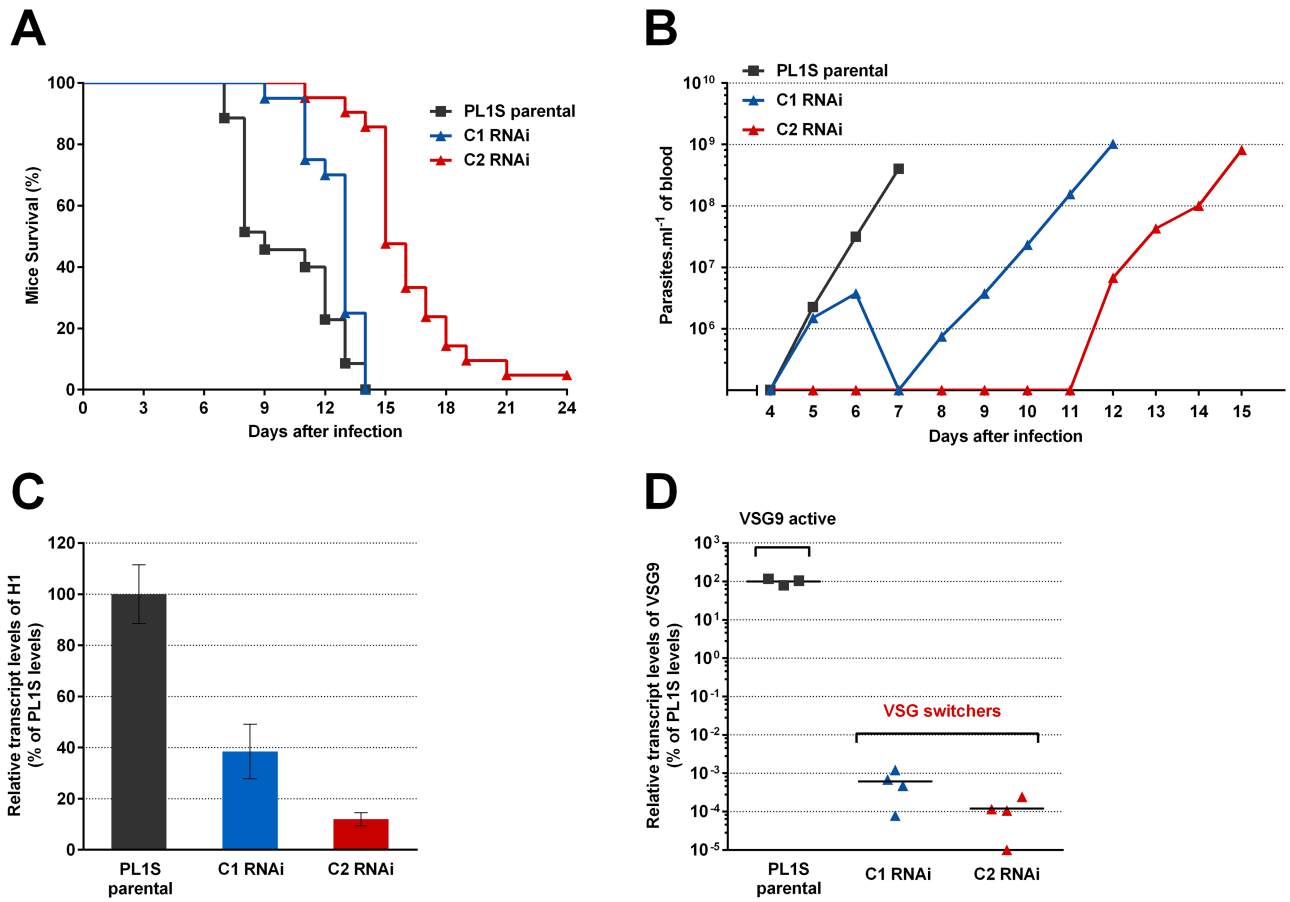


Fig. 7. H1 depletion compromises *T. brucei* fitness *in vivo*.
A. Survival of mice infected with 100 parasites of parental cell-line PL1S or the two H1-depleted clones (C1 or C2). Three to four independent experiments were analysed. Mice $n = 20–35$ per group.
B. Representative examples of parasitaemia in mice infected with *T. brucei* parasites depleted of H1. While PL1S parasites typically reached a peak in parasite number in 7–12 days, histone H1-depleted parasites had a delay of several days in parasitaemia.
C. Depletion of histone H1 in *T. brucei* parasites isolated from the blood of infected mice. PL1S parasites were collected at day 7 post-infection and C1 and C2 parasites at days 9–14 post-infection. Transcript levels of histone H1 were quantified with primers for all histone H1 classes by qPCR. Results are shown as percentage of transcript levels relative to PL1S levels.
D. Quantification of *VSG9* expression in parasites isolated from the blood of mice (same parasites analysed in (C)). C1 and C2 parasites underwent VSG switching and no longer express *VSG9*. Transcript levels were measured by qPCR. Mice $n = 3–4$ per group. Results are shown as mean ± SEM.

Burri *et al.*, 1994; 1995; Bednar *et al.*, 1998) and *in vivo* (Fan *et al.*, 2005; Masina *et al.*, 2007; Hashimoto *et al.*, 2010; Povelones *et al.*, 2012) for *T. brucei* and several other eukaryotes using multiple experimental assays. Using FAIRE and ChIP of H3, we showed here that depletion of H1 resulted in a global increase of chromatin accessibility (2.9-fold) and a 36% reduction of histone H3 of most loci, suggesting that H1 maintains chromatin compacted throughout the genome (Fig. 2B–C). The first exception to this rule is the actively transcribed *VSG* and *BSD*^{R} genes (BES1) and *rDNA 18S* gene, which probably already have a maximal chromatin decompaction and therefore chromatin remained essentially open in the presence of normal or reduced levels of H1. The most important exception to this general trend was the chromatin of silent BES promoters, which became 10-fold more open and lost more than 60% of histone H3. Procyclin promoter regions showed an average decompaction by FAIRE, but a significant drop in H3 content (48%), suggesting that the chromatin of these loci is also relatively sensitive to H1 loss. A similar trend was observed by Povelones *et al.* (2012) using an MNase sensitivity assay. Chromatin immunoprecipitation showed that silent BES promoters have more H1 than procyclin promoters (Povelones *et al.*, 2012), which may explain why depletion of H1 results in a more dramatic chromatin opening at the silent BESs than procyclin promoter regions. We conclude that, although *T. brucei* H1 has a global role in chromatin condensation, there are genomic sites, such as the BES promoters, in which H1 role seems to be more important or less redundant.

Consistent with an important function of H1 in the chromatin of BESs and somewhat procyclin promoter regions, almost all derepressed genes identified by RNA-Seq are from Pol I loci (*VSG*s, procyclins and *PAGs*) (Fig. 3D). It should be noted that rRNA genes, although transcribed by Pol I, were not included in our analysis due to their repetitive nature. How does histone H1 actually affect Pol I gene expression? To test if H1 acts at the transcriptional level, we chose to measure transcription rate by metabolically labelling nascent transcripts with 4sU nucleoside analogue. The advantage of this approach over the classic Nuclear Run-On is that the assay is done in unperturbed cells, instead of permeabilized cells or isolated nuclei. Despite a low yield of recovery, we clearly show that a 4sU-labelled RNA sample is enriched in molecules of RNA that contain intergenic sequences, which are present in primary, unprocessed transcripts (so-called nascent transcripts). Moreover, we showed that 4sU-labelling is dependent on transcription, suggesting that the transcript levels in 4sU-labelled RNAs are a reflection of transcription rates of these genes. Using this metabolic labelling, we observed that depletion of H1 apparently leads to an increased transcription rate of silent procyclin loci and BESs, but no changes in transcription of Pol II loci (Fig. 5). Because we detected BESs derepression with a pair of primers located immediately downstream of the BESs silent promoters, we can conclude that H1 acts at least in part as an inhibitor of transcription initiation. These results are consistent with a recently published work that showed that BES monoallelic transcription is controlled at least partially at the level of transcription initiation, which is dependent on the CITFA complex (Nguyen *et al*., 2014). To our knowledge this is the first mechanistic evidence in *T. brucei* on how depletion of a chromatin component results in an increase of steady-state transcripts by increasing transcription rate, particularly at BESs.

It is interesting to notice that, upon H1 depletion, the FAIRE analysis revealed that chromatin of a silent BES opened 10-fold at the promoter, 7-fold at the luciferase gene (1.2 kb downstream of BES promoter) and 2.6-fold at *VSG2* (56.4 kb downstream BES promoter). Such gradual decreasing effect of chromatin decompaction with increasing proximity to telomeres was also detected by H3 ChIP and it suggests that changes in chromatin is not homogeneous throughout the entire BES. These chromatin changes are also reflected at the RNA level, since a gene close to the promoter of BES1 (Luciferase) was more derepressed than VSG2, which is at the telomeric end of the same BES (Fig. 4A and RNA-Seq, data not shown). Such disparities of derepression at the beginning and end of a silent BES have been observed in other mutants, namely in mutants for TbISWI, TbSPT16, TbDAC3 and TbNLP, in which the BES promoter region was derepressed, but not *VSG*s (Hughes *et al*., 2007; Denninger *et al*., 2010; Wang *et al*., 2010; Narayanan *et al*., 2011). In contrast, deletion of TbRAP1 and TbMCM-BP resulted in a stronger derepression of VSG genes than silent BES promoters (Yang *et al*., 2009; Kim *et al*., 2013). Although it is not clear how these factors interact with each other to regulate VSG expression, these studies are in agreement with previously described silencing mechanisms, which act either at transcription elongation level (Vanhamme *et al*., 2000) or emanate from telomeric silencing (Yang *et al*., 2009). In this context, we propose that histone H1 helps maintaining a compact chromatin structure throughout silent BESs, but with a predominant role at the promoter regions. When H1 is depleted, chromatin of silent BES promoter regions opens dramatically, facilitating transcription initiation by RNA polymerase I. At the telomeric end of BESs, H1 depletion also leads to opening of chromatin but not so dramatically, thus inhibiting transcription elongation. Most likely, the same mechanism happens at procyclin loci.

Although depletion of H1 caused global changes in chromatin condensation, the number of genes with significantly altered expression levels was low (only 26 out of 8996 expressed genes). Similar trends have been observed in the past: only 26 of over 6000 genes were downregulated in H1 knockout mutants of yeast, (Hellauer *et al*., 2001); and 29 genes in mammalian cells showed an altered expression of more than twofold (Fan *et al*., 2005). In contrast to these organisms, however, *T. brucei* H1 acts only as a negative-regulator of gene expression, since H1 depletion results in higher transcripts levels for all significant genes. We conclude that (i) H1 is not a global regulator of gene expression in *T. brucei*, (ii) H1 predominantly represses transcription of silent Pol I genes, and (iii) H1 inhibits at least in part transcription initiation from silent BESs. These specific roles of histone H1 in Pol I loci is likely due to the fact that Pol I genes are probably the only ones that are transcriptionally regulated in *T. brucei*. In fact, because most Pol II-dependent genes are organized in polycistronic units, their expression is believed to be constitutive and gene regulation of Pol II to occur mainly post-transcriptionally.

H1 plays a role in DNA repair

It is interesting to notice that the discrepancy between a global role of H1 in chromatin compaction and the limited role in transcription regulation has also been observed in other organisms (Hellauer *et al*., 2001; Fan *et al*., 2005). It is therefore likely that *T. brucei* H1 plays important but yet unsolved roles that go beyond transcription regulation. Indeed, in *T. brucei*, as in *S. cerevisiae* (Downs *et al*., 2003), H1 is inhibitory of DNA repair (Fig. 6). This could be due to the role of H1 in keeping chromatin closed and thus

refractory to DNA repair machinery. These results are also consistent with the observations that H1-depleted cells undergo more frequent *VSG* gene conversion, a recombination-based switching mechanism (Povelones *et al*., 2012).

H1 depletion showed no changes of parasite growth in culture but revealed important differences of parasite fitness in vivo

In *T. brucei*, when 61% of H1 is depleted by RNA interference a modest defect on cell growth was observed for cultivated bloodstream parasites (Fig. 1 D) (Povelones *et al*., 2012). This suggests that, like other lower eukaryotes, H1 may not be essential for *T. brucei* survival in culture. An alternative explanation is that the remaining H1 (20–30%) is sufficient for *T. brucei* survival. When we infected mice with H1-depleted parasites, these mutant parasites were substantially less infective, allowing mice to survive up to twice longer (Fig. 7A–B). This is probably associated to the lower parasitaemia in the first days of infection. Why is parasitaemia lower in an infection with H1-depleted parasites? This cannot be simply explained by the slower parasite growth rate detected *in vitro*. Indeed, based on *in vitro* growth rate, we estimated a delay in parasitaemia of up to 1 day, but *in vivo* the delay ranged from 1 to 10 days, which indicates that H1 is necessary for adapting to the host environment at the onset of infection. Several mechanisms could explain the reduced fitness of H1-depleted parasites. First, it is possible that the simple deregulation of gene expression could have a negative impact on putative mechanisms that may allow Trypanosomes to sense and adapt to the host environment. These genes could be among the 26 differentially expressed genes, or other genes whose repetitive nature prevents unambiguous mapping of RNA-Seq reads and accurate expression analysis (ESAGs and rRNAs, for example). Consistent with this hypothesis, overexpression of H1 in *L. major* results in a delay in differentiation of promastigotes into amastigotes (Smirlis *et al*., 2006). Second, derepression of silent *VSG*s or increased *VSG* switching frequency may make the first peak of parasitaemia a heterogeneous one that is more easily controlled by the immune system. Indeed, Povelones *et al.* showed that deletion of H1 results in a slight increase in VSG switching (Povelones *et al*., 2012).

When a second population of parasites establishes (switchers) (Fig. 7D), parasitaemia increases very rapidly, eventually killing the mouse. One possible explanation for this rapid growth is that, after 9–14 days of infection, parasites have activated a compensatory pathway for H1 function, which could involve replacement by other small chromatin-interacting proteins such as High-Mobility Group proteins (Hill and Reeves, 1997).

In this work we showed that *T. brucei* H1 keeps chromatin globally compact, inhibits DNA repair and represses the transcription of Pol I silent genes, including *VSG*s and procyclins, which become around 15-fold derepressed on average. We showed that such derepression is associated to an opening of chromatin at silent BESs and silent procyclin loci, which likely results in the observed increase in transcription rate from both types of promoters. Our data show that H1 keeps chromatin of silent BES promoter and to a lesser extent chromatin of procyclin promoters in a condensed state, keeping transcription rate of these loci low. It would be interesting to determine if, in other organisms, histone H1 also acts as an inhibitor of Pol I transcription.

Experimental procedures

Cell-lines and growth medium

Trypanosoma brucei bloodstream-form (BSF) parasites (strain Lister 427, antigenic type MiTat 1.2, clone 221a) (Johnson and Cross, 1979) were cultured in HMI-11 as described in Hirumi and Hirumi (1989). PL1S cell-line was described in Yang *et al.* (2009). Cells were diluted according to their doubling times in order to maintain log-phase growth. Cell growth was measured by counting cells in a Neubauer chamber. H1 RNAi cell-line clones (C1 and C2) were obtained by transfecting *T. brucei* PL1S cells with a construct (pLF100i), p2T7TA RNAi vector in which a MAKASA H1 gene is under the inducible control of two opposing T7 promoters (Alibu *et al*., 2005). RNA interference (RNAi) was induced by adding 1 μg ml^{-1} of tetracyclin (Fisher Scientific) to the medium for a maximum of 6 days. Cloning was performed by recombination of PCR-amplified fragments using the In-Fusion® HD Cloning system (Clontech).

Real-time quantitative PCR analysis

Parasites were harvested by centrifugation at 650 *g* for 10 min, 4°C and immediately resuspended in PureZOL (Bio-Rad) or TRIzol (Invitrogen). RNA was isolated following the manufacturer's instructions and RNA quantity and quality was assessed on a NanoDrop 2000 (Thermo Scientific). The procedure used to isolate nascent RNA labelled with 4sU is described in a section below. cDNA was generated using a Superscript cDNA Synthesis Kit (Invitrogen), according to manufacturer's protocol. Quantitative PCR (qPCR) was performed using 1 × SYBR Green PCR Master Mix (Applied Biosystems). Negative controls lacking reverse transcriptase (RT-) were confirmed by qPCR. Primer efficiencies were determined using standard curves with 3-logs coverage. Amplification reactions were performed in duplicates. Relative quantification was performed based on the CT (cycle threshold) value and the PCR efficiency-corrected method of Pfaffl (Pfaffl, 2001). DNA and transcript levels were normalized to those of *18S*, which did not change significantly after histone H1 depletion. Primer sequences are listed in Table S3.

FAIRE

Formaldehyde-assisted isolation of regulatory elements (FAIRE) was adapted to *T. brucei* as described previously (Figueiredo and Cross, 2010). The main difference is that an external DNA spike consisting of the commercial vector pBluescript-SK (Promega) was added to the samples prior to sonication. Quantification of the FAIRE and total DNA samples was performed by real-time qPCR as described above.

Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed essentially as described (Siegel *et al*., 2009). Histone H3 custom made antibodies were provided by Christian Janzen.

RNA-Seq data analysis

Total RNA was extracted 4 days after induction of H1 depletion *in vitro*. Ribosomal RNA was depleted with RiboMinus kit (Invitrogen). PolyA-containing RNA was purified using Dynabeads oligo-(dT) (Invitrogen) and cDNA was synthesized using random hexamers according to Complete RNA-seq Library System kit (NuGEN) instructions. Samples were sequenced in an Illumina HiSeq2000 platform (EMBL). Sequence reads were 100 nucleotides long and paired-ended. Their quality was accessed using the FASTQC quality control tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Contiguous read segments for which the quality score at each base was greater than 28 (-h 28), and longer than 25 nucleotides (-l 25) were selected using the DynamicTrim and LengthSort Perl-based software respectively (Cock *et al*., 2010). The sequencing data were then aligned to a hybrid genome composed of the Tb927 genome (TriTrypDB v 4.1), in which the *VSG* coding regions in subtelomeric *VSG* arrays were removed and replaced with the Tb427 *VSG* coding regions as a separate chromosome. Alignments were obtained using bowtie 2 (Langmead *et al*., 2009), allowing for one mismatch. Non-unique reads (i.e. aligning in different locations) were excluded and the number of reads mapping to each feature (gene) was measured with the htseq-count software (http://www-huber.embl.de/users/anders/HTSeq/doc/count.html) using the 'intersection-strict' mode. These sequence data have been submitted to the ArrayExpress database (EBI-EMBL) under Accession No. E-MTAB-1715.

Differential gene expression was analysed in R (v 3.0.2), using DESeq (v 1.14.0) (Anders and Huber, 2010) from Bioconductor (v 2.6). DESeq uses the negative binomial distribution for modelling read counts per genomic feature. We first estimated the relative library sizes with the function estimateSizeFactors: for a matrix of raw read counts (where samples are represented in columns and genes in rows), each column is divided by the geometric means of the rows and the median of these ratios is used as the size factor for that specific column. Then, we used the function estimateDispersions: for each condition, it first computes an empirical distribution value for each gene, then fits by regression a dispersion-mean relationship and finally chooses for each gene, from the empirical and the fitted value, a dispersion parameter that will be used in subsequent tests. The function nbinomTest was finally used to test for differences between the base means of the two conditions (RNAi and PL1S in this case).

Feature regions were defined as gene coding sequence (CDS) regions, except in cases where the number of uniquely mappable positions inside the CDS is smaller than 100, in which case UTRs (annotated by Nilsson *et al*., 2010) were also included. To determine uniquely mappable positions we extracted all 100bp sequences (pseudo-reads) appearing in the reference genome and mapped them back using the same software and parameters used previously. After filtering out non-unique alignments, positions where one pseudo-read was successfully mapped were considered mappable. We corrected for multiple comparisons by controlling the FDR, following the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995), which relies on the *P*-values being uniformly distributed under the null hypothesis consists of sorting the *P*-values in ascending order, and then dividing each observed *P*-value by its percentile rank to get an estimated FDR (Noble, 2009). Genes with a FDR-adjusted *P*-value < 0.05 were considered significant.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) was performed as described in Subramanian *et al*. (2005). GSEA allows us to investigate if specific gene sets (i.e. groups of genes that share common biological properties of interest) show statistically significant, concordant differences between two biological states and are therefore correlated with the phenotypic distinction under study. Non-expressed genes were removed from the analysis and the remaining were ranked by differential expression *P*-value. For a given gene set S (in Fig. 3D 'Genes inside BESs and procyclin loci'; in Fig. S7 'Genes inside BESs', 'VSGs inside BESs', 'VSGs outside BESs' and 'Genes inside procyclin loci'), an enrichment score was calculated by walking down the ranked list of genes increasing a running-sum statistic when the gene is in S and decreasing it otherwise and extracting its maximum value at the end. Thus, a high enrichment score indicates a preference for the genes in S to be among the more significantly differentially expressed genes. Statistical significance of this enrichment was determined by comparing the observed maximum enrichment score for gene set S with the distribution of maximum scores obtained by randomly permuting the gene ranks 1000 times (histogram in the top right corner), allowing the estimate of the associated FDR-adjusted *P*-value. That significance is also illustrated by comparing the distribution of scores across observed gene ranks (blue line) with a simulation of scores distribution if gene ranks were randomly selected (grey line). This simulation was generated by 1000 random permutations of the gene ranks and represented as the 95th percentile of the enrichment scores obtained.

Luciferase activity

Luciferase activity was measured with a Firefly Luciferase Kit (Biotium) from 1.5×10^6 parasites following the manu-

facturer's protocol. Luminescence was measured in an Infinite M200 plate reader (Tecan).

4sU-labelling and purification of nascent transcripts

For metabolic labelling of newly transcribed RNA, 150–185 million bloodstream parasites were collected by centrifugation at 970 *g* for 10 min, at room temp. and, washed three times each with 25, 50 and 2 ml of trypanosome dilution buffer (TDB) (5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$, 20 mM Na_2HPO_4, 2 mM NaH_2PO_4, 20 mM glucose, pH 7.7). Cells were resuspended in 2 ml of TDB and 4-thiouridine (4sU) (Sigma) was added to a final concentration of 500 μM for 2, 5 or 10 min at 37°C. In experiments where transcription was inhibited, actinomycin D (Sigma) was added to a final concentration of 20 μg ml^{-1}, 5 min before starting 4sU-labelling. A negative control sample, with parasites with no 4sU addition (NO 4sU sample), was also included for each experiment. Total RNA was extracted using TRIzol reagent (Invitrogen) following the manufacturer's protocol and RNA was dissolved in RNase-free Tris-EDTA buffer (TE) (10 mM Tris-HCl pH 8, 1 mM EDTA). Sample quantity and quality was assessed in a NanoDrop2000 (Thermo Scientific). Biotinylation and purification of 4sU-labelled RNA was performed essentially as described before (Dolken *et al.*, 2008). Labelled RNA was biotinylated with EZ-Link Biotin-HPDP (Pierce), dissolved in dimethylformamide (DMF). Biotinylation reaction was carried out in labelling buffer (10 mM Tris-HCl pH 8, 1 mM EDTA) and 0.2 mg ml^{-1} Biotin-HPDP for 1.5 h at room temperature. An amount of 50–90 μg of total RNA was used for the biotinylation reaction. Unbound Biotin-HPDP was removed by chloroform/isoamyl alcohol (24:1) extraction. RNA was precipitated with a 1:10 volume of 5 M NaCl, an equal volume of isopropanol and 15 μg of GlicoBlue (Life Technologies) at 20000 *g* for 20 min, 4°C. The pellet was washed with an equal volume of 75% ethanol and resuspended in RNase-free TE. Biotinylated RNA was captured using μMACS streptavidin beads and columns (Miltenyi). Up to 58 μg of biotinylated RNA was incubated with 100 μl of beads for 15 min at room temperature, with rotation. Beads were magnetically fixed and washed 3× with washing buffer (100 mM Tris-HCl pH 7.5, 10 mM EDTA, 1 M NaCl, 0.1% Tween20) at 65°C followed by 3× washes with washing buffer at room temperature. 4sU-labelled RNA was eluted in two rounds by adding 100 μl of freshly prepared 100 mM dithiothreitol (DTT). Eluted RNA was precipitated with a 1:10 volume of 5 M NaCl, 2.5× volume of 100% ethanol and 30 μg of GlicoBlue, at –80°C for at least 30 min. RNA was washed in 75% ethanol and recovered by centrifugation as described above and resuspended in RNase-free water. cDNA synthesis and qPCR analysis was performed as described in a previous section. Levels of nascent transcripts were assessed by calculating the ratio of intergenic/CDS levels for *18S* and *β-tub*. In this case, an internal normalization for *18S* was not needed since two types of transcripts are being compared inside the same sample. The fold-change relative to NO 4sU samples was then calculated and plotted. In the experiment where transcription was blocked by actinomycin D treatment, an equal amount of RNA was used to synthesize cDNA and no further normalization was applied. Transcript levels for each locus in the MACS-isolated fraction were plotted as fold-change to the NO 4sU sample. To compare nascent transcript levels in H1 RNAi clones relative to PL1S, qPCR levels were first normalized to those of *18S* transcripts and afterwards fold-change was calculated for H1 RNAi clones at each locus, relative to its expression in PL1S.

MMS-induced DNA damage

Parasites in which RNAi has been pre-induced for 2 days with tetracycline (1 μg ml^{-1}), were incubated in a 96-well plate (200 μl per well) with sixfold range dilutions of methyl methanesulphonate (MMS; Sigma) in HMI-11, at a final concentration of 0.5–4 × 10^4 cells ml^{-1}. After 2 days, cell viability was measured with AlamarBlue® (Sigma). Ten microlitres of Alamar Blue® was added per well followed by incubation at 37°C, 4 h and fluorescence was measured (530ex/590em nm). RNAi induction with tetracycline and selection with Hygromycin was maintained during the 2-day exposure to MMS. MMS concentration is expressed in parts per million (ppm) of volume.

Mouse infections

Inbred C57BL/6 wild-type mice (Charles River) were housed in the pathogen-free facilities of the Instituto de Medicina Molecular (IMM). The animal facility and the experimental procedures complied with European Guideline 86/609/EC, followed the Federation of European Laboratory Animal Science Associations (FELASA) guidelines and recommendations concerning laboratory animal welfare and were approved by the Instituto de Medicina Molecular Animal Care and Ethics Committee (AEC_2011_006_LF_TBrucei _IMM). Mice were infected intraperitoneally with 100 parasites of *T. brucei* Lister 427 PL1S and RNAi clones C1 and C2 (with and without induction). RNAi induction was initiated in culture with 1 μg ml^{-1} tetracycline 2 days prior to mouse infection. RNAi was maintained *in vivo* in mice by watering them with 1 mg ml^{-1} of doxycycline (doxycycline hyclate, Sigma). Mice received doxycycline 2 days prior to infection and during the whole course of infection. Parasitaemia was monitored throughout infection by collecting blood from the mouse tail. For RNA extraction of parasites in the blood, between day 7 and 14 post-infection, 25–30 μl of blood from the tail of infected mice (parasitaemia ~10^8 parasites ml^{-1} of blood) was collected in red blood cell lysis buffer (150 mM NH_4Cl, 10 mM $KHCO_3$, 1 mM EDTA.Na_2, pH 7.4), washed in 1× TDB and extracted with PureZOL, according to manufacturer's instructions (Bio-Rad). cDNA synthesis and qPCR were performed as described above.

Statistical analysis

To estimate statistical significance of data calculated as a fold increase, we computed a log-normal posterior distribution of the fold-change, using an empirical Bayes approach (FAIRE, ChIP and 4sU experiments). Survival curves were compared by a Log-rank (Mantel-Cox) test. The statistical significance for remaining comparisons was given by Mann–Whitney *U* tests. A *P*-value < 0.05 was considered significant. For the statistical analysis of RNA-Seq data, see RNA-Seq methods above.

Acknowledgements

The authors thank all members of the Parasitology group at IMM and the Lisbon Chromatin Club for helpful discussions; George A.M. Cross for allowing L.M.F. to begin some preliminary experiments during her post-doc, for sequence information on VSGnome, and for critically reading the manuscript; Christian Janzen for the generous gift of H3 antibody; Nina Papavasiliou and Danae Schulz for useful comments; Lars Dölken, Noélia Custódio and Célia Carvalho for the helpful recommendations on the 4sU experiments. This work was supported in part by FCG (P-105333), FCT (PTDC/BIA-BCM/115864/2009), Marie Curie Reintegration Grant (FP7-PEOPLE-RG-2009 'HistoneH1Tryp'), EMBO Installation grant (Project 2151) and Howard Hughes Medical Institute International Early Career Scientist Program (55007419) to L.M.F. and PTDC/BIA-BCM/101575/2008 to M.C.-F. L.P.C. was supported by FCT (grant PTDC/SAU-GMG/115652/2008). A.C.P., F.R.-F., F.A.-B. and R.V.-D. were supported by the fellowships SFRH/BD/73998/2010, SFRH/BD/51286/2010, SFRH/BD/80718/2011 and SFRH/BD/90231/2012 respectively. N.L.B.-M. was supported by a Marie Curie International Outgoing Fellowship (FP7-PEOPLE-2010-IOF 'EvoAltSplice'). The authors declare that they have no conflict of interest.

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*. *Mol Biochem Parasitol* **139:** 75–82.

Allan, J., Hartman, P.G., Crane-Robinson, C., and Aviles, F.X. (1980) The structure of histone H1 and its location in chromatin. *Nature* **288:** 675–679.

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. *Mol Biochem Parasitol* **144:** 142–148.

Anders, S., and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol* **11:** R106.

Bednar, J., Horowitz, R.A., Grigoryev, S.A., Carruthers, L.M., Hansen, J.C., Koster, A.J., and Woodcock, C.L. (1998) Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc Natl Acad Sci USA* **95:** 14173–14178.

Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J Roy Stat Soc B* **57:** 289–300.

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

Burri, M., Schlimme, W., Betschart, B., and Hecker, H. (1994) Characterization of the histones of *Trypanosoma brucei brucei* bloodstream forms. *Acta Trop* **58:** 291–305.

Burri, M., Schlimme, W., Betschart, B., Lindner, H., Kampfer, U., Schaller, J., and Hecker, H. (1995) Partial amino acid sequence and functional aspects of histone H1 proteins in *Trypanosoma brucei brucei*. *Biol Cell* **83:** 23–31.

Clayton, C., and Shapira, M. (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol* **156:** 93–101.

Cock, P.J., Fields, C.J., Goto, N., Heuer, M.L., and Rice, P.M. (2010) The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res* **38:** 1767–1771.

Dacks, J.B., Walker, G., and Field, M.C. (2008) Implications of the new eukaryotic systematics for parasitologists. *Parasitol Int* **57:** 97–104.

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.

Dolken, L., Ruzsics, Z., Radle, B., Friedel, C.C., Zimmer, R., Mages, J., *et al*. (2008) High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA* **14:** 1959–1972.

Donelson, J.E. (2003) Antigenic variation and the African trypanosome genome. *Acta Trop* **85:** 391–404.

Downs, J.A., Kosmidou, E., Morgan, A., and Jackson, S.P. (2003) Suppression of homologous recombination by the *Saccharomyces cerevisiae* linker histone. *Mol Cell* **11:** 1685–1692.

Fan, Y., Nikitina, T., Morin-Kensicki, E.M., Zhao, J., Magnuson, T.R., Woodcock, C.L., and Skoultchi, A.I. (2003) H1 linker histones are essential for mouse development and affect nucleosome spacing *in vivo*. *Mol Cell Biol* **23:** 4559–4572.

Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., *et al*. (2005) Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* **123:** 1199–1212.

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.

Giresi, P.G., Kim, J., McDaniell, R.M., Iyer, V.R., and Lieb, J.D. (2007) FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin. *Genome Res* **17:** 877–885.

Grüter, E., and Betschart, B. (2001) Isolation, characterisation and organisation of histone H1 genes in African trypanosomes. *Parasitol Res* **87:** 977–984.

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

Haenni, S., Renggli, C.K., Fragoso, C.M., Oberle, M., and Roditi, I. (2006) The procyclin-associated genes of *Trypanosoma brucei* are not essential for cyclical transmission by tsetse. *Mol Biochem Parasitol* **150:** 144–156.

Hashimoto, H., Takami, Y., Sonoda, E., Iwasaki, T., Iwano, H., Tachibana, M., *et al*. (2010) Histone H1 null vertebrate cells exhibit altered nucleosome architecture. *Nucleic Acids Res* **38:** 3533–3545.

Hellauer, K., Sirard, E., and Turcotte, B. (2001) Decreased expression of specific genes in yeast cells lacking histone H1. *J Biol Chem* **276:** 13587–13592.

Hertz-Fowler, C., Figueiredo, L.M., Quail, M.A., Becker, M.,

Jackson, A., Bason, N., *et al*. (2008) Telomeric expression sites are highly conserved in *Trypanosoma brucei*. *PLoS ONE* **3:** e3527.

Hill, D.A., and Reeves, R. (1997) Competition between HMG-I(Y), HMG-1 and histone H1 on four-way junction DNA. *Nucleic Acids Res* **25:** 3523–3531.

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. *J Parasitol* **75:** 985–989.

Hughes, K., Wand, M., Foulston, L., Young, R., Harley, K., Terry, S., *et al*. (2007) A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. *EMBO J* **26:** 2400–2410.

Izzo, A., Kamieniarz, K., and Schneider, R. (2008) The histone H1 family: specific members, specific functions? *Biol Chem* **389:** 333–343.

Johnson, J.G., and Cross, G.A. (1979) Selective cleavage of variant surface glycoproteins from *Trypanosoma brucei*. *Biochem J* **178:** 689–697.

Kasinsky, H.E., Lewis, J.D., Dacks, J.B., and Ausio, J. (2001) Origin of H1 linker histones. *FASEB J* **15:** 34–42.

Kim, H.S., Park, S.H., Gunzl, A., and Cross, G.A. (2013) MCM-BP is required for repression of life-cycle specific genes transcribed by RNA polymerase I in the mammalian infectious form of *Trypanosoma brucei*. *PLoS ONE* **8:** e57001.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10:** R25.

Laybourn, P.J., and Kadonaga, J.T. (1991) Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. *Science* **254:** 238–245.

Lecordier, L., Walgraffe, D., Devaux, S., Poelvoorde, P., Pays, E., and Vanhamme, L. (2005) *Trypanosoma brucei* RNA interference in the mammalian host. *Mol Biochem Parasitol* **140:** 127–131.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389:** 251–260.

Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A.S., and Helleday, T. (2005) Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable *in vivo* DNA double-strand breaks. *Nucleic Acids Res* **33:** 3799–3811.

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

Masina, S., Zangger, H., Rivier, D., and Fasel, N. (2007) Histone H1 regulates chromatin condensation in *Leishmania* parasites. *Exp Parasitol* **116:** 83–87.

Murga, M., Jaco, I., Fan, Y., Soria, R., Martinez-Pastor, B., Cuadrado, M., *et al*. (2007) Global chromatin compaction limits the strength of the DNA damage response. *J Cell Biol* **178:** 1101–1108.

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.

Navarro, M., Penate, X., and Landeira, D. (2007) Nuclear architecture underlying gene expression in *Trypanosoma brucei*. *Trends Microbiol* **15:** 263–270.

Nguyen, T.N., Muller, L.S., Park, S.H., Siegel, T.N., and Gunzl, 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 Res* **42:** 3164–3176.

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

Noble, W.S. (2009) How does multiple testing correction work? *Nat Biotechnol* **27:** 1135–1137.

Noll, M., and Kornberg, R.D. (1977) Action of micrococcal nuclease on chromatin and the location of histone H1. *J Mol Biol* **109:** 393–404.

Patterton, H.G., Landel, C.C., Landsman, D., Peterson, C.L., and Simpson, R.T. (1998) The biochemical and phenotypic characterization of Hho1p, the putative linker histone H1 of *Saccharomyces cerevisiae*. *J Biol Chem* **273:** 7268–7276.

Pays, E., Vanhamme, L., and Perez-Morga, D. (2004) Antigenic variation in *Trypanosoma brucei*: facts, challenges and mysteries. *Curr Opin Microbiol* **7:** 369–374.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29:** e45.

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.

Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., *et al*. (2011) Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat Biotechnol* **29:** 436–442.

Robinson, P.J., and Rhodes, D. (2006) Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol* **16:** 336–343.

Rudenko, G., Bishop, D., Gottesdiener, K., and Van der Ploeg, L.H. (1989) Alpha-amanitin resistant transcription of protein coding genes in insect and bloodstream form *Trypanosoma brucei*. *EMBO J* **8:** 4259–4263.

Sanicola, M., Ward, S., Childs, G., and Emmons, S.W. (1990) Identification of a *Caenorhabditis elegans* histone H1 gene family. Characterization of a family member containing an intron and encoding a poly(A) + mRNA. *J Mol Biol* **212:** 259–268.

Shen, X., and Gorovsky, M.A. (1996) Linker histone H1 regulates specific gene expression but not global transcription *in vivo*. *Cell* **86:** 475–483.

Shen, X., Yu, L., Weir, J.W., and Gorovsky, M.A. (1995) Linker histones are not essential and affect chromatin condensation *in vivo*. *Cell* **82:** 47–56.

Shimamura, A., Sapp, M., Rodriguez-Campos, A., and Worcel, A. (1989) Histone H1 represses transcription from minichromosomes assembled *in vitro*. *Mol Cell Biol* **9:** 5573–5584.

Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., *et al*. (2009) Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev* **23:** 1063–1076.

Simpson, A.G., and Roger, A.J. (2004) The real 'kingdoms' of eukaryotes. *Curr Biol* **14:** R693–R696.

Smirlis, D., Bisti, S.N., Xingi, E., Konidou, G., Thiakaki, M., and Soteriadou, K.P. (2006) Leishmania histone H1 overexpression delays parasite cell-cycle progression, parasite differentiation and reduces *Leishmania* infectivity *in vivo*. *Mol Microbiol* **60:** 1457–1473.

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

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., *et al*. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102:** 15545–15550.

Talbert, P.B., Ahmad, K., Almouzni, G., Ausio, J., Berger, F., Bhalla, P.L., *et al*. (2012) A unified phylogeny-based nomenclature for histone variants. *Epigenetics Chromatin* **5:** 7.

Thoma, F., Koller, T., and Klug, A. (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J Cell Biol* **83:** 403–427.

Vanhamme, L., Poelvoorde, P., Pays, A., Tebabi, P., Van Xong, H., and Pays, E. (2000) Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. *Mol Microbiol* **36:** 328–340.

Vujatovic, O., Zaragoza, K., Vaquero, A., Reina, O., Bernues, J., and Azorin, F. (2012) Drosophila melanogaster linker histone dH1 is required for transposon silencing and to preserve genome integrity. *Nucleic Acids Res* **40:** 5402–5414.

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.

Windhager, L., Bonfert, T., Burger, K., Ruzsics, Z., Krebs, S., Kaufmann, S., *et al*. (2012) Ultrashort and progressive 4sU-tagging reveals key characteristics of RNA processing at nucleotide resolution. *Genome Res* **22:** 2031–2042.

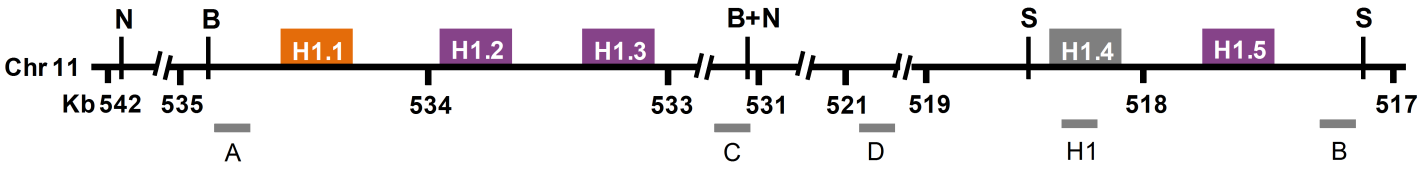
Yang, X., Figueiredo, L.M., Espinal, A., Okubo, E., and Li, B. (2009) RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. *Cell* **137:** 99–109.

Supporting information

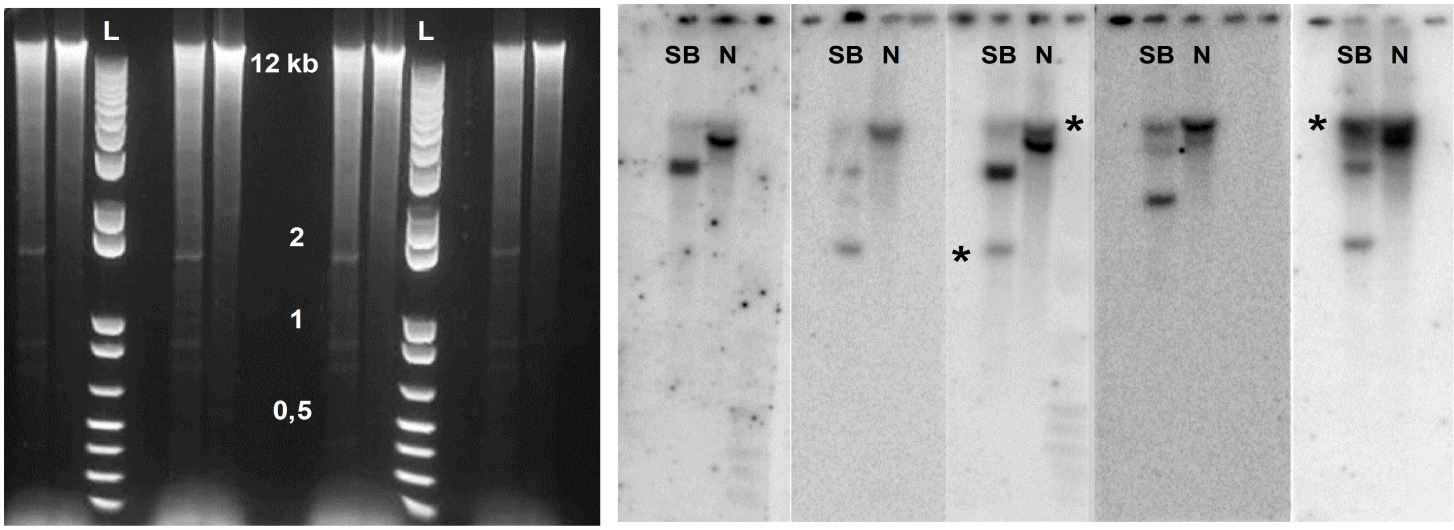
Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1

A



B



C

H1 locus	Primers	Amplicon (Kb)	N-terminal sequences identified
H1.1-1.3	Fwd 5'-ggaaagtagaaaggaaaataaaat-3' Rev 5'-cgttcgttgtacacaggaac-3'	1.65	5'- MAKTT, MAKASA, MAKASA -3'
H1.4-1.5	Fwd 5'-ggcggcaacttactgcag-3' Rev 5'-gtccgatgaacagatgcattc-3'	1.8	5'- MNNTT, MAKASA -3'
Putative new H1 locus	Fwd 5'-ggcggcaacttactgcag-3' Rev 5'-cgttcgttgtacacaggaac-3'	3	MNNTT

Figure S2

A

5'UTR

```
% Identity
44,8%  H1.1  GAAAGTAGAAAGGAAAATAAAAT                    23
76,5%  H1.2  --AAGTCGCAATCTTATCAACACTCGGAAGT            29
76,5%  H1.3  --AAGTCGCAATCTTATCAACACTCGGAAGT            29
76,5%  H1.5  --AAGTCGCAATCTTATCAACACTCGGAAGT            29
45,9%  H1.4  --AAAGCTC---TTTATCGACTCCCCACAAGA           27
```

B

3'UTR

```
% identity
                    10        20        30        40        50        60        70        80
96,9%  H1.1  GCGCATCCGCTGCTGCCCGCTATTAGACACGCTATGAGGTTTACCTGAGTGTGGGAGAAAGCTGTCACACGTTTCAGGAC
94,3%  H1.2  GCGCATCCGCTGCTGCCCGCTATTAGACACGCTATGAGGTTTACCTGAGTGTGGGAGAAAGCTGTCACACGTTTCAGGAC
95,1%  H1.3  GCGCATCCGCTGCTGCCCGCTATTAGACACGCTATGAGGTTTACCTGAGTGTGGGAGAAAGCTGTCACACGTTTCAGAAG
96,2%  H1.5  GCGCATCCGCTGCTGCCCGCTATTAGACACGCTATGAGGTTTACCTGAGTGTGGGAGAAAGCTGTCACACGTTTCAGGAC
97,3%  H1.4  GCGCATCCGCTGCTGCCCGCTATTAGACACGCTATGAGGTTTACCTGAGTGTGGGAGAAAGCTGTCACACGTTTCAGGAC

                    90        100       110       120       130       140       150       160
H1.1  GTCCTCGTGCGTCCCTCCAGGACGGAGTTAGATTTTTCCTATCTTACTTATTTAGTTCCCTTCTACCGTTTTTATTGGAT
H1.2  GTCCTCGTGCGTCTCTCCAGGACGGAGTTAGATTTTTCCTATCTTACTTTTTTATTTCTTTTTCATT-TTTTTATTGGAT
H1.3  GTACTTGTGCGCCCGTCAAGGACGGAGTTAGATTTTTCCTATCTTACTTGTTTAGTTCCCTTCTACCGTTTTTATTGGAT
H1.5  GTCCTCGTGCGTCCCTCCAGGACGGAGTTAGATTTTTCCTATCTTTTTGTTTAGTTCCCTTCTACCGTTTTTATTGG
H1.4  GTCCTCGTGCGTCTCTCCAGGACGGAGTTAGATTTTTCCTATCTTTTTGTTTAGTTCCCTTCTACCGTTTTTATTGGAT

                    170       180       190       200       210       220       230       240
H1.1  AT-TTTCATTTGTGGGTTGCGTCTTATGTACCGCCATGCGGTGTTGGTGTCGTAGCGTTGCAAAGAGCATATCATCCTGA
H1.2  ATGTTTCGTTTGTGGGTTGCGTCTTATGTACCGCCATGCGGTGTTGGTGTCGTAGCGTTGCAAAGAGCATATCATCCTGA
H1.3  ATGTTTCGTTTGTGGGTTGCGTCTTATGTACCGCCATGCGGTGTTGGTGTCGTAGCGTTGCAAAGAGCATATCATCCTGA
H1.5
H1.4  ATGTTTCGTTTGTGGGTTGCGTCTTATGTACCGCCATGCGGTGTTGGTGTCGTAGCGTTGCAAAGAGCATATCATCCTGA

                    250       260       270
H1.1  TGTGTGGCTATTTT                                  253
H1.2  TGTGTGGCTATACTAACTGCCTGTGTATGGTTGTGGTCC         278
H1.3  TGTGTGGCT                                       249
H1.5                                                  158
H1.4  TGTGTGGCTATTTTAACTGCCTGTGTATGGTTGTGGTCC         279
```

Figure S3

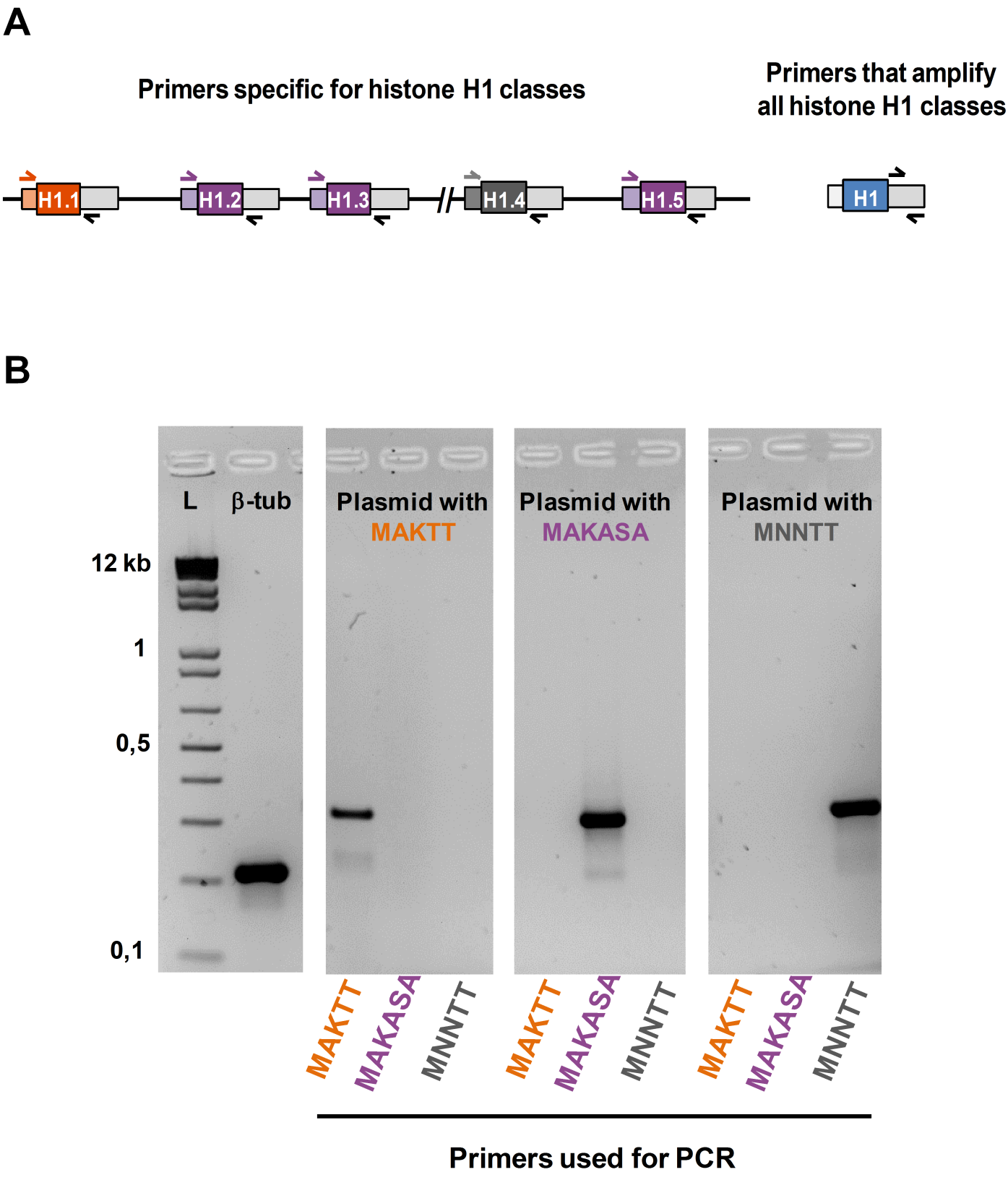

A
Primers specific for histone H1 classes
Primers that amplify all histone H1 classes
H1.1
H1.2
H1.3
H1.4
H1.5
H1
B
L
β-tub
12 kb
1
0,5
0,1
Plasmid with MAKTT
Plasmid with MAKASA
Plasmid with MNNTT
MAKTT
MAKASA
MNNTT
MAKTT
MAKASA
MNNTT
MAKTT
MAKASA
MNNTT
Primers used for PCR

Figure S4

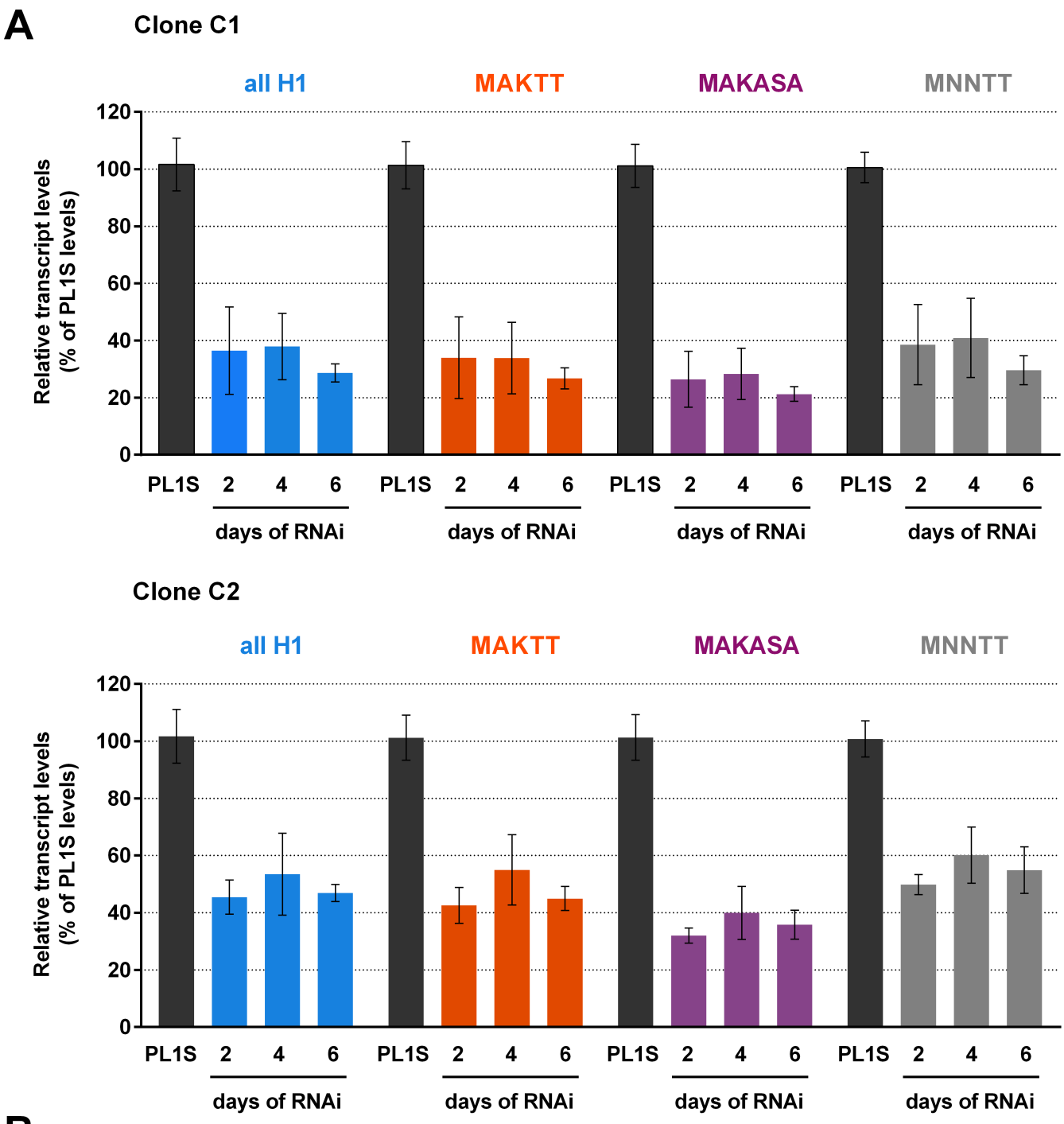

A
Clone C1
all H1
MAKTT
MAKASA
MNNTT
Relative transcript levels (% of PL1S levels)
120
100
80
60
40
20
0
PL1S
2
4
6
days of RNAi
Clone C2
all H1
MAKTT
MAKASA
MNNTT
Relative transcript levels (% of PL1S levels)
PL1S
2
4
6
days of RNAi
B
all H1
Relative transcript levels (% of PL1S levels)
PL1S
2
4
6
non-induced RNAi

Figure S5

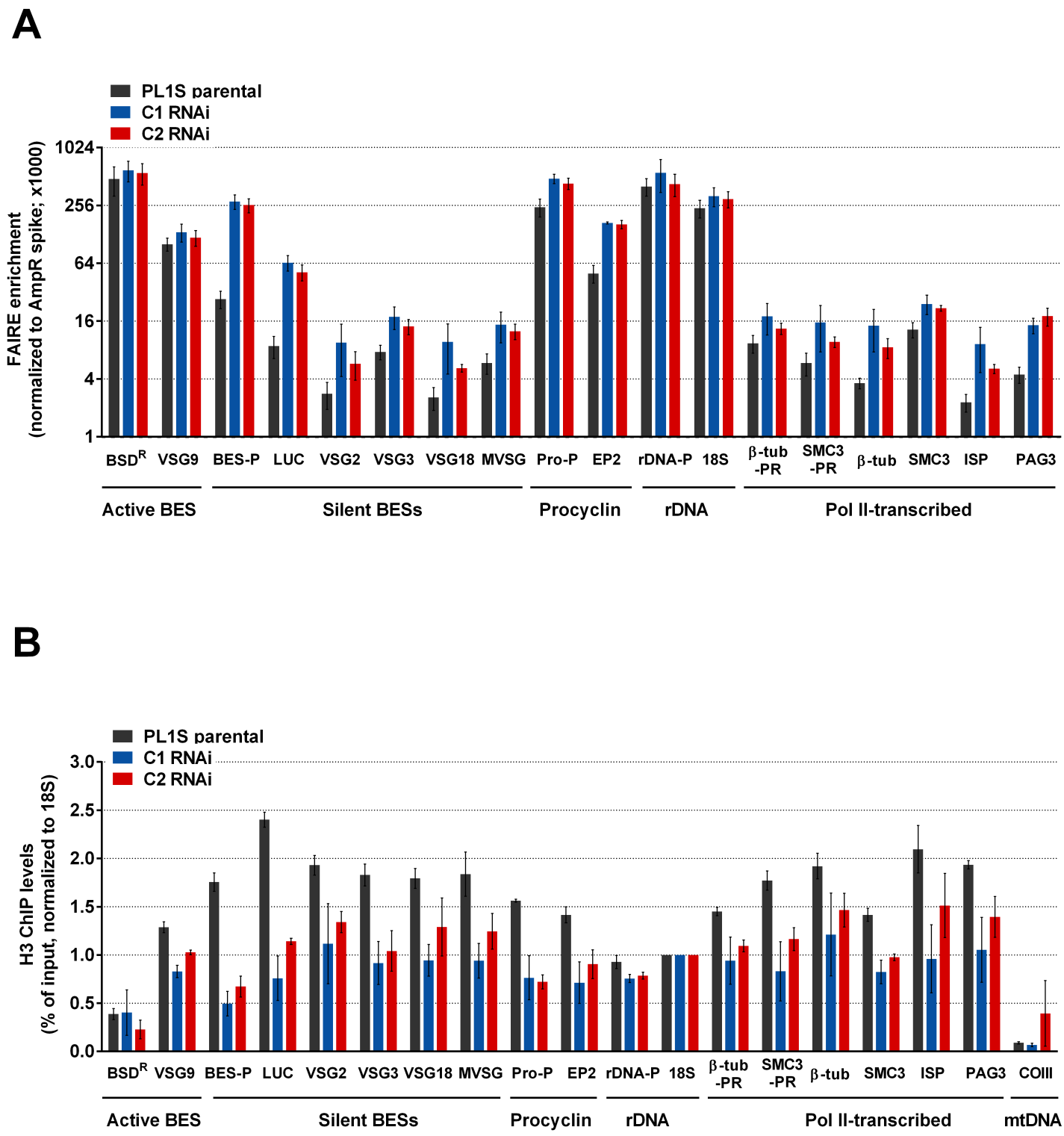

A
PL1S parental
C1 RNAi
C2 RNAi
FAIRE enrichment
(normalized to AmpR spike; x1000)
1024
256
64
16
4
1
BSD^R
VSG9
BES-P
LUC
VSG2
VSG3
VSG18
MVSG
Pro-P
EP2
rDNA-P
18S
β-tub -PR
SMC3 -PR
β-tub
SMC3
ISP
PAG3
Active BES
Silent BESs
Procyclin
rDNA
Pol II-transcribed
B
PL1S parental
C1 RNAi
C2 RNAi
H3 ChIP levels
(% of input, normalized to 18S)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
BSD^R
VSG9
BES-P
LUC
VSG2
VSG3
VSG18
MVSG
Pro-P
EP2
rDNA-P
18S
β-tub -PR
SMC3 -PR
β-tub
SMC3
ISP
PAG3
COIII
Active BES
Silent BESs
Procyclin
rDNA
Pol II-transcribed
mtDNA

Figure S6

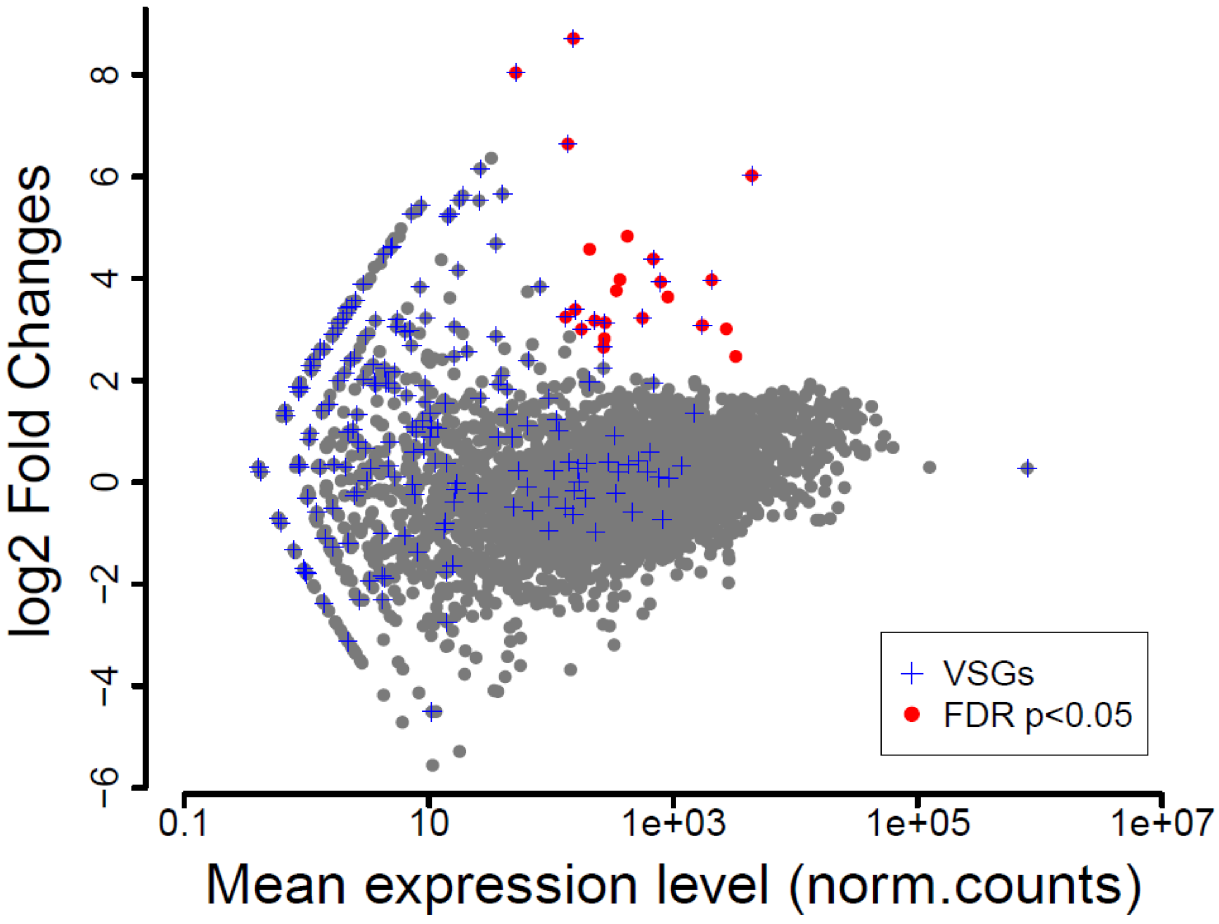
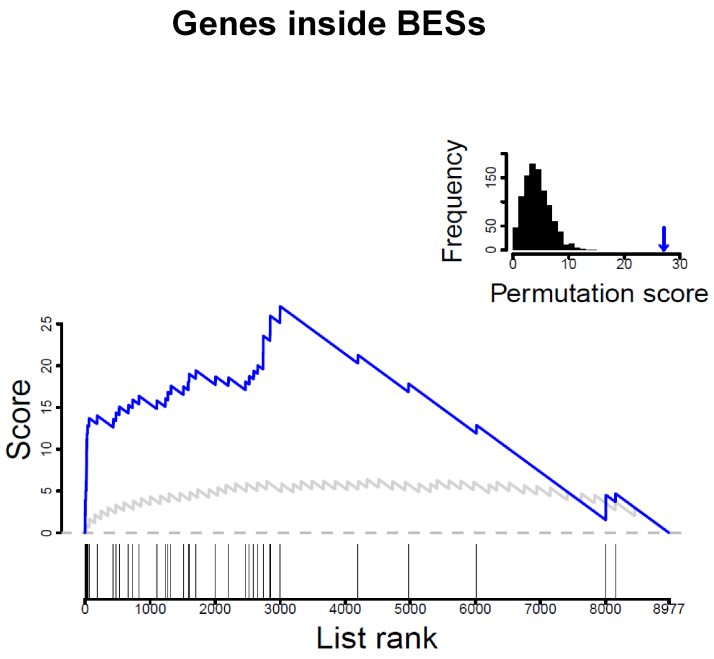
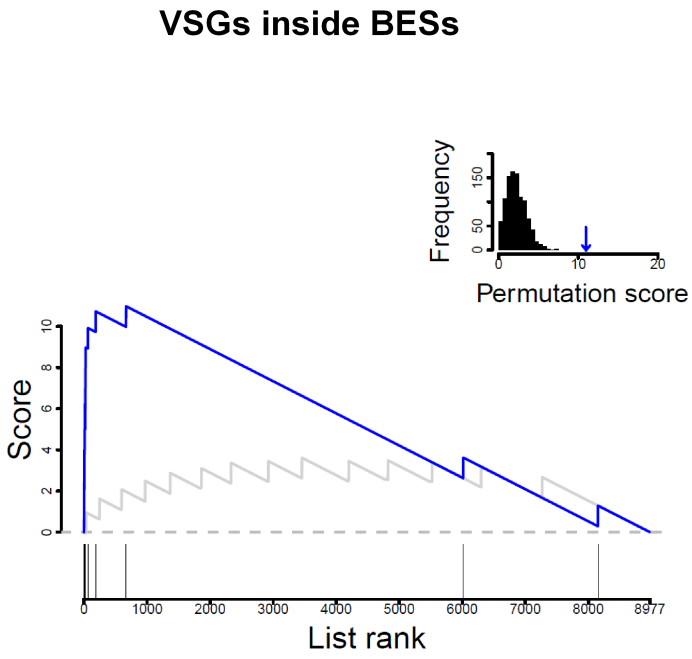
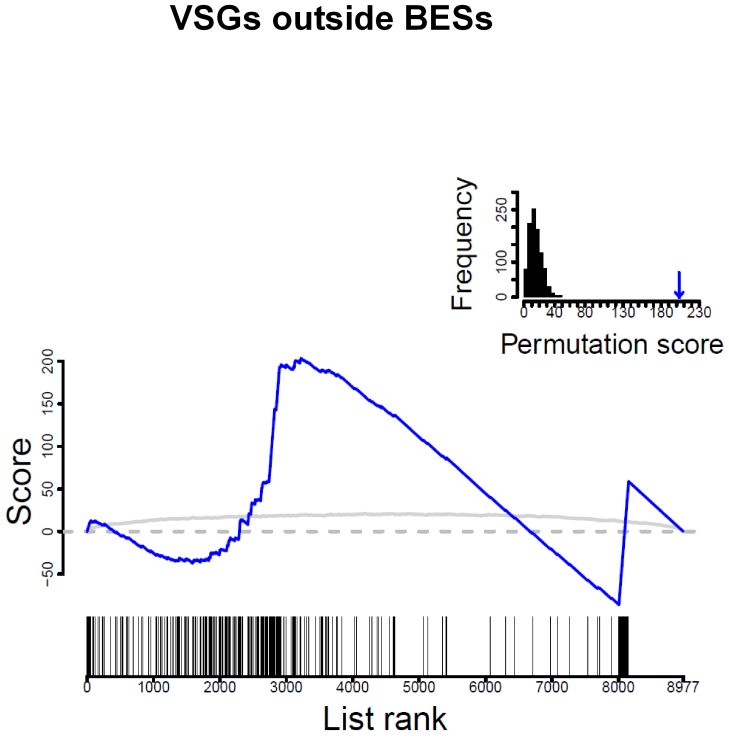

log2 Fold Changes
8
6
4
2
0
−2
−4
−6
0.1
10
1e+03
1e+05
1e+07
Mean expression level (norm.counts)
VSGs
FDR p<0.05

Figure S7


Genes inside BESs
Frequency
Permutation score
Score
List rank


VSGs inside BESs
Frequency
Permutation score
Score
List rank


VSGs outside BESs
Frequency
Permutation score
Score
List rank

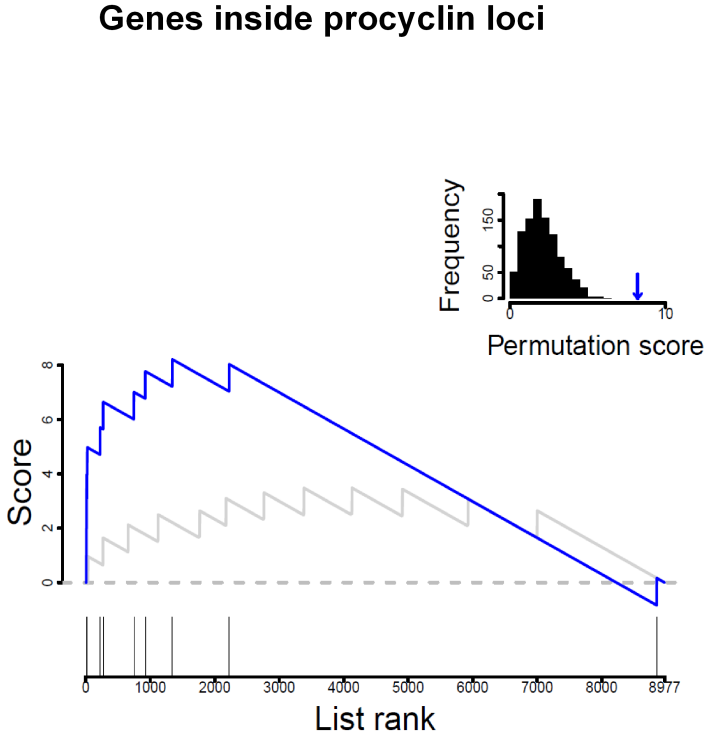

Genes inside procyclin loci
Frequency
Permutation score
Score
List rank

Figure S8

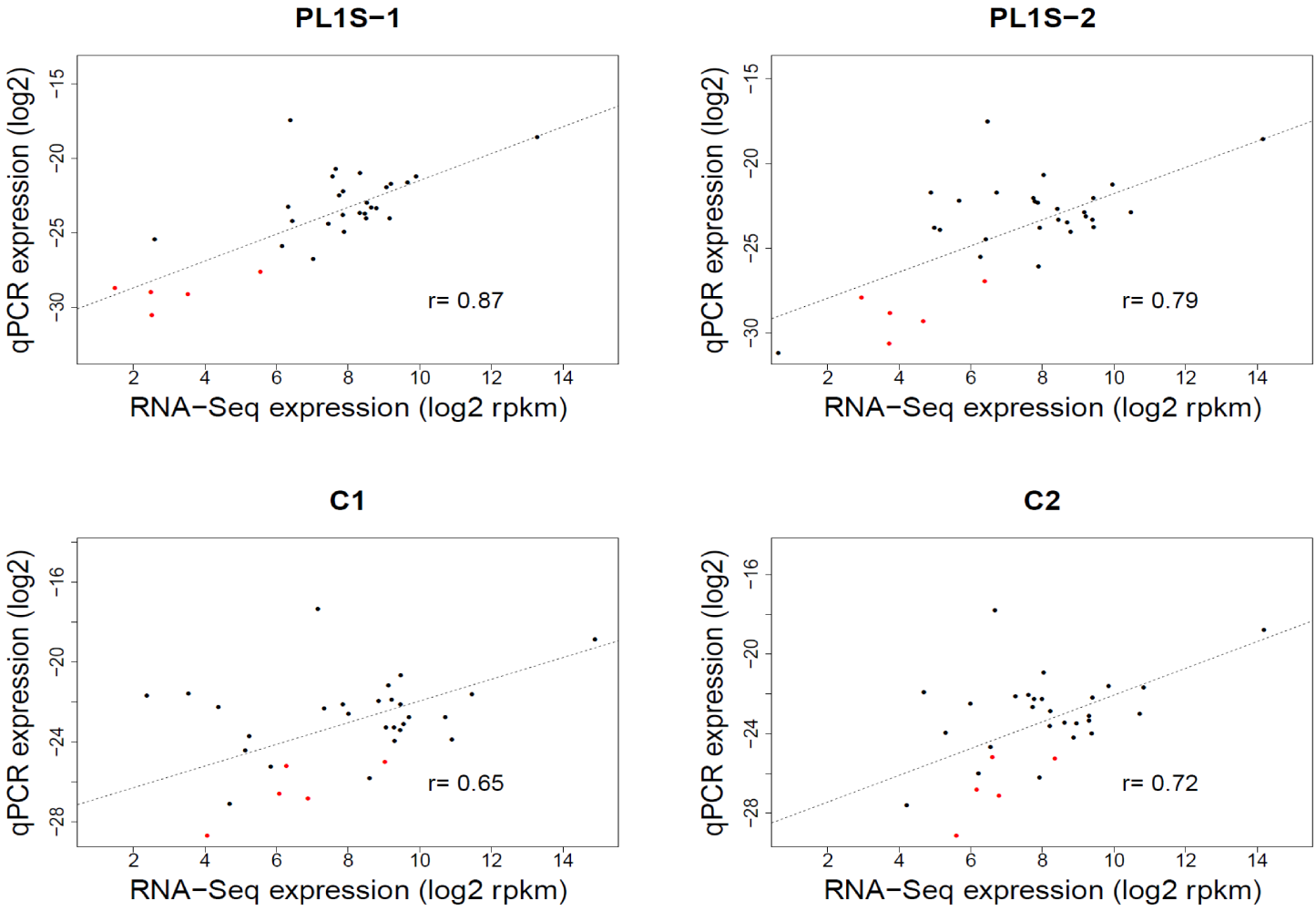

PL1S−1
qPCR expression (log2)
RNA−Seq expression (log2 rpkm)
r= 0.87
PL1S−2
qPCR expression (log2)
RNA−Seq expression (log2 rpkm)
r= 0.79
C1
qPCR expression (log2)
RNA−Seq expression (log2 rpkm)
r= 0.65
C2
qPCR expression (log2)
RNA−Seq expression (log2 rpkm)
r= 0.72

Figure S9

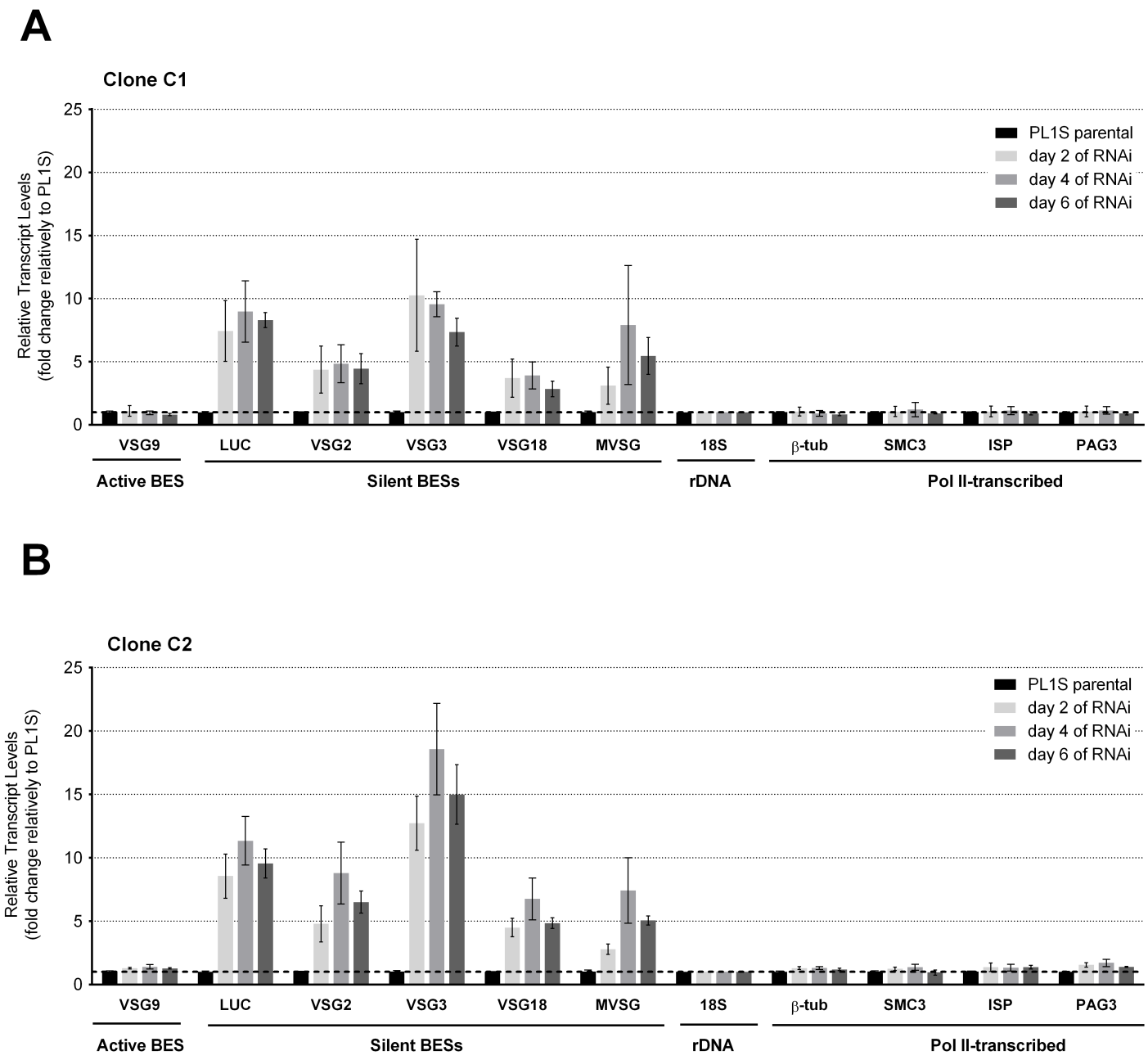

A
Clone C1
Relative Transcript Levels
(fold change relatively to PL1S)
25
20
15
10
5
0
PL1S parental
day 2 of RNAi
day 4 of RNAi
day 6 of RNAi
VSG9
LUC
VSG2
VSG3
VSG18
MVSG
18S
β-tub
SMC3
ISP
PAG3
Active BES
Silent BESs
rDNA
Pol II-transcribed
B
Clone C2
Relative Transcript Levels
(fold change relatively to PL1S)
25
20
15
10
5
0
PL1S parental
day 2 of RNAi
day 4 of RNAi
day 6 of RNAi
VSG9
LUC
VSG2
VSG3
VSG18
MVSG
18S
β-tub
SMC3
ISP
PAG3
Active BES
Silent BESs
rDNA
Pol II-transcribed

Figure S10

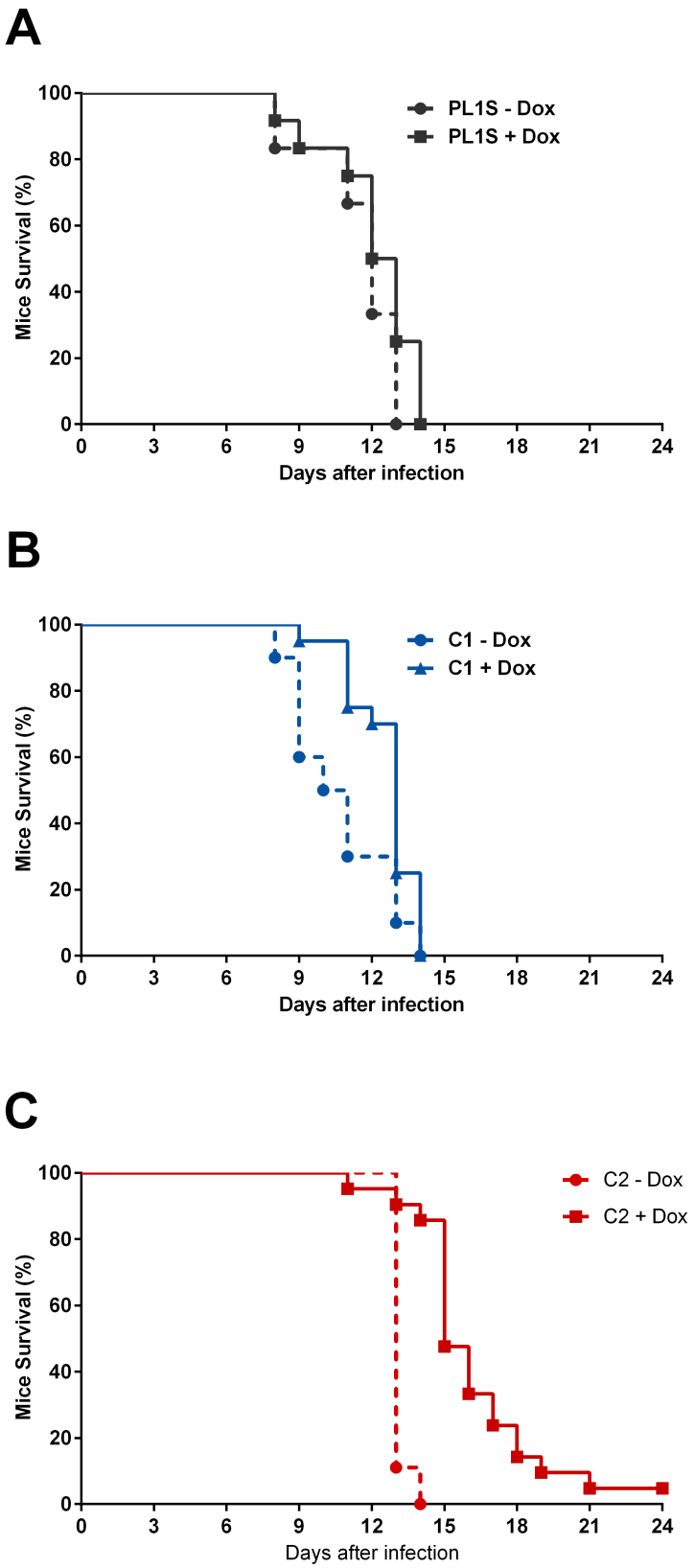

A
100
80
60
40
20
0
Mice Survival (%)
PL1S - Dox
PL1S + Dox
0
3
6
9
12
15
18
21
24
Days after infection
B
100
80
60
40
20
0
Mice Survival (%)
C1 - Dox
C1 + Dox
0
3
6
9
12
15
18
21
24
Days after infection
C
100
80
60
40
20
0
Mice Survival (%)
C2 - Dox
C2 + Dox
0
3
6
9
12
15
18
21
24
Days after infection

Figure S11

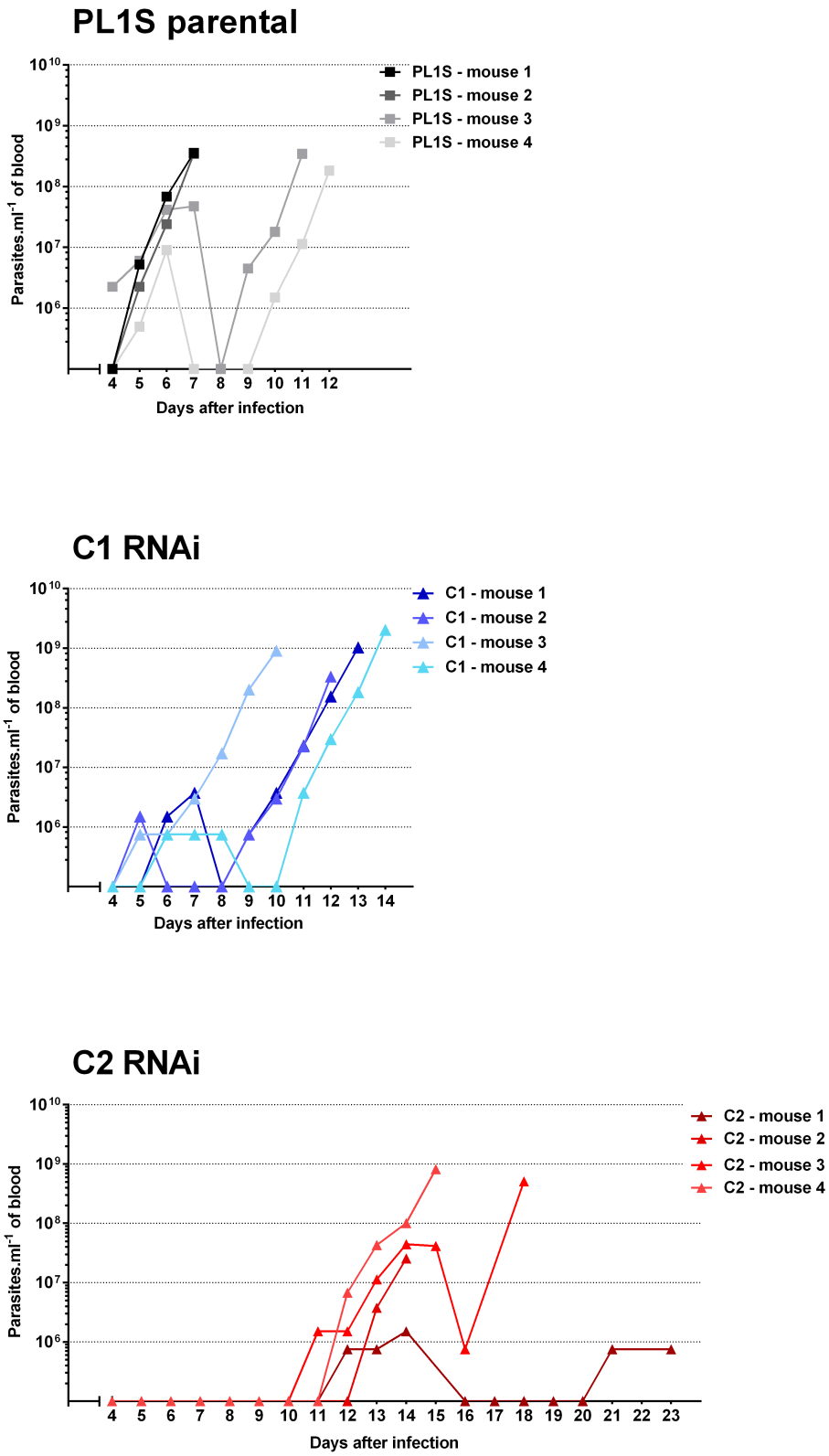

PL1S parental
PL1S - mouse 1
PL1S - mouse 2
PL1S - mouse 3
PL1S - mouse 4
Parasites.ml-1 of blood
Days after infection
C1 RNAi
C1 - mouse 1
C1 - mouse 2
C1 - mouse 3
C1 - mouse 4
Parasites.ml-1 of blood
Days after infection
C2 RNAi
C2 - mouse 1
C2 - mouse 2
C2 - mouse 3
C2 - mouse 4
Parasites.ml-1 of blood
Days after infection

Figure S1. Genotyping of histone H1 loci. (**A**) Diagram of H1 loci with hybridization probes and restriction enzyme target sequences represented. H1 loci are colored according to the code associated with the 5'UTR-specific primers and protein N-terminal sequences defined for Figure 1. (**B**) Southern blot analysis from *T. brucei* Lister427 genomic DNA. DNA was double digested with *Sma*I (S) and *Bam*HI (B) (SB lane) and single digested with *Nhe*I (N lane). DNA restriction fragments were separated by standard agarose gel electrophoresis. DNA probes were made by PCR amplification of specific gene sequences (probe A GeneID: Tb11.42.0004; probe B GeneID: Tb11.55.0002; probe C GeneID: Tb11.39.0010; probe D GeneID: Tb11.39.0006 and probe H1 GeneID: Tb11.39.0008) and ^{32}P-radiolabeled. Blots were hybridized and visualized by phosphorimaging. There were three unexpected bands (*) which suggest the existence of an additional H1 locus. (**C**) Summary table of PCR and sequencing results of each H1 loci. A putative new locus was identified by PCR amplification and sequencing, being composed by a unique histone H1 gene.

Figure S2. Each histone H1 class has a unique 5'UTR. (**A**) Alignment of the predominant 5'UTRs and (**B**) 3'UTRs of histone H1 variants of *T. brucei*. Average percentage of sequence identity between each variant is shown on the left and sequence length on the right of each sequence. Divergent nucleotides are indicated in gray. Gaps (-) were introduced for best alignment. Predominant UTRs in BSFs were derived from Siegel *et al*. (Siegel *et al.*, 2010) and Manful *et al*. (Manful *et al.*, 2011).

Figure S3. Specificity of histone H1 primers. (**A**) Class-specific amplification of histone H1 was obtained by using forward primers that annealed with the class-specific 5'UTR and a reverse primer that annealed with the conserved 3'UTR sequence common to all H1 classes. Amplification of all H1 classes was obtained by using a forward and a reverse primer that annealed both to 3'UTR sequences common to all H1 classes. (**B**) Class-specificity of the primers was confirmed by PCR using as a template plasmids containing either MAKASA, MAKTT or MNNTT gene. Primers for a given class only amplify a fragment of the corresponding class.

Figure S4. Efficiency of Histone H1 depletion. (**A**) Percentage of depletion of histone H1 classes during RNAi induction for 2, 4 and 6 days in RNAi clones C1 and C2 relatively to the parental cell-line PL1S. Primers for all H1 classes and primers specific for each H1 classes were used. Three independent experiments were analyzed. (**B**) Non-induced RNAi clones C1 and C2 show a reduction of histone H1 from 77% to 63% of PL1S levels. One and two independent experiments were analyzed for C1 and C2, respectively. Results are shown as mean ± SEM. Statistical significance was determined by a Mann-Whitney U test.

Figure S5. H1 compacts chromatin at different levels across genome. (**A**) FAIRE enrichment and (**B**) H3 ChIP levels in individual H1 RNAi clones C1 and C2 after 6 days of RNAi induction. For FAIRE an AmpR gene contained in a plasmid spike was used as a normalizer for DNA input. For H3 ChIP, DNA was quantified by qPCR, compared to the total input material and normalized to 18S DNA. Results are expressed as fold-change relatively to parental cell-line PL1S. Note that FAIRE enrichment is represented

in a logarithmic scale. Two to three independent experiments were analyzed. Results are shown in mean ± SEM. BSD^R, blasticidin resistance gene; VSG, variant surface glycoprotein gene; BES-P, promoter region of silent BESs; LUC, luciferase gene; MVSG; metacyclic VSG gene; Pro-P: procyclin promoter region; EP2: procyclin EP2 gene; rDNA-PR, rDNA promoter region; 18S, ribosomal 18S gene; β-tub, β-tubulin gene; SMC3, structural maintenance of chromosome 3 gene; β-tub-PR, promoter region of β-tub polycistronic unit; SMC3-PR, promoter region of SMC3 polycistronic unit ; ISP, inhibitor of serine peptidase gene; PAG3, procyclin associated gene 3; COIII, cytochrome c oxidase subunit III gene; mtDNA, mitochondrial DNA. Statistical significance was determined by an empirical Bayes approach.

Figure S6. Gene expression MA-plot of H1-depleted clones relative to parental clone. Mean expression levels (normalized counts) across samples are plotted on the x-axis and the corresponding log2 fold-changes between H1 RNAi and PL1S samples are plotted on the y-axis. Each point represents a gene. Genes with false discovery rate (FDR) adjusted p-values<0.05 are considered statistical significant and are indicated in red; *VSG* genes are indicated in blue.

Figure S7. H1 regulates transcription of genes located in bloodstream expression sites and procyclin loci. Gene set enrichment analysis plots showing, for the subset of genes in each plot title, the enrichment score distributions across genes ranked by decreasing statistical evidence of differential expression (blue line). Light gray line represents 95^{th} percentile of the enrichment scores obtained by 1,000 random permutations of the gene ranks. The histogram at the top-right corner of the figure represents the distribution of maximum scores obtained by random permutation, with the arrow indicating the experimental enrichment score (i.e. the “peak” in the main plot). The diagrams at the bottom show where the members of each gene set appear in the ranked list of genes.

Figure S8. Validation of RNA-Seq results by qPCR. Scatter plots of log2 qPCR levels and log2 RNA-Seq RPKMs for each sample (PL1S and H1-depleted clones, C1 and C2). 34 genes with mean expression levels ranging from 12 to 512,586 RPKMs by RNA-Seq were randomly selected and quantified by qPCR. VSG genes are indicated in red. Pearson's correlation coefficients (r) are indicated.

Figure S9. Loss of histone H1 leads to *VSG* de-repression. Quantification of gene expression 2, 4 and 6 days after RNAi induction in clone C1 (**A**) and clone C2 (**B**), relative to the parental-cell line PL1S. Silent BESs and *VSGs* become derepressed when histone H1 is depleted. Three independent experiments were analyzed. Results are shown as mean ± SEM.

Figure S10. Survival of PL1S (**A**), H1 RNAi clone C1 (**B**) and clone C2 (**C**), when RNAi was non-induced (-) or induced (+) with doxycycline. Mice n=6-21 per group. Two to four independent experiments were analyzed. Survival curves were compared by a Log-rank (Mantel-Cox) test.

Figure S11. Histone H1 depletion delays parasite appearance *in vivo*. Representative example of parasitemia (number of parasites per ml of blood) in mice infected with PL1S and H1-depleted C1 and C2 clones. Three to four independent experiments were analyzed.

Table S1. Comparison of p-values for fold-increase of H1 depleted clones relatively to PL1S in FAIRE and ChIP experiments.

Table S2. Comparison of gene expression between H1 depleted clones and parental cell-line PL1S. Genes are ordered by decreasing statistical evidence of differential expression. Genes with altered expression (FDR $p<0.05$) upon loss of histone H1 in BSF of *T. brucei* are shaded in red; in gray are non-significant genes; in blue are those genes which would be significant if a less conservative statistical method had been used (FDR>0.05 and $p<0.05$)

Table S3. Primers used for real-time qPCR analysis.

Table S1. Comparison of p-values for fold-change levels of H1-depleted clones relative to PL1S in FAIRE and ChIP experiments. Non-significant p-values (>0.05) are highlighted in blue.

	Fold relative to PL1S		P value	
Loci	**FAIRE**	**ChIP-H3**	**FAIRE**	**ChIP-H3**
BSD	1,28	0,59	7.8155e-02	4.3350e-02
VSG9	1,25	0,74	1.1154e-01	3.8466e-02
BES-P	10,42	0,38	2.3594e-45	1.2662e-08
LUC	7,08	0,47	1.1271e-32	4.3546e-07
VSG2	2,59	0,70	2.7651e-08	5.1371e-03
VSG3	2,05	0,59	8.0545e-06	3.8335e-04
VSG18	2,77	0,68	3.5050e-09	4.0048e-03
MVSG	2,30	0,63	2.3810e-07	1.6109e-03
Pro-P	1,95	0,53	3.1702e-05	6.1951e-05
EP2	3,60	0,60	2.1755e-14	1.7638e-03
rDNA-P	1,18	0,86	1.9650e-01	2.3235e-01
18S	1,29	1,00	6.1862e-02	-
β-tub-PR	1,67	0,77	1.9647e-03	3.4793e-02
SMC3-PR	2,12	0,61	1.3259e-05	7.5486e-04
β-tub	2,99	0,75	4.0836e-10	1.5674e-02
SMC3	1,79	0,70	2.2672e-04	1.2807e-02
ISP	2,98	0,62	1.9575e-10	6.7962e-04
PAG3	3,80	0,71	2.8251e-15	5.3581e-04

Table S3 – Primers used for real-time

Primer name /GeneID	Used to validate RNA-Seq	Amplified region	Forward primer (5'-3')	Reverse primer (5'-3')
All H1		All classes of H1 genes	GGAGAGAGCTGTCACACG	ATGGCGGTACATAAGACGC
BES-P		Silent BESs promoter region (97bp downstream TSS)	AAAACCCTTAGCGTTACCAC	AAACAACACACCAGCCCTC
BSD^{R}		Blasticidin resistance gene	CGGCTACAATCAACAGCATC	ACGATACAAGTCAGGTTGCC
b-tub	✓	b-tubulin gene	TTCAGGCTGGCCAATGCG	TACGGAGTCCATTGTACCTG
b-tub-PR		b-tubulin promoter region	AGCCGCTACTCGATTTCACA	TCGCCATACCGACACAGC
COIII		Cytochrome c oxidase subunit III gene	GAGGGAACGGGAGAGGAACG	TGTTTCTGCAAACAATCTCATCTGG
EP1	✓	Procyclin EP2 gene (3'UTR)	AATGTCTTATTAACCATCGCCTG	AAAATTATGGAATACGCAACCG
EP2	✓	Procyclin EP2 gene (3'UTR)	GGTGGCCCCAGTTATTTCTTT	TATGCAAGTGTCCTGTCGCC
ISP		Inhibitor of serine peptidase gene	GTCGTGTGATGGAGGATGG	TGAAGTGCTTGGTCGGAAC
LUC	✓	Luciferase reporter gene	ATGTCCGTTCGGTTGGCAG	CATACTGTTGAGCAATTCACG
MAKTT		MAKTT class H1 gene	GGAAAGTAGAAAGGAAAATAAAAT	TCTAATAGCGGGCAGCAG
MAKASA		MNNTT class H1 gene	TCGCAATCTTATCAACACTCG	TCTAATAGCGGGCAGCAG
MNNTT		MAKASA class H1 gene	CTTTATCGACTCCCCACAAG	TCTAATAGCGGGCAGCAG
MVSG		Metacyclic variant surface glycoprotein 639 gene	TCGCACTTTCAGCTCTGCTC	GCCGACCACTCGCTGTCC
PAG3		Procyclin-associated gene 3	GCAAATGCTCCTCTTCTCC	TGGCTGCCGAAGCAAGCG
Pro-P		Procyclin promoter region (10bp downstream TSS)	AGTTTAAGATGTTCTCGTGAT	CTTTTTGGTGTAATTGAAGTC
rDNA-P = 18S Int		Ribosomal DNA promoter region (15bp downstream TSS)	CTGACCGCTCTCAGACCG	ACACACGTATTACACACACTC
SMC3	✓	Structural maintenance of chromosome 3 gene	GAGGCTGGATGATGAGAGG	ATTCAGCTTCACTGATGATGG
SMC3-PR		SMC3 promoter region	GGCGGTGTGGGTGTATCC	ATCGTCGCCCCTAAGTGC
VSG2	✓	Variant surface glycoprotein 2 gene	AGCAGCCAAGAGGTAACAGC	CAACTGCAGCTTGCAAGGAA
VSG3	✓	Variant surface glycoprotein 3 gene	GCTTATTTTGTGTCTGTCGC	GACGCAGCAGAATCAACAC
VSG9	✓	Variant surface glycoprotein 9 gene	ACTAAGCTCGTGGCGCAC	CGCGTAGTTGACGCATGAC
VSG13	✓	Variant surface glycoprotein 13 gene	ATAACGCATGGCCATCTTGAC	GTCGTTGCTGTGGATTGCTC
VSG18	✓	Variant surface glycoprotein 18 gene	ACAGACCGCCGACAGTATC	GTATCTTTGTAGGCCGCTGC
VSG21	✓	Variant surface glycoprotein 21 gene	CAGCGCAAGTACAGGACG	TGCTTCGTCGTCGCTTAC
18S		18S ribosomal DNA	ACGGAATGGCACCACAAGAC	GTCCGTTGACGGAATCAACC
Tb927.10.2080	✓	hypothetical protein, conserved	AGCTGCAGCCACAGATATTC	TTGTTGCATTGGTGGCGTAC
Tb927.10.2400	✓	hypothetical protein	TTGTGGGTGCCAGAACTTTG	TGTTTCTCGTGCATCCCTTC
Tb11.02.5800	✓	calmodulin, putative	ACGCGTGAGTATTTGAAGCG	TCGTAGTCTGCTTCAGCCTTC
Tb927.7.4520	✓	hypothetical protein, conserved	ACAAGAAAGCAACGGGGAAG	TGGTGAAGCTCTGCAAGAAC
Tb09.211.2140	✓	hypothetical protein, unlikely	CGTTCCTTTTATTTCGCTCACG	AGATGATAAAGCAATGGGAGCTC
Tb09.211.4750	✓	hypothetical protein, unlikely	GGTGGCATGGTTGGGTAAAG	TTCCGGGTGTATGCAATTGC
Tb11.03.0370	✓	hypothetical protein, conserved	AGCAGCCTCTTTTGTGCTTC	CGTGATACGCAAAAGGGAAAC
Tb927.4.3430	✓	hypothetical protein, conserved	ATCCATGGCTTTCGCAAACG	ACGACACATTTCTGGAACGC
Tb09.211.3190	✓	hypothetical protein, unlikely	TGGGCGGTGATTCTCTTAAT	TTTAACCGAGGTGTAACGTGAA
Tb09.160.0380	✓	hypothetical protein, conserved	TCAGTTGTTGCGGAATGTCC	AACATCACGCACGCCATTG
Tb927.10.10760	✓	hypothetical protein	TTCCACGAGAGCGAAAAGTG	TCGAGCATTGACAAGGAACC

Tb927.10.11060	✓	hypothetical protein, conserved	TATGTTCGTGAACCGCACTG	AAAAGTCATCGTCGCGACTC
Tb09.160.2440	✓	hypothetical protein, conserved	TTTTCTTCTGGCAGCATCGC	AACCCGCGCTGTAACAATTC
Tb927.8.2470	✓	hypothetical protein, conserved	TGTTCAGCGCGAGAAAGTAC	ACTTTGGGCTATGCTCCTTCTC
Tb11.03.0935	✓	RNA polymerase subunit, putative (RPB6z)	GCACGCCAAATTGTGAATGG	ATGAATAGGATCGACGGATGCC
Tb09.211.2150	✓	Nuclear poly(A)-binding protein 1 (PABP2), putative (PABP2)	AGACATGCGGAACAAACAGC	ATGCGCGTGTTCATGAAGTG
Tb11.01.6220	✓	procyclin-associated gene 4 (PAG4) protein, putative	TGGAAACGAAAAGGGCAAGG	ATGACTTTCGTTCGCCCATG
Tb927.10.2100	✓	elongation factor 1-alpha,EF-1-alpha (TEF1)	AACATGATCACCGGCACATC	TGTTGCAGCACACAACCATC
Tb09.244.2400	✓	BARP protein (BARP)	AGTGTACTGATACAGCGGAAGG	TGCCGCAACTTTCTAGCAAC
Tb927.10.12100	✓	RNA-binding protein, putative (RBP7B)	TTCTACGGTGATGTGCTGCA	TTGTGAAGTCCGAGGATGGC
Tb927.7.5940	✓	Protein Associated with Differentiation (TbPAD2)	ATCGCAGCGTTCTCAAATGG	TTGTTGCGTGTACCACTCAC
Tb11.02.0740	✓	60S ribosomal protein L44	GCGAAGACGACCAAGAAGATTG	TTGTCGTTCAGCTCGAAGTG
Tb11.42.0004	✓	hypothetical protein, conserved	TGGCTTGGTCCGCATTTTTG	GCCACAATGCTGTTGGAATG

Cell Host & Microbe
Article

Trypanosoma brucei Parasites Occupy and Functionally Adapt to the Adipose Tissue in Mice

Sandra Trindade,[1,9] Filipa Rijo-Ferreira,[1,2,3,9] Tânia Carvalho,[1] Daniel Pinto-Neves,[1] Fabien Guegan,[1] Francisco Aresta-Branco,[1] Fabio Bento,[1] Simon A. Young,[7] Andreia Pinto,[1] Jan Van Den Abbeele,[5,6] Ruy M. Ribeiro,[4,8] Sérgio Dias,[1] Terry K. Smith,[7] and Luisa M. Figueiredo[1,*]

[1]Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, 1990-375 Lisboa, Portugal
[2]Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, TX 75390-9111, USA
[3]Graduate Program in Areas of Basic and Applied Biology, Instituto de Ciências Biomédicas Abel Salazar, Universidade do Porto, 4099-002 Porto, Portugal
[4]Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[5]Department of Biomedical Sciences, Unit of Veterinary Protozoology, Institute of Tropical Medicine Antwerp, B-2000 Antwerp, Belgium
[6]Department of Physiology, Laboratory of Zoophysiology, University of Ghent, B-9000 Ghent, Belgium
[7]Biomedical Sciences Research Complex, University of St Andrews, North Haugh, St Andrews, Fife KY16 9ST, UK
[8]Guest Professor, Faculdade de Medicina, Universidade de Lisboa, 1990-375 Lisboa, Portugal
[9]Co-first author
*Correspondence: lmf@medicina.ulisboa.pt
http://dx.doi.org/10.1016/j.chom.2016.05.002

SUMMARY

***Trypanosoma brucei* is an extracellular parasite that causes sleeping sickness. In mammalian hosts, trypanosomes are thought to exist in two major niches: early in infection, they populate the blood; later, they breach the blood-brain barrier. Working with a well-established mouse model, we discovered that adipose tissue constitutes a third major reservoir for *T. brucei*. Parasites from adipose tissue, here termed adipose tissue forms (ATFs), can replicate and were capable of infecting a naive animal. ATFs were transcriptionally distinct from bloodstream forms, and the genes upregulated included putative fatty acid β-oxidation enzymes. Consistent with this, ATFs were able to utilize exogenous myristate and form β-oxidation intermediates, suggesting that ATF parasites can use fatty acids as an external carbon source. These findings identify the adipose tissue as a niche for *T. brucei* during its mammalian life cycle and could potentially explain the weight loss associated with sleeping sickness.**

INTRODUCTION

Human African trypanosomiasis (HAT), also known as sleeping sickness, is a neglected tropical disease that is almost always fatal if left untreated. This disease is caused by *Trypanosoma brucei*, a unicellular parasite that lives in the blood, lymphatic system, and interstitial spaces of organs (reviewed in Kennedy, 2013). Disease pathology often correlates with sites of accumulation of the infectious agent within its host, including the brain, which is associated with characteristic neuropsychiatric symptoms and sleep disorder. Weight loss is another typical clinical feature of sleeping sickness pathology (Kennedy, 2013), but is essentially unstudied.

T. brucei is transmitted through the bite of a tsetse and quickly adapts to the mammalian host to become what is known as a "slender" bloodstream form (BSF). As parasitemia increases, slender forms are capable of sensing population density, and this triggers differentiation to the stumpy form, which is pre-adapted to life in the transmitting tsetse vector and, once there, further differentiates into procyclic form (PCF). Several studies have shown 10%–30% of genes being differentially expressed between BSFs and PCFs (reviewed in Siegel et al., 2011), including genes involved in metabolism, organelle activity, cell-cycle regulation, and endocytic activity. Recent proteomic studies also revealed around 33% of proteins that are developmentally regulated (Butter et al., 2013).

A major difference between BSFs and PCFs is their energy production, with the former utilizing glucose via glycolysis within the glycosome and the latter utilizing proline and, to a lesser extent, other amino acids as their carbon source, via the Krebs cycle in mitochondrion (reviewed in Szöör et al., 2014). To date, no fatty acid β-oxidation has been observed as a carbon source in any life cycle stage of this parasite. This has been a puzzling observation, as the genes required for productive β-oxidation, including the carnitine-acyltransferases (for mitochondrial import of fatty acids), are present in the genome.

Here, we describe an additional form of *T. brucei* in mammalian hosts: we demonstrate that *T. brucei* accumulates in adipose tissue, consistent with recent studies showing accumulation of parasites in the lower abdomen (Claes et al., 2009; McLatchie et al., 2013). Adipose tissue resident *T. brucei* have a different metabolic profile from either slender or stumpy forms in the blood, and this profile is consistent with their utilization of fatty acids (myristate) as a carbon source. These experiments describe an additional form of *T. brucei* life cycle and possibly explain weight loss (wasting), one of the characteristic pathological features of sleeping sickness.

RESULTS

T. brucei Parasites Are Heterogeneously Distributed in Mice

The well-established mouse model (C57BL/6J mice with a pleomorphic clone AnTat1.1^{E}) was used to confirm weight loss during infection as observed in humans with sleeping sickness. Parasitemia followed a previously described pattern: the first peak of parasitemia occurred 5–6 days post-infection, at around 2 × 10^8 parasites/mL, and after approximately 4 days of undetectable parasitemia, parasites could be detected again with a fluctuating parasitemia of 10^6–10^7 parasites/mL (Figure 1A). After the first peak of parasitemia, all infected animals showed reduced food intake and a 10%–15% decrease in body weight. Eventually, all mice recovered normal food intake, although their body weight remained 5% lower than that of non-infected animals (Figures 1B and S1A, available online). The weight of most organs from mice sacrificed on days 6 and 28 post-infection showed minimal changes relative to day 0, except fat depots, which decreased on average 43% ± 12%. Spleen size and weight increased dramatically as previously reported (Figure 1C). Infected mice died 35 ± 2.5 days post-infection (Figure 1D).

To assess the parasite load in different organs, we used immunohistochemistry at different days of infection (Figures 1E and S1B). Parasites were consistently detected in the fat 6 days post-infection and at later time points, while in other organs they were seen sporadically and at very low densities. As infection progressed, we observed an increase in parasite load in most organs, with fat, heart, brain, lung, and kidney being the most visibly infiltrated. Parasites were always found extracellularly within the interstitium of these organs. In the brain, our data corroborate the extensively reported evidence for the localization of parasites being restricted to the choroid plexuses and meninges (Kennedy, 2013) (Figure 1E).

Histologically, thymus, lymph nodes, bone marrow, skin (of the head and neck), salivary glands, spleen, gastrointestinal mucosa, testis, and liver displayed few or no parasites (Figure 1E; data not shown). Although parasites in the stroma of the testis were absent, the epidymal fat body and stroma of the epididymis (a small paired organ in the posterior end of the testis) contained a significant number of parasites, many of which appeared as debris, but which could explain the bioluminescence detected by Claes et al. (2009).

Early in Infection, *T. brucei* Accumulate in Adipose Tissue

Immunohistochemical staining showed parasites in the stroma of several fat depots: gonadal, mediastinic, mesenteric, retroperitoneal, perirenal, and interscapular (Figure 2A). Transmission electron microscopy (TEM) confirmed that these parasites were indeed extravascular, as numerous trypanosomes were observed in the interstitial space, either between adjacent adipocytes or between the adipocytes and the capillaries (Figure 2B).

To quantify parasite density, we used as a proxy Trypanosome genomic DNA (gDNA), which was quantified at 6 and 28 days post-infection in the organs/tissues where parasites had been detected by histology, i.e., fat, lung, heart, kidney, brain, and blood (Figures 2C and S2A). The blood had the highest parasite density on day 6. Among solid organs/tissues, for the same day of infection, parasite density was relatively low, except for fat, which had on average 60-fold more parasites than lung, heart, kidney, and brain and 7-fold less than blood. On day 28 of infection, parasite density remained equally high in fat (10^4–10^5 parasites/mg), while it increased, on average, 20-fold in the brain, heart, and lung (Figure 2C). The overall high parasite density was detected in all fat depots characterized in this study, with no significant differences between white and brown fat depots (Figure S2B; Supplemental Experimental Procedures). The blood was the only site where we observed a decrease in parasite density during infection, which is consistent with parasitemia dynamics (Figures 1A and 2C). As a consequence, on day 28, fat was the compartment with the highest parasite density (linear mixed-effects model [LME], $p < 0.0001$).

Overall, the density of parasites per milligram of organ/tissue correlated well with the density calculated as a ratio of parasite gDNA versus mouse gDNA in each tissue (Figure S2C). We also observed essentially the same pattern of parasite density and the same preferential accumulation in the fat when we quantified parasite RNA (qRT-PCR) instead of DNA (Figure S2D), suggesting that gDNA quantification reflects accurately the number of live parasites. Immunohistochemistry also showed accumulation of parasites in fat regardless of parasite strain (EATRO1125 AnTat1.1^{E}, Lister 427), infection route (intraperitoneal, intravascular), mouse strains (C57BL/6J, BALB/c), animal gender (male, female), or rodent species (mice, rat) (Figure S2E).

Fat represents around 14% of the body weight of a healthy mouse (Jackson Mouse Phenome Database); thus, it is potentially a very large reservoir of parasites. The number of parasites in the organs/tissues (parasite load) was determined by multiplying parasite density by the weight of the organ at the corresponding time of infection. For fat, we used the weight of the six depots characterized in this work, which comprises around 25% of the total body fat. We observed that while 6 days post-infection, the blood contained the majority of the parasites (around 10^8 parasites), on day 28 the six depots of fat contained overall more parasites than the blood, brain, and all other tested organs combined (LME, $p < 0.0001$) (Figure 2D). A similar preference for accumulation in fat was observed when the mouse infection was initiated by a tsetse bite, which deposits metacyclic forms in the skin of the mouse (Figure S3) (LME, $p < 0.0001$). Overall, these data revealed that fat represents a major reservoir of parasites, regardless of whether the infection was initiated by BSFs or metacyclic forms.

Adipose Tissue Contains Replicative and Infective Parasites

In the blood, parasites can be either replicative slender, G1-arrested stumpy forms or intermediate forms that are not fully differentiated (reviewed in MacGregor et al., 2012). To investigate whether the parasites from fat (referred to hereafter as adipose tissue forms [ATFs]) are replicative or not, we infected mice with a *GFP::PAD1*$_{utr}$ reporter cell-line, in which a *GFP* gene is followed by a *PAD1* 3′ UTR that confers maximum expression in stumpy forms (J. Sunter, A. Schwede, and M. Carrington, personal communication; MacGregor et al., 2012). Four and six days post-infection, blood and fat were collected, and parasites isolated and purified. As described by MacGregor et al. (2011), on day 4 we observed that most

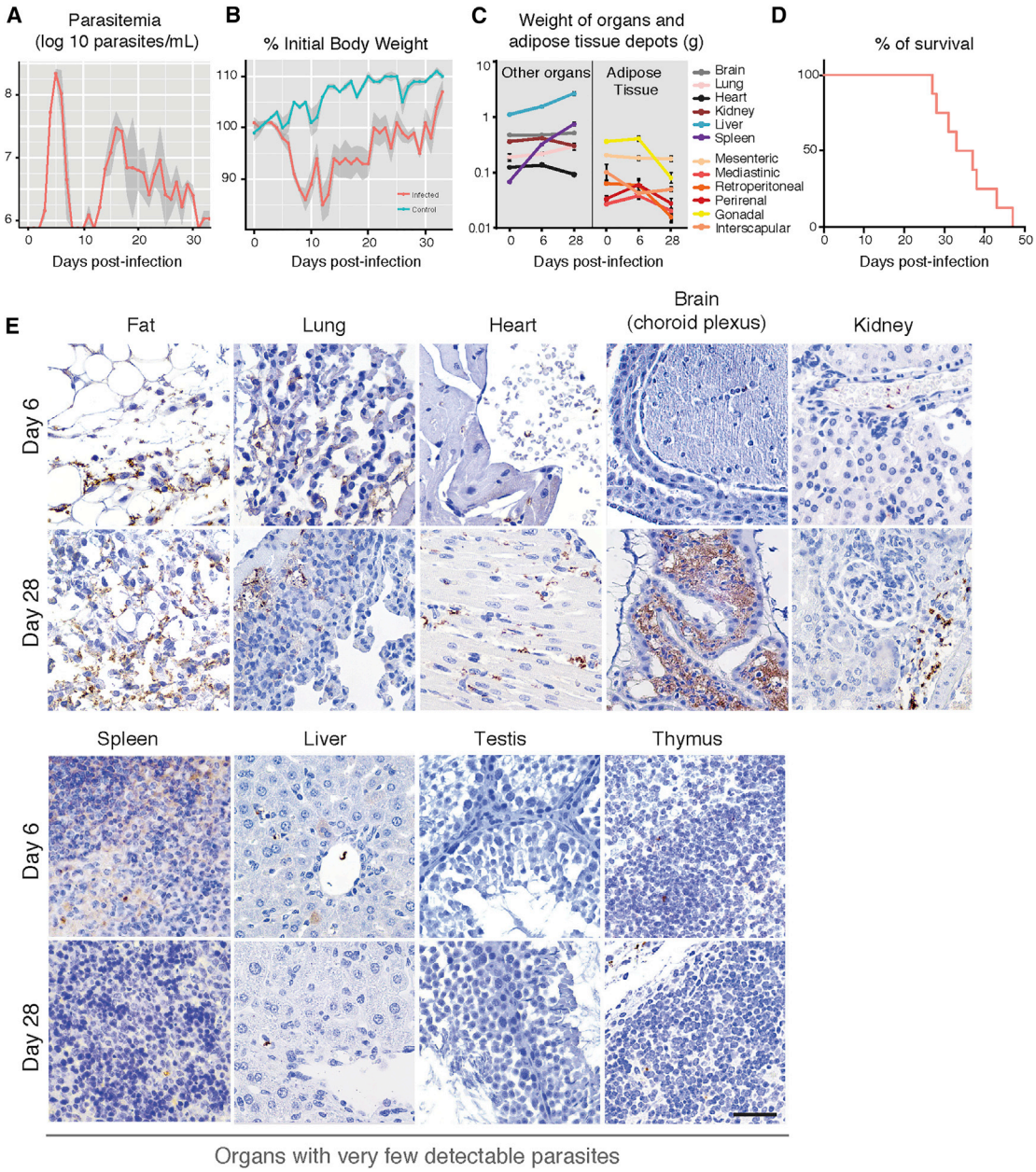


Figure 1. Tissue Distribution of *T. brucei* during a Mouse Infection Is Heterogeneous

(A) Mean parasitemia profile of 20 mice infected with *T. brucei* AnTat1.1^{E}. Parasitemia was assessed from tail blood using a hemocytometer (limit of detection is around 4 × 10^5 parasites/mL). Light gray shaded area represents SEM.

(B) Variation of body weight during infection. Daily body weight measurement of control and infected mice (n = 15 per group). Light gray shaded area represents SEM.

(C) Variation of organ weight during infection (n = 4 per group).

(D) Survival curve of *T. brucei* infected mice (n = 8).

(E) Representative brightfield micrographs of *T. brucei* distribution in several organs/tissues at days 6 and 28 post-infection, assessed by immunohistochemistry with a non-purified rabbit anti-VSG antibody (parasites appear in brown). n = 5 per time point. Scale bar, 50 μm.

See also Figures S1 and S2.

parasites in the blood were GFP negative (98% ± 0.3%), while on day 6 most parasites expressed high levels of GFP (86% ± 2.6%). Interestingly, from day 5, we noted the presence of parasites expressing lower levels of GFP, which likely correspond to differentiating intermediate forms (data not shown). As the yield of isolation of ATF parasites was very low on days 4 and 5, we analyzed these parasites only on day 6. The majority of the ATF parasites were GFP negative (79% ± 4.6%), while 21% ± 4.6% expressed GFP, indicating that on day 6 fat contains fewer stumpy/intermediate forms than blood (Figure 3A).

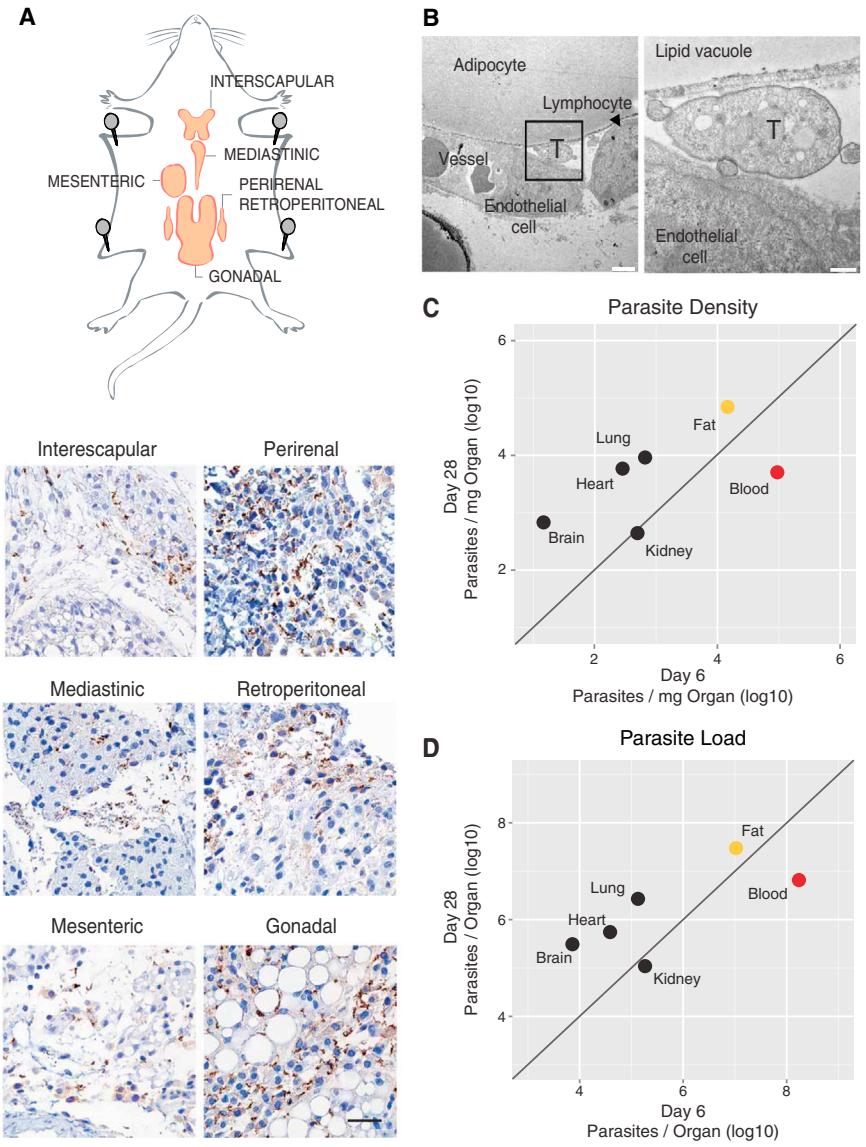


Figure 2. Fat Depots Are a Major Parasite Reservoir

(A) Schematic representation of mice fat depots and anti-VSG immunohistochemistry images of six different fat depots, collected 28 days post-infection. Scale bar, 50 μm.

(B) Transmission electron micrograph of a gonadal fat depot 6 days post-infection. Trypanosome (T) and lymphocyte in the interstitial space, adjacent to an adipocyte and next to a small capillary. Scale bars, 2 and 0.5 μm in the left and right panels, respectively.

(C) Parasite density in multiple organs/tissues (6 and 28 days post-infection) was measured by qRT-PCR of gDNA (quantification of *T. brucei* 18s rDNA relative to the tissue/organ weight). Blood density was assumed 1.05 g/mL. Fat value is the average of quantification of the six depots indicated in (A). Each point represents the geometric mean of the parasite density on days 6 (n = 3–9) and 28 post-infection (n = 3–6).

(D) Parasite load in multiple organs/tissues estimated by multiplying parasite density with organ weight at the corresponding day of infection. Each point represents the geometric mean of the parasite density on days 6 (n = 3–9) and 28 post-infection (n = 3–6).

See also Figures S1–S3.

To confirm if GFP-negative ATF parasites are replicative and GFP-positive parasites are cell-cycle arrested, we stained the parasite nuclear DNA with propidium iodide and quantified it by flow cytometry. In all samples (blood day 4, blood day 6, and fat day 6), we observed that GFP-negative parasites displayed a cell-cycle profile characteristic of replicative cells (around 60%–70% of cells in G1, 5% in S-phase and 20%–30% in G2/M), while GFP-positive cells were cell-cycle arrested in G1/G0 (90%–95%) (Figure 3B). Similar data were obtained by performing cell-cycle analysis with DyeCycle Violet (Figures S4A and S4B), further confirming the presence of slender and stumpy/intermediate forms in fat.

To validate the presence of stumpy/intermediate forms in fat, we used fluorescence microscopy on an intact gonadal depot infected with *GFP::PAD1*$_{utr}$-expressing parasites (Figure 3C). LipidTOX stains the lipids in the large lipid droplet of adipocytes. Among the adipocytes, we could clearly observe many green foci, which represent the parasite nuclei where GFP accumulates, thus confirming the presence of GFP-positive parasites (stumpy and/or intermediate forms) in close proximity to adipocytes. The presence of replicating parasites in intact tissue was confirmed by immunohistochemistry with an anti-H2A antibody. Dividing nuclei were clearly visible in close proximity to adipocytes (Figure S4C), further confirming the fluorescence-activated cell sorting (FACS) cell-cycle data.

To test whether ATF parasites are capable of establishing a new infection, an infected donor mouse was sacrificed and perfused, and several organs were collected, homogenized, and injected intraperitoneally into recipient naive mice. Parasitemia was assessed daily thereafter and scored on the first day it became detectable (Figure 3D). Mice that received blood or a fat homogenate showed parasitemia earlier (around 3 days post-transplantation) than animals injected with heart and brain homogenates (around 4 days post-transplantation), consistent with the observed parasite load in these organs (Figures 2D and 3D). Transplant of intact gonadal fat depot also led to successful infection of the recipient naive mice (data not shown), suggesting that parasites can exit from an intact tissue. These results showed that parasites from fat, heart, and brain are capable of reinvading the bloodstream and establishing a new infection.

Morphology of Adipose Tissue Forms

Although *T. brucei* is always extracellular, its morphology changes during the life cycle, which may reflect a specific adaptation to the host niche (Wheeler et al., 2013; Bargul et al., 2016).

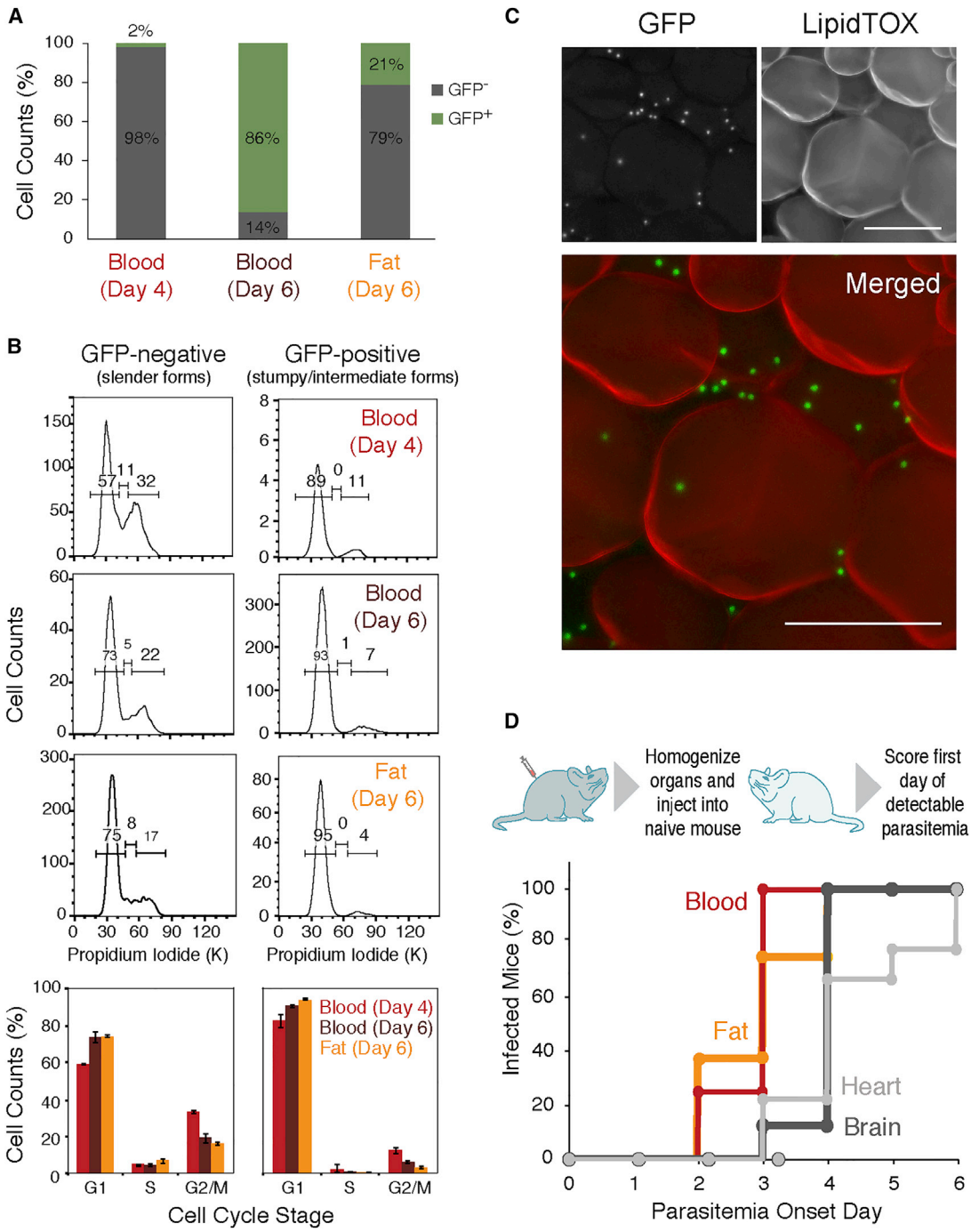


Figure 3. Fat Harbors Replicative Forms that Can Establish a New Infection

(A) Frequency of GFP expression measured by flow cytometry in parasites isolated from blood and fat, 4 and 6 days post-infection with a *GFP::PAD1*$_{utr}$ *T. brucei* reporter cell line (n = 2–3).

(B) Cell-cycle analysis assayed by flow cytometry of propidium iodide-stained parasites (n = 2–3). The values represented are the means of the percentage of the cell population in each cell-cycle stage and their SEM.

(C) Fluorescence microscopy of gonadal adipose tissue from a mouse infected for 6 days with *GFP::PAD1*$_{utr}$ reporter cell line. Lipid droplets were stained with LipidTOX (red), and nuclei of GFP-expressing parasites (stumpy and/or intermediate forms) are green. Scale bar, 50 μm.

(D) Onset of parasitemia curves in mice that were injected intraperitoneally with infected organs/tissues lysates from a donor mouse. Lysates from blood, heart, brain, and gonadal fat depot were prepared from mice sacrificed between 21 and 28 days post-infection to ensure presence of a larger number of parasites (n = 9).

See also Figure S4.

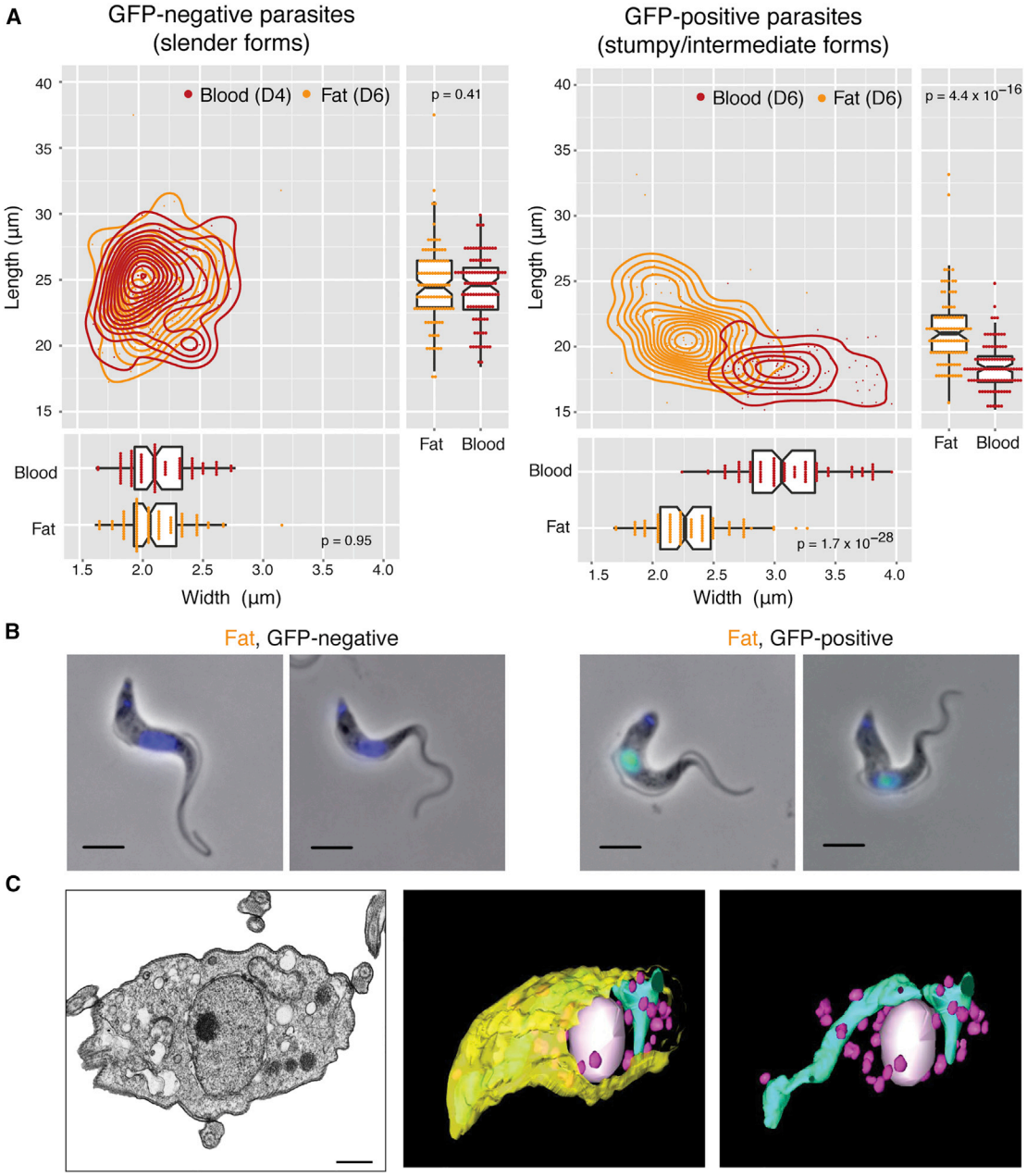


Figure 4. Fat Is Populated by Slender, Intermediate, and Stumpy Forms

(A) Morphological features (length and width) of fixed parasites isolated from fat and blood of mice infected with *GFP::PAD1*$_{utr}$ reporter. Fat gonadal tissue was collected on day 6 post-infection. The blood "controls" were obtained as follows: GFP-negative parasites were collected on day 4 post-infection (mostly slender forms), and GFP-positive parasites were collected on day 6 post-infection (mostly stumpy forms). Morphometric measurements were scored from phase contrast microscopy images, analyzed via HTIAoT, and confirmed by manual measurement. GFP negative, slender form; GFP positive, stumpy and intermediate forms. n = 100 per group, from three independent mouse experiments. Statistical significance was assessed using a Wilcoxon rank-sum test.

(B) Representative images of parasites isolated from fat. Replicating parasites (such as the second from the left) were excluded from morphometric analysis. DNA was stained with DAPI (blue). GFP protein (green) is localized in the nucleus of intermediate and stumpy forms. Scale bar, 4 µm.

(C) Transmission electron micrograph and 3D tomography images of a parasite isolated from gonadal adipose tissue. Mitochondrion is represented in cyan, glycosomes in pink, nucleus in white, and plasma membrane in yellow. Scale bar, 500 nm.

See also Figure S5 and Movie S1.

ATF parasites, like BSFs, have an undulating appearance, with a flagellum attached to the cell body and with kinetoplast DNA positioned between nucleus and flagellar pocket (Figure 4B).

To characterize in more detail the morphology of ATF parasites, we compared the length and width of *GFP::PAD1*$_{utr}$ parasites isolated from adipose tissue (day 6 post-infection) and

blood (days 4 and 6 post-infection) (Figures 4A and 4B). Automatic measurements of phase contrast microscopy images were generated via HTIAoT (Wheeler et al., 2012) and confirmed with manual measurements (Figure S5A). We observed that slender forms (GFP negative) from blood and adipose tissue were very similar both in length (blood, 24.39 ± 2.50 μm; fat, 24.57 ± 2.99 μm) and width (blood, 2.15 ± 0.26 μm; fat, 2.12 ± 0.26 μm). This average length is consistent with previous reports (Tyler, 1998; Tyler et al., 2001; Bargul et al., 2016). In blood day 6 post-infection, GFP-positive parasites were, as expected, shorter (18.43 ± 1.81 μm) and wider (3.11 ± 0.38 μm) than slender counterparts of day 4, corresponding to the morphology of stumpy forms (Tyler et al., 2001). Interestingly, in adipose tissue we found not only stumpy forms, but also GFP-positive parasites that morphologically were in between slender and stumpy forms (length, 21.32 ± 2.73 μm; width, 2.29 ± 0.31 μm) (Tyler, 1998). These probably correspond to the previously described blood intermediate forms, which, as the name suggests, are not fully differentiated into stumpy forms, but could already express GFP::PAD1 (MacGregor et al., 2011, 2012). These results indicate that adipose tissue is populated by parasites whose morphology has been previously found in the blood. The only significant difference is their relative distribution: on day 6 of infection, while blood is mostly populated by stumpy forms, adipose tissue appears to be "delayed," as we detected both intermediate and stumpy forms.

At the ultrastructure level, ATF parasites contain all major structures described in other stages of life cycle (Gull, 1999), including an electron-dense coat, nucleus, mitochondrion, endoplasmic reticulum, Golgi apparatus, glycosomes, dense granules and numerous vesicles compatible with endosomes, an internal subpellicular corset of microtubules underneath plasma membrane, and a flagellum attached to the cell body (Figures 4C and S5B). Using serial 3D tomography, we observed that the single mitochondrion of ATF parasites occupies a small volume of parasite body and is not highly branched (Figure 4C; Movie S1). This organization was confirmed by Mitotracker Green staining, which showed no major differences between the mitochondrion of parasites in blood and adipose tissue (Figure S5C; Movie S1).

Transcriptome of ATF Parasites Reveals Differences in Several Key Regulatory Processes

During its life cycle, *T. brucei* adapts to its environment by changing gene expression (reviewed in Siegel et al., 2011). To test whether parasites within fat also adapted to the new environment, total RNA was extracted from infected gonadal fat depot (n = 3) on day 6, along with parasites from blood (n = 2) on day 4 (maximizing slender and minimizing stumpy/intermediate forms), and was subjected to RNA sequencing (RNA-seq) analysis. As expected, sequence reads from blood samples corresponded mainly to parasite transcripts, while sequence reads from fat corresponded mainly to host transcripts. Nevertheless, the 1%–9% of the sequence reads from *T. brucei* provided enough statistical power to detect changes in the transcriptomes of ATF parasites (Table S1). Two previously published RNA-seq datasets of BSF parasites grown in culture were also included in this analysis (Pena et al., 2014).

Unbiased clustering of gene expression profiles revealed that ATF parasites clustered separately from parasites isolated from blood or culture (Figure 5A), suggesting significant changes in their overall transcriptome. Changes were identified using three methods of differential expression analysis, and only those genes identified by at least two methods with an adjusted p value <0.01 were considered. These analyses showed that 2,328 genes (around 20% of transcriptome) were differentially expressed between BSF and ATF parasites: 1,160 were upregulated in ATF parasites and 1,178 were upregulated in BSFs (Figures 5B and 5C; Table S2).

Significant changes were found in genes involved in gene expression regulation, cell cycle, and cell signaling (Table S3). RNA-binding proteins play an important role in gene expression and differentiation throughout the *T. brucei* life cycle. For example, RNA-binding protein 42 (RBP42; TriTrypDB: Tb927.6.4440, http://tritrypdb.org/tritrypdb/) binds many mRNAs involved in cellular energy metabolism (Das et al., 2012). Upregulation of RBP42 in ATF parasites could be involved in the metabolic rewiring when parasites enter the fat (Table S3).

ATF parasite transcriptome also showed dramatic changes in gene expression of various post- and co-translational modifying enzymes that might have considerable influence on diverse cellular processes (Table S3). A small number of genes potentially acting in the cell cycle and cytokinesis was identified with significant differential expression, including the cytoskeleton-associated AIR9 protein and spastin, which were upregulated (Table S3), suggesting differences in cell-cycle regulation in these parasites. Consistent with a minor stumpy form population in fat, we did not find enrichment of stumpy-specific genes in the transcriptome of ATF parasites. Extracellular signaling mechanisms also seem to be affected in ATF parasites, including upregulation of TOR3, which can relate the supply of external nutrients to internal energy levels to regulate cellular growth (de Jesus et al., 2010).

Interestingly, although by TEM an electron-dense coat can be observed around the parasite (Figure 2B), we found that the transcript levels of the active variant surface glycoprotein (VSG) (VSG AnTat1.1, CAA25971.1) are 3-fold downregulated in adipose tissue, suggesting VSG downregulation or VSG switching within the tissue. As the VSGnome of AnTat1.1E clone is currently unknown, we could not test whether silent VSGs were upregulated as a compensatory mechanism. Genes encoding for other surface molecules, such as the haptoglobin receptor and most procyclins, were not differentially expressed.

Transcriptome of ATF Parasites Reveals Metabolic Adaptations

One of the most evident changes in ATF transcriptome was the upregulation of many metabolic pathways, including glycolysis, pentose phosphate, purine salvage, sterol and lipid metabolism, and, surprisingly, β-oxidation. Thirteen of the 14 enzymatic steps of glycolysis were upregulated relative to BSFs (Table S3). This may either be a response to the lower glucose concentration in fat interstitial fluid relative to bloodstream, or an upregulated gluconeogenesis, which relies mostly on the same enzymes.

In ATF parasites, genes involved in three out of the five biosynthetic steps in the pentose-phosphate pathway were upregulated, including the rate-limiting glucose-6-phosphate

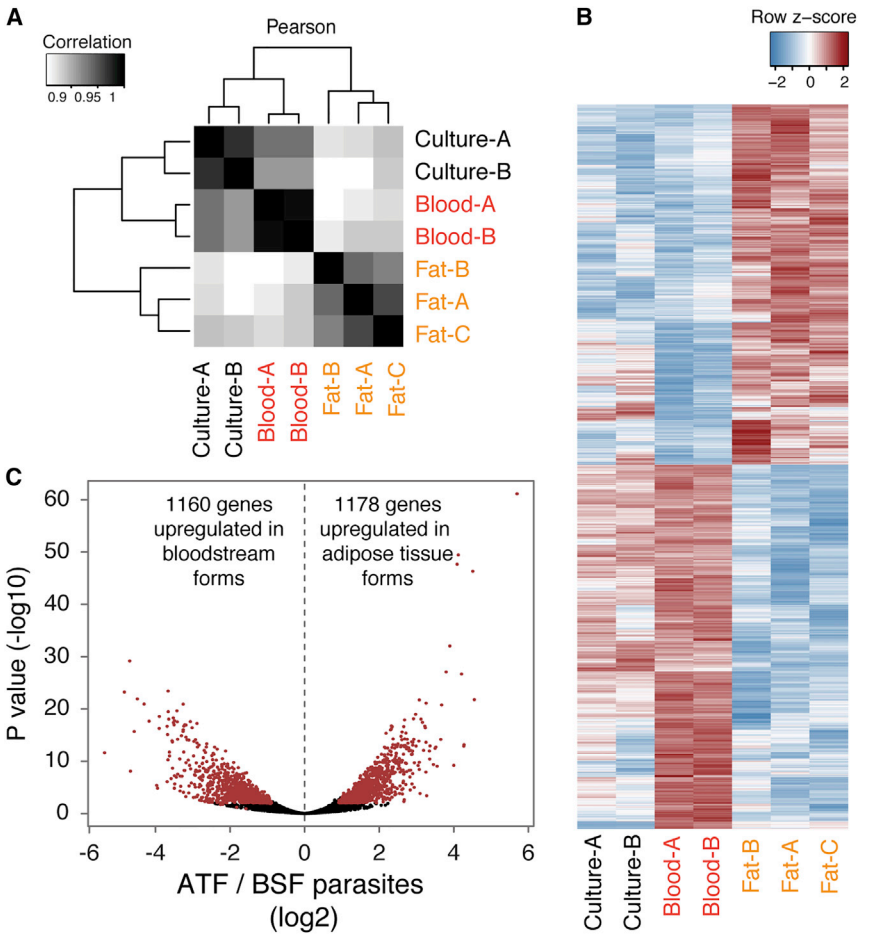


Figure 5. ATF Parasites Are Transcriptionally Different from BSFs

(A) Hierarchically clustered heat map of Pearson correlations of transcript levels (log_2 transformed RPKM) from independent RNA-seq datasets: Lister427 parasites grown in culture (Pena et al., 2014) (n = 2), parasites isolated from blood of AnTat1.1-infected mice on day 4 post-infection (n = 2), and parasites isolated from gonadal fat on day 6 post-infection (n = 3).

(B) Heat map view of relative transcript levels for differentially expressed genes from culture and in vivo in parasites isolated from the two tissues (adjusted $p < 0.01$ in at least two of three methods).

(C) Volcano plot displaying in red the differentially expressed genes represented in (B). Displayed p values and fold changes are from DESeq2.

See also Tables S1 and S2.

dehydrogenase (TriTrypDB: Tb927.10.2490). This observation, taken together with the fact that numerous enzymes (16 in total) involved in purine salvage pathway were also upregulated, suggests that ATF parasites may increase purine production. Interestingly, the purine phosphatases (TriTrypDB: Tb927.8.3800 and Tb927.7.1930) and cAMP phosphodiesterase PDEA (TriTrypDB: Tb927.10.13000) are up- and downregulated, respectively, suggesting that the increased purine production may be directed toward cAMP signaling (Table S3).

ATF parasites showed significant upregulation of the alanine and aspartate aminotransferases and the glutamate shunt, which feed products into the tricarboxylic acid (TCA) cycle. Additionally, this cycle also appeared to be more active, given the upregulation in three key steps, allowing it to process succinate, fumarate, and 2-oxoglutarate, resembling the TCA cycle of the *T. brucei* insect form (reviewed in Szöör et al., 2014). These changes suggest that the F0/F1 ATP synthase complex is functional and that the associated electron transport chain is operating in ATF parasites in a manner similar to that in PCF parasites.

Significant changes in gene expression of lipid and sterol metabolic pathways were also observed in ATF parasites. However, one of the most striking observations in the RNA-seq data was the potential presence of an active fatty acid β-oxidation, which produces energy from fatty acid catabolism. This was unexpected, as β-oxidation activity has never been detected in any *T. brucei* life cycle stage to date. ATF parasites showed upregulation of the putative genes responsible for the second and fourth steps of the β-oxidation cycle (enoyl-CoA hydratase and 3-ketoacyl-CoA thiolase, respectively) (Figure 6A). Moreover, fatty acid transport across the mitochondrial membrane (facilitated by acyl-CoA synthases and carnitine-acyltransferases) was upregulated, while fatty acid elongases 2 and 4 were downregulated, suggesting that in ATF parasites, endocytosed fatty acids are not being elongated and anabolized into more complex molecules. Instead, they may enter the glycosomes and/or mitochondrion, where they are catabolized via a β-oxidation pathway to form acetyl-CoA (experimentally validated; see below and Figure 6), which feeds into the now-active TCA cycle.

ATF Parasites Have Active Fatty Acid β-Oxidation

To investigate whether ATF parasites are capable of β-oxidation, labeled myristate was used in a pulse-chase experiment with living trypanosomes, and potential labeled β-oxidation intermediates were identified by gas chromatography-mass spectrometry (GC-MS). Myristate (C14:0) was chosen, as it is efficiently taken up and incorporated into lipids and glycosylphosphatidylinositol (GPI) anchors in the slender BSF parasites (reviewed in van Hellemond and Tielens, 2006). Isolated ATF parasites were labeled with deuterated myristate (D_{27}-C14:0) for 1 hr and then chased with serum, following which labeled myristate metabolites were identified. As expected, BSFs showed accumulation of D_{27}-C14:0 during the pulse and chase periods (Figures 6B and 6C). ATF parasites also showed D_{27}-C14:0 accumulation during the pulse (Figure 6B, upper panel, and Figure 6C), but the amount of D_{27}-C14:0 decreased significantly during the chase (Figure 6B, lower panel, and Figure 6C). Importantly, the decrease of

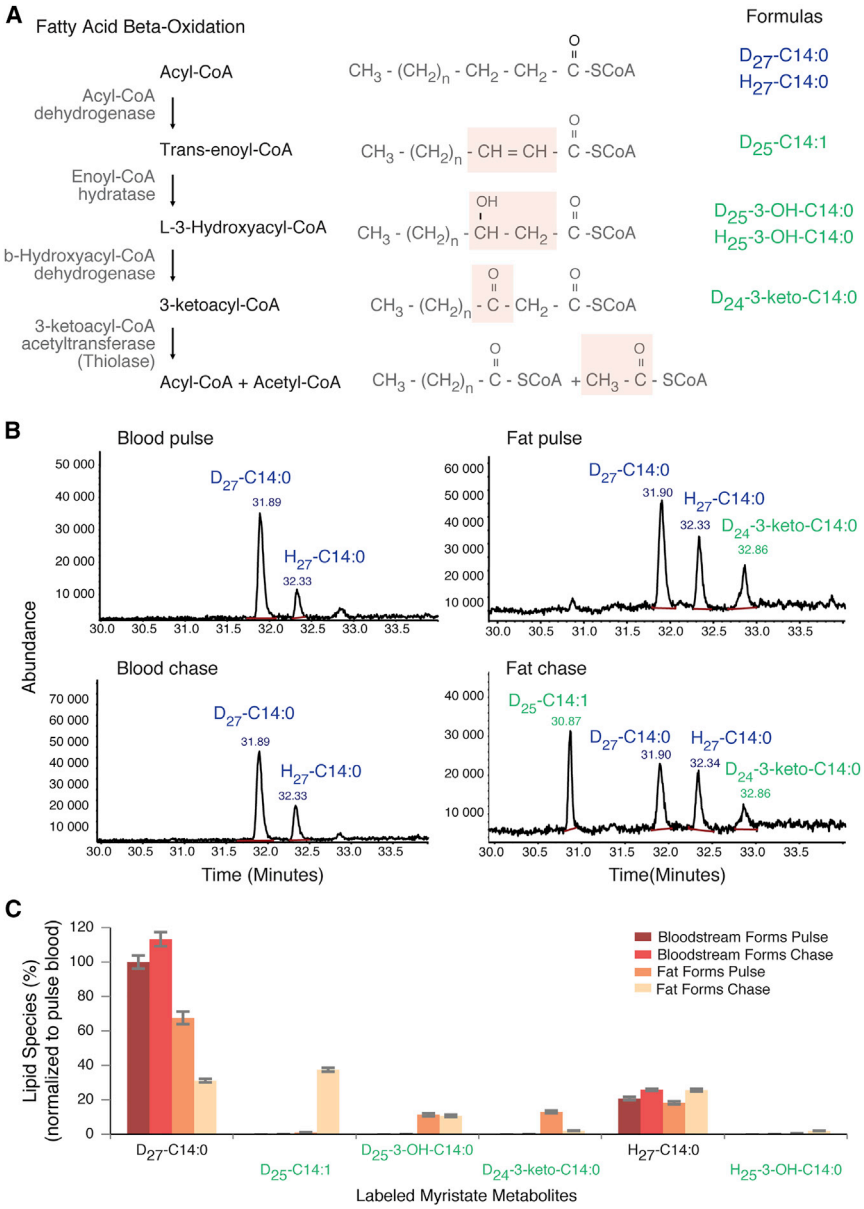


Figure 6. Fatty Acid β-Oxidation Is Active in ATF Parasites

(A) Schematic of fatty acid β-oxidation pathway. Four enzymatic modifications are indicated by shaded box on the fatty acid structures where biotransformation takes place. Formulas in blue and green indicate the myristate and β-oxidation metabolites from the non-labeled and labeled myristate, respectively, identified in this work.

(B) Fatty acid methyl ester (FAME) analysis by GC-MS of D_{27}-C14:0-labeled BSF (left) or ATF (right) parasites for 1 hr (upper) and chased for a further 1 hr (lower). GC-MS trace shows 30–34 min (n = 3).

(C) Uptake of D_{27}-C14:0 and β-oxidation metabolites after normalization to the added internal standard C17:0. 100% equates to the amount of D_{27}-C14:0 taken up by bloodstream form in the 1 hr labeling (pulse) (n = 3). The values represented are the means and the respective SEM.

See also Figures S6 and S7.

DISCUSSION

A well-established feature of the unicellular, extracellular *T. brucei* parasite is its ability to invade the CNS. Here, we show that while blood is the major site of parasite accumulation on the first peak of parasitemia, fat contains the highest density and total number of parasites later in infection (around 100- to 800-fold more than the brain). Although the reason(s) why parasites accumulate in adipose tissue remain unknown, we clearly show it has dramatic consequences for the parasites. They functionally adapt to the tissue environment by rewiring gene expression, including the possibility of using lipids/fatty acids as a carbon source.

Possible Advantages to Parasite Accumulation in the Adipose Tissue

Accumulation in the adipose tissue could be due to several non-mutually exclusive reasons that may have provided a selective advantage during evolution: parasites may be less efficiently eliminated by adipose tissue-specific immune response, parasites may grow at a faster rate, parasite differentiation may be delayed, and/or parasite entry in adipose tissue may be more efficient in adipose tissue than in other organs/tissues. Depending on the dynamics of parasite movement to/from adipose tissue, it is possible that fat acts as a source of parasites that can repopulate the blood. This reversible movement between blood and fat could have important implications for (1) the transmission dynamics, since stumpy formation is triggered by a quorum-sensing mechanism (MacGregor et al., 2011), and (2) antigenic variation, if fat, for example, would favor the appearance of new VSG variants that could later go to the blood (Mugnier et al., 2015).

An intriguing question is whether stumpy forms could be directly ingested by a tsetse fly from the subcutaneous fat.

D_{27}-C14:0 in ATF parasites coincided with the detection of β-oxidation metabolites derived from the labeled myristate, including myristoleic acid (D_{25}-C14:1), 3-hydroxy-myristate (D_{25}-3-OH-C14:0), and 3-oxo-myristic acid (D_{24}-3-keto-C14:0) (Figure 6B, right panels, and Figures 6C, S6, and S7). The latter two metabolites were also observed to some minor extent during the pulse (Figure 6B, upper panel, and Figure 6C), while D_{25}-C14:1 was present in higher amount during the chase period. Minor amounts of unlabeled 3-hydroxy-myristrate (H_{26}-3-OH-C14:0) were also observed in ATF parasites, but not BSFs (Figure 6C).

Collectively these data show that ATF parasites are able to actively take up exogenous myristate and form β-oxidation intermediates, demonstrating the existence of this pathway in trypanosomes and suggesting that ATF parasites could in part use fatty acid β-oxidation to satisfy their energy requirements.

Although in our histological analysis we did not find a significant number of parasites in this fat depot, it is possible we missed a preferential skin location. Moreover, we performed this analysis in mice infected by intraperitoneal injection, which bypasses the skin as the first entry point of metacyclic forms. So it remains to be determined whether subcutaneous fat is important for accumulation of metacyclics and/or ATF parasites and how it impacts transmission.

Not all *Trypanosoma* species occupy the same niche in the host. *T. brucei* and *T. evansi* are mainly tissue-invading parasites, while *T. congolense* stays in smaller capillaries and venules of tissues and *T. vivax* remains mainly in circulation (Losos and Ikede, 1972). These differences have been associated with the different swimming properties of each *Trypanosoma* species (Bargul et al., 2016). Previous reports had indirectly suggested that *T. brucei* parasites could be present in adipose tissue (Fernandes et al., 1997; Giroud et al., 2009). Our study demonstrates that *T. brucei* parasites accumulate in high numbers in the fat of rodents. Although mouse is an accepted model to study *T. brucei* infection (Giroud et al., 2009), we cannot exclude the possibility that accumulation in adipose tissue is a result of the selection process that happens when *T. brucei* infects a non-natural host. In the future, it would be interesting to confirm whether fat preference is a common feature of this and other *Trypanosoma* species in their natural hosts and to compare their swimming properties in different tissues/organs.

It is intriguing to note that several pathogens infect adipose tissue. *T. cruzi*, the causative agent of Chagas disease in Latin America, invades adipocytes during acute infections in mice and humans (Ferreira et al., 2011). Also, *Plasmodium berghei*, a causative agent of rodent malaria, sequesters in lungs and fat (Franke-Fayard et al., 2005). *Mycobacterium tuberculosis* infects adipocytes, where it accumulates in intracytoplasmic lipid inclusions and survives in a ''dormant'' non-replicating state that is insensitive to anti-mycobacterial drugs (Neyrolles et al., 2006). HIV takes advantage of the fat as a viral reservoir during the chronic stage of infection, and persistence on this reservoir is an obstacle for treatment (Chun et al., 2015). It is possible that persistence of *T. brucei* in the fat may also account for some of the treatment failures in humans (Richardson et al., 2016).

Functional Adaptation to Host Adipose Tissue

A major observation in this work is that 20% of the genes are differentially expressed between ATFs and BSFs, which is comparable with the differences between BSFs and PCFs (around 30%) and between slender and stumpy BSF forms (around 12%) (reviewed in Siegel et al., 2011). Parasites adapt to the fat environment by changing transcript levels of genes involved in metabolism, signaling, cell-cycle control, and RNA binding. Using biochemical assays, we confirmed that ATF can utilize fatty acids, i.e., myristate, and catabolize them via β-oxidation, which could lead to the production of ATP via the TCA cycle and oxidative phosphorylation. Therefore, it seems that parasites can sense and adapt to the adipose environment by rewiring their gene expression, including the ability to use lipid/fatty acid as a carbon source.

The major carbon source of BSF and PCF parasites is glucose and proline, respectively, both of which are readily available nutrients in the host (reviewed in Szöör et al., 2014). Fat is, in its essence, a lipid-rich environment. Therefore, it is possible that fatty acids or some other form of lipid are released from the host adipocytes and are endocytosed or actively transported via a receptor-mediated process by ATF parasites (Vassella et al., 2000). So far, only one receptor has been identified in *T. brucei* as necessary for the import of LDL particles (Coppens et al., 1987). Its transcript levels are not altered in ATF parasites, suggesting either that this protein can be upregulated post- or co-translationally or that lipid/fatty acid import is mediated by yet-uncharacterized transporters.

Consumption/utilization of host's lipids during a *T. brucei* infection could contribute to the weight loss observed in patients with sleeping sickness, cattle with Nagana, and mice infected with *T. brucei* (Kennedy, 2013; Ranjithkumar et al., 2013). Interestingly, obese mice ($db^{-/-}$ knockout mice) infected with *T. brucei* live 3-fold longer than their littermates, suggesting that having more adipose tissue partially protected mice from a *T. brucei* infection (Amole et al., 1985). Because obesity is associated with persistent low-grade chronic inflammation in adipose tissue (Ouchi et al., 2011), it is possible that in obese mice, parasites get more efficiently eliminated (or controlled), thus prolonging the survival of the host.

Most of what is known today about the mechanisms of virulence, persistence, and transmission of *T. brucei* results from studies performed in BSF parasites. The identification of adipose tissue as an additional major reservoir of functionally differentiated *T. brucei* brings a unique perspective to our understanding of this parasite and raises several questions. What is the relative contribution of BSF and ATF parasites for pathogenicity and host metabolic alterations? Could fat act as a source of parasites expressing novel VSGs? What are the implications of such a large reservoir of ATF parasites in terms of transmission? What are the dynamics of parasite entry and exit from fat? Given that the brain is a lipid-rich organ, which is also invaded by *T. brucei*, it is obvious to ask whether these parasites also adapt their gene expression and how this impacts brain-associated pathology. Do ATF parasites induce changes in the host metabolism, providing an advantage to the parasitic infection? Is the immune response of the adipose tissue more permissive to *T. brucei* parasites? Are anti-trypanosome drugs equally efficient at eliminating ATFs and BSFs? In sum, our findings have important consequences for the understanding of parasite biology, disease, and drug treatment efficacy.

EXPERIMENTAL PROCEDURES

Detailed experimental procedures can be found in Supplemental Experimental Procedures.

Animal Experiments

Animal experiments were performed according to EU regulations and approved by the Animal Ethics Committee of Instituto de Medicina Molecular (AEC_2011_006_LF_TBrucei_IMM). Tsetse fly infections were performed in compliance with the regulations for biosafety and animal ethics (VPU2014_1) and under approval from the environmental administration of the Flemish government. Unless otherwise indicated, all infections were performed in wild-type male C57BL/6J mice, 6–10 weeks old (Charles River, France), by intraperitoneal injection of 2,000 *T. brucei* AnTat 1.1^{E} 90-13 parasites. For parasite counts, blood samples were taken daily from the tail vein. Organs/tissues of infected mice were collected at days 6 and 28 post-infection unless otherwise stated. Animals were sacrificed by CO_2 narcosis, blood collected

by heart puncture, and mice immediately perfused. Collected organs were snap frozen in liquid nitrogen or fixed in 10% neutral-buffered formalin. In transplants, homogenates as well as 600 μL of blood were transplanted into age- and sex-matched naive mice.

Histology and Electron and Fluorescence Microscopy

Formalin-fixed organs were immunostained with a non-purified rabbit serum anti-*T. brucei* VSG13 antigen (which crossreacts with many VSGs) and a non-purified rabbit serum anti-*T. brucei* H2A. For TEM, ultra-thin sections (70 nm) were screened in a Hitachi H-7650 microscope. 3D reconstruction of isolated trypanosome was done using the IMOD software package version 4.7.3 for alignment and modeling (Kremer et al., 1996).

For fluorescence analysis, the gonadal depot was stained with LipidTox, fixed in 10% neutral-buffered formalin and embedded in Fluoromount-G. Fluorescence images were taken using a 40× objective in a Zeiss Cell Observer wide-field microscope. For morphometry analysis, isolated parasites were fixed with paraformaldehyde, DAPI stained, and embedded in vectashield. Images were taken using a 63× oil objective with optional optovar magnification (1.6×) in the same wide-field microscope. Parasite measurements were taken essentially as described in Wheeler et al. (2012).

The mitochondrion of isolated parasites was labeled using MitoTracker Green (Invitrogen/Molecular Probes, M-7514) according to the manufacturer's instructions. Fluorescence and DIC images were acquired using a confocal laser point-scanning microscope (Zeiss LSM 710).

Parasite Quantification

T. brucei 18S rDNA genes were amplified from gDNA of a known mass of tissue and converted into parasite number using a standard curve. For RNA quantification, the ΔΔCt method was used by amplifying *TbZFP3* and mouse *Gapdh* genes from tissue total RNA.

Parasite Isolation from Tissues

Bloodstream parasites were purified over a DEAE column (Taylor et al., 1974), while ATF parasites were isolated from gonadal fat depot by incubating the depot in MEM or HMI11 at 37°C and 150 rpm agitation for up to 40 min.

Flow Cytometry

Cell-cycle analysis was performed using propidium-iodide (PI) or Vybrant DyeCycle violet (DCV) in fixed or live cells, respectively. PI staining was done according to Aresta-Branco et al. (2016). For DCV staining, cell suspensions were washed, and 0.5 μL DCV was added per each million isolated parasites and incubated for 10 min at 37°C. PI, DCV, and GFP intensities were measured with BD LSRFortessa cell analyzer.

RNA-Seq

RNA and cDNA library of both blood and gonadal fat depot from days 4 and 6 of infection, respectively, were prepared as described (Pena et al., 2014), and samples sequenced in an Illumina HiSeq2000 platform. Reads were processed and mapped to the *T. brucei* TREU927 genome. Differential gene expression was analyzed, and genes were considered differentially expressed if they were detected by at least two of the three considered algorithms (p adjusted < 0.01).

Myristate Metabolic Labeling

To evaluate myristate incorporation and metabolism, the fat isolation protocol was performed in lipid-free minimum essential medium (MEM). Parasites were placed in a vented tube with 1 mL MEM and starved for 30 min at 37°C. Starved parasites were then labeled with 0.4 mg of radiolabeled D_{27}-C14:0 precoupled with defatted BSA for 1 hr. A total of 450 μL of the cell suspension was washed, snap frozen in liquid nitrogen, and lyophilized in glass vials (pulse sample). The remaining parasites were re-suspended in 500 μL MEM and 100 μL HMI11 for 1 hr at 37°C, and at the end processed as for pulse sample (chase sample). Metabolite extraction, identification, and quantification were conducted as described in Oyola et al. (2012), with the exception that fatty acids were released by acid hydrolysis (200 μL 6M HCl at 110°C for 16 hr).

Statistical Analysis

Statistical analyses were performed by fitting LME models with mice as random effects unless otherwise indicated. At least three independent experiments were considered in each case and statistical significance was set to $\alpha = 0.05$ level. Data were analyzed after logarithm transformation.

ACCESSION NUMBERS

The accession number for Lister427 culture parasites is ArrayExpress: E-MTAB-1715. Sequence data generated as part of this study have been submitted to the ArrayExpress database (EMBL-EBI) under accession number ArrayExpress: E-MTAB-4061.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, three tables, and one movie and can be found with this article online at http://dx.doi.org/10.1016/j.chom.2016.05.002.

AUTHOR CONTRIBUTIONS

S.T., F.R.-F., T.K.S., and L.M.F. designed the experiments and wrote the paper; S.T., F.R.-F., T.C., F.G., F.A.-B., F.B., A.P., and T.K.S. conducted the experiments; D.P.-N. and S.A.Y. analyzed the RNA-seq data; J.V.D.A. designed and conducted experiments with tsetse flies; R.M.R. conducted the statistical analysis; and S.D. designed experiments.

ACKNOWLEDGMENTS

The authors thank Keith Matthews (University of Edinburgh) for providing AnTat1.1^E clone, Christian Janzen (University of Wurzburg) for the *GFP::PAD1$_{utr}$* cell line, Dave Barry (University of Glasgow) for critical reading of the manuscript, Catarina Gadelha (University of Nottingham) for guidance in electron microscopy studies, Rita Ventura (ITQB) for support with lyophilization of samples, João Rodrigues (iMM) and Eugénia Carvalho (University of Coimbra) for useful discussions of the data, Leonor Pinho and Margarida Vaz for help with mouse work, and Fernando Augusto for drawing the graphical abstract. The authors would like to acknowledge the Rodent Facility, the Histology and Comparative Pathology and Bioimaging Laboratories of the Instituto de Medicina Molecular, and all members of the L.M.F. lab. This work was supported by 55007419 (HHMI) and 2151 (EMBO) to L.M.F., D.P.-N., F.B., and F.G.; FCT fellowships to S.T., F.R.-F., and F.A.-B. (SFRH/BPD/89833/2012, SFRH/BD/51286/2010, and SFRH/BD/80718/2011, respectively); Wellcome Trust grant (093228), MRC MR/M020118/1, and European Community Seventh Framework Programme under grant agreement No. 602773 (Project KINDRED) to S.A.Y. and T.K.S.; and PAI 7/41 (Belspo) and ERC-NANOSYM to J.V.D.A.

Received: December 10, 2015
Revised: March 21, 2016
Accepted: April 30, 2016
Published: May 26, 2016

REFERENCES

Amole, B.O., Wittner, M., Hewlett, D., and Tanowitz, H.B. (1985). *Trypanosoma brucei*: infection in murine diabetes. Exp. Parasitol. *60*, 342–347.

Aresta-Branco, F., Pimenta, S., and Figueiredo, L.M. (2016). A transcription-independent epigenetic mechanism is associated with antigenic switching in *Trypanosoma brucei*. Nucleic Acids Res. *44*, 3131–3146.

Bargul, J.L., Jung, J., McOdimba, F.A., Omogo, C.O., Adung'a, V.O., Krüger, T., Masiga, D.K., and Engstler, M. (2016). Species-specific adaptations of *Trypanosome* morphology and motility to the mammalian host. PLoS Pathog. *12*, e1005448.

Butter, F., Bucerius, F., Michel, M., Cicova, Z., Mann, M., and Janzen, C.J. (2013). Comparative proteomics of two life cycle stages of stable isotope-labeled *Trypanosoma brucei* reveals novel components of the parasite's host adaptation machinery. Mol. Cell. Proteomics *12*, 172–179.

Chun, T.W., Moir, S., and Fauci, A.S. (2015). HIV reservoirs as obstacles and opportunities for an HIV cure. Nat. Immunol. *16*, 584–589.

Claes, F., Vodnala, S.K., van Reet, N., Boucher, N., Lunden-Miguel, H., Baltz, T., Goddeeris, B.M., Büscher, P., and Rottenberg, M.E. (2009). Bioluminescent imaging of *Trypanosoma brucei* shows preferential testis dissemination which may hamper drug efficacy in sleeping sickness. PLoS Negl. Trop. Dis. *3*, e486.

Coppens, I., Opperdoes, F.R., Courtoy, P.J., and Baudhuin, P. (1987). Receptor-mediated endocytosis in the bloodstream form of *Trypanosoma brucei*. J. Protozool. *34*, 465–473.

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

de Jesus, T.C., Tonelli, R.R., Nardelli, S.C., da Silva Augusto, L., Motta, M.C., Girard-Dias, W., Miranda, K., Ulrich, P., Jimenez, V., Barquilla, A., et al. (2010). Target of rapamycin (TOR)-like 1 kinase is involved in the control of polyphosphate levels and acidocalcisome maintenance in *Trypanosoma brucei*. J. Biol. Chem. *285*, 24131–24140.

Fernandes, J.H., Atouguia, J.M., Peleteiro, M.C., Jennings, F.W., and Rosário, V.E. (1997). Post-treatment hind-leg paralysis in mice infected with *Trypanosoma brucei brucei:* a light microscopic study. Acta Trop. *63*, 179–184.

Ferreira, A.V., Segatto, M., Menezes, Z., Macedo, A.M., Gelape, C., de Oliveira Andrade, L., Nagajyothi, F., Scherer, P.E., Teixeira, M.M., and Tanowitz, H.B. (2011). Evidence for Trypanosoma cruzi in adipose tissue in human chronic Chagas disease. Microbes Infect. *13*, 1002–1005.

Franke-Fayard, B., Janse, C.J., Cunha-Rodrigues, M., Ramesar, J., Büscher, P., Que, I., Löwik, C., Voshol, P.J., den Boer, M.A., van Duinen, S.G., et al. (2005). Murine malaria parasite sequestration: CD36 is the major receptor, but cerebral pathology is unlinked to sequestration. Proc. Natl. Acad. Sci. USA *102*, 11468–11473.

Giroud, C., Ottones, F., Coustou, V., Dacheux, D., Biteau, N., Miezan, B., Van Reet, N., Carrington, M., Doua, F., and Baltz, T. (2009). Murine models for *Trypanosoma brucei* gambiense disease progression—from silent to chronic infections and early brain tropism. PLoS Negl. Trop. Dis. *3*, e509.

Gull, K. (1999). The cytoskeleton of trypanosomatid parasites. Annu. Rev. Microbiol. *53*, 629–655.

Kennedy, P.G. (2013). Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). Lancet Neurol. *12*, 186–194.

Kremer, J.R., Mastronarde, D.N., and McIntosh, J.R. (1996). Computer visualization of three-dimensional image data using IMOD. J. Struct. Biol. *116*, 71–76.

Losos, G.J., and Ikede, B.O. (1972). Review of pathology of diseases in domestic and laboratory animals caused by *Trypanosoma congolense, T. vivax, T. brucei, T. rhodesiense* and *T. gambiense*. In Veterinary Pathology, D.C. Dodd, ed. (S. Karger), pp. 1–71.

MacGregor, P., Savill, N.J., Hall, D., and Matthews, K.R. (2011). Transmission stages dominate trypanosome within-host dynamics during chronic infections. Cell Host Microbe *9*, 310–318.

MacGregor, P., Szöőr, B., Savill, N.J., and Matthews, K.R. (2012). Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act. Nat. Rev. Microbiol. *10*, 431–438.

McLatchie, A.P., Burrell-Saward, H., Myburgh, E., Lewis, M.D., Ward, T.H., Mottram, J.C., Croft, S.L., Kelly, J.M., and Taylor, M.C. (2013). Highly sensitive in vivo imaging of *Trypanosoma brucei* expressing "red-shifted" luciferase. PLoS Negl. Trop. Dis. *7*, e2571.

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.

Neyrolles, O., Hernández-Pando, R., Pietri-Rouxel, F., Fornès, P., Tailleux, L., Barrios Payán, J.A., Pivert, E., Bordat, Y., Aguilar, D., Prévost, M.C., et al. (2006). Is adipose tissue a place for Mycobacterium tuberculosis persistence? PLoS ONE *1*, e43.

Ouchi, N., Parker, J.L., Lugus, J.J., and Walsh, K. (2011). Adipokines in inflammation and metabolic disease. Nat. Rev. Immunol. *11*, 85–97.

Oyola, S.O., Evans, K.J., Smith, T.K., Smith, B.A., Hilley, J.D., Mottram, J.C., Kaye, P.M., and Smith, D.F. (2012). Functional analysis of *Leishmania* cyclopropane fatty acid synthetase. PLoS ONE *7*, e51300.

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. Mol. Microbiol. *93*, 645–663.

Ranjithkumar, M., Malik, T.A., Saxena, A., Dan, A., Sakthivel, P.C., and Dey, S. (2013). Hyperlipidaemia in trypanosomiasis of naturally infected horses: possible cachexia-anorexia syndrome? Trop. Anim. Health Prod. *45*, 417–421.

Richardson, J.B., Evans, B., Pyana, P.P., Van Reet, N., Sistrom, M., Büscher, P., Aksoy, S., and Caccone, A. (2016). Whole genome sequencing shows sleeping sickness relapse is due to parasite regrowth and not reinfection. Evol. Appl. *9*, 381–393.

Siegel, T.N., Gunasekera, K., Cross, G.A., and Ochsenreiter, T. (2011). Gene expression in *Trypanosoma brucei*: lessons from high-throughput RNA sequencing. Trends Parasitol. *27*, 434–441.

Szöőr, B., Haanstra, J.R., Gualdrón-López, M., and Michels, P.A. (2014). Evolution, dynamics and specialized functions of glycosomes in metabolism and development of trypanosomatids. Curr. Opin. Microbiol. *22*, 79–87.

Taylor, A.E., Lanham, S.M., and Williams, J.E. (1974). Influence of methods of preparation on the infectivity, agglutination, activity, and ultrastructure of bloodstream trypanosomes. Exp. Parasitol. *35*, 196–208.

Tyler, K. (1998). Differentiation and division of *Trypanosoma brucei* in the mammalian bloodstream. PhD thesis (Manchester: University of Manchester).

Tyler, K.M., Matthews, K.R., and Gull, K. (2001). Anisomorphic cell division by African trypanosomes. Protist *152*, 367–378.

van Hellemond, J.J., and Tielens, A.G. (2006). Adaptations in the lipid metabolism of the protozoan parasite *Trypanosoma brucei*. FEBS Lett. *580*, 5552–5558.

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

Wheeler, R.J., Gull, K., and Gluenz, E. (2012). Detailed interrogation of trypanosome cell biology via differential organelle staining and automated image analysis. BMC Biol. *10*, 1.

Wheeler, R.J., Gluenz, E., and Gull, K. (2013). The limits on trypanosomatid morphological diversity. PLoS ONE *8*, e79581.

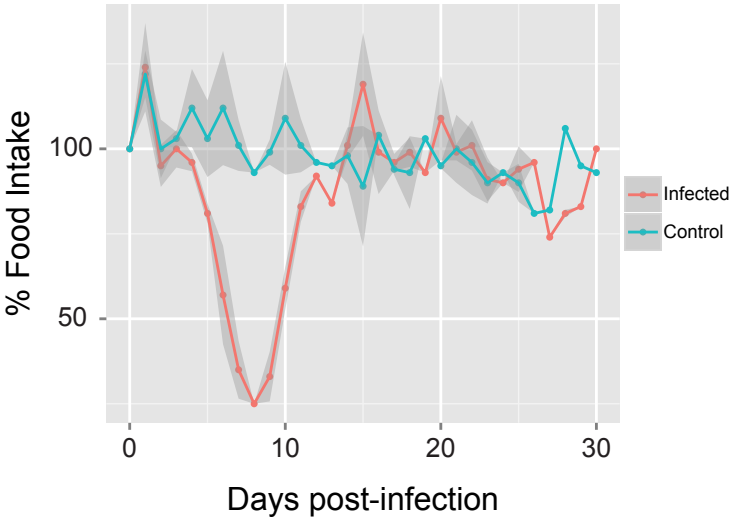
Cell Host & Microbe, Volume *19*

Supplemental Information

Trypanosoma brucei Parasites Occupy and Functionally Adapt to the Adipose Tissue in Mice

Sandra Trindade, Filipa Rijo-Ferreira, Tânia Carvalho, Daniel Pinto-Neves, Fabien Guegan, Francisco Aresta-Branco, Fabio Bento, Simon A. Young, Andreia Pinto, Jan Van Den Abbeele, Ruy M. Ribeiro, Sérgio Dias, Terry K. Smith, and Luisa M. Figueiredo

A



B

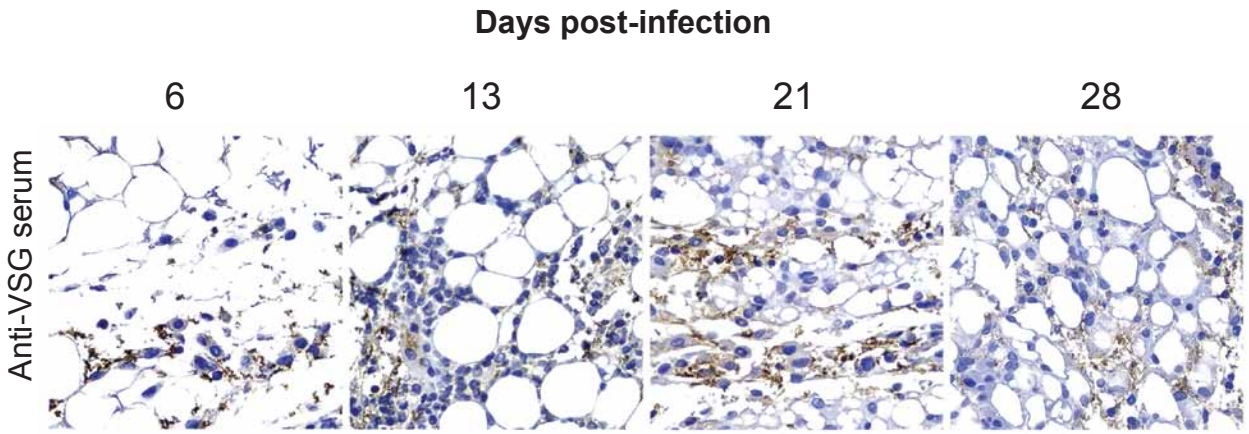
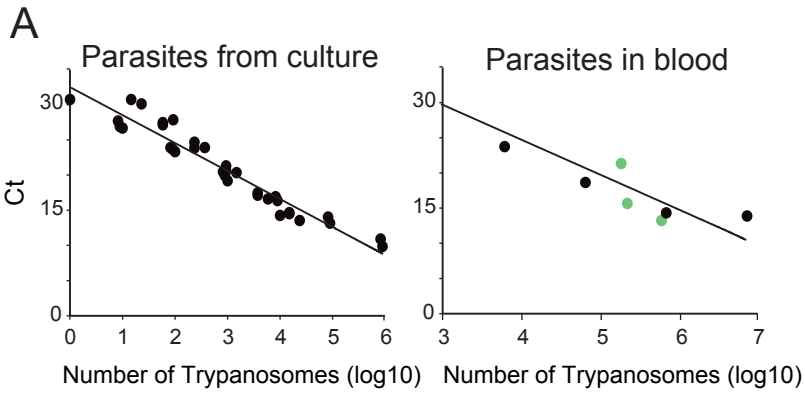
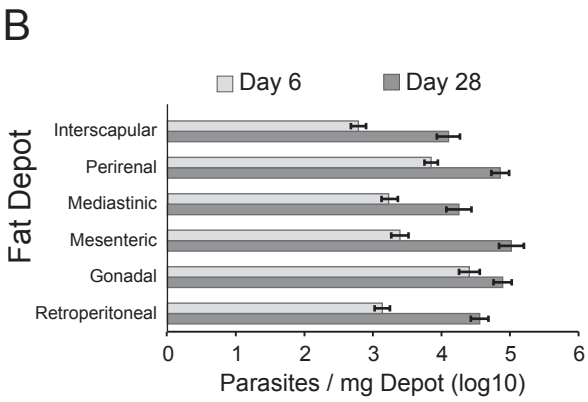
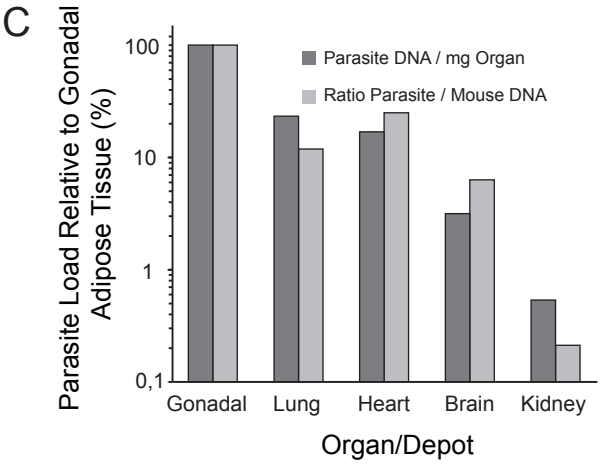


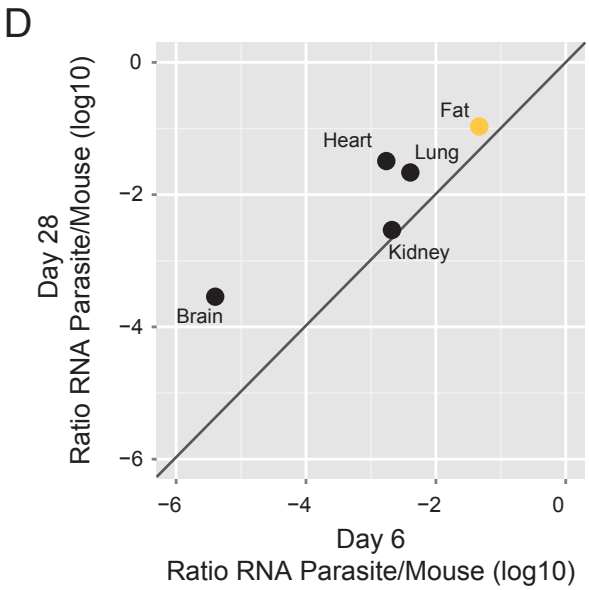
Figure S1. Related to Figure 1 and Figure 2. Clinical and histology details during *T. brucei* infection.

C57BL/6J mice were injected i.p. with 2000 AnTat1.1E parasites.
(A) Variation of food intake during infection. Animals (n = 15 per group) were group housed and food intake was measured daily by weighting the food and dividing by the number of mice per cage (n = 4 per condition). Light grey shaded area represents SEM.
(B) Representative brightfield micrographs of gonadal fat depots at different days post-infection, assessed by immunohistochemistry with anti-VSG antibody (parasites appear in brown). Original magnification, 400x.


A
Parasites from culture
Parasites in blood
Ct
30
15
0
0 1 2 3 4 5 6
3 4 5 6 7
Number of Trypanosomes (log10)
Number of Trypanosomes (log10)


B
Day 6
Day 28
Fat Depot
Interscapular
Perirenal
Mediastinic
Mesenteric
Gonadal
Retroperitoneal
0 1 2 3 4 5 6
Parasites / mg Depot (log10)


C
Parasite Load Relative to Gonadal Adipose Tissue (%)
100
10
1
0,1
Parasite DNA / mg Organ
Ratio Parasite / Mouse DNA
Gonadal
Lung
Heart
Brain
Kidney
Organ/Depot


D
Day 28
Ratio RNA Parasite/Mouse (log10)
0
−2
−4
−6
Fat
Heart
Lung
Kidney
Brain
−6 −4 −2 0
Day 6
Ratio RNA Parasite/Mouse (log10)

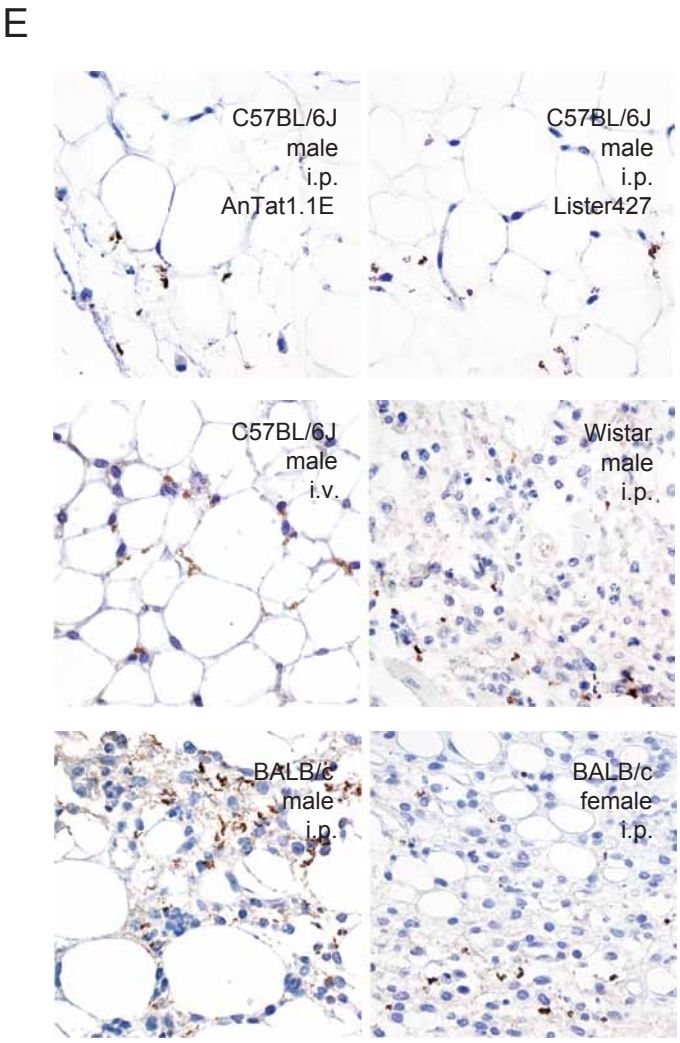

E
C57BL/6J male i.p. AnTat1.1E
C57BL/6J male i.p. Lister427
C57BL/6J male i.v.
Wistar male i.p.
BALB/c male i.p.
BALB/c female i.p.

Figure S2. Related to Figure 1 and 2. Further validation of preferential accumulation of parasites in adipose tissue.

(A) Calibration curve obtained from four independent *in vitro* cultures of known cell density. gDNA was extracted, serially diluted and amplified by quantitative PCR using *T. brucei* 18S rDNA primers. The goodness of fit of the linear regression is $R^2 = 0.933$. Calibration curve obtained from blood from infected mice (n = 3) (green dots) and from a culture of parasites diluted in blood from naive mice (n = 1) (black dots). The goodness of fit of the linear regression is $R^2 = 0.925$.
(B) Parasite density in six fat depots on day 6 (n = 6-12) and 28 (n = 3-6) post-infection determined by qPCR of gDNA. For each depot, significant differences were found between days 6 and 28 post infection (Student's unpaired t test, $P < 0.05$).
(C) Two different gDNA qPCR methods were compared on day 28 post-infection: i. number of parasites per mg of organ (n = 3 - 9) and ii. ratio between *T. brucei and* mouse 18S gDNA (n = 3 - 9). Both methods show a similar parasite density in different tissues (LME, $P = 0.72$).
(D) Parasite density on day 6 and 28 post-infection determined by qRT-PCR of RNA. Transcripts of the parasite *TbZFP3* gene were normalized to the mouse *Gapdh*. Each point represents the geometric mean of the parasite density on day 6 (n = 3 - 4) and on day 28 post-infection (n = 3 - 5). RNA quantification validates the conclusions taken from gDNA qPCR: the adipose tissue is the major parasite reservoir (LME, $P = 0.0006$). The relative contribution of other organs is similar to what was measured by gDNA, except for brain, in which parasite density was lower than expected. Perhaps *TbZFP3* is downregulated in this organ.
(E) Representative brightfield micrographs of *T. brucei* in gonadal adipose tissue in different models of infection, assessed by immunohistochemistry with anti-VSG antibody (parasites appear in brown). Original magnification, 400x.

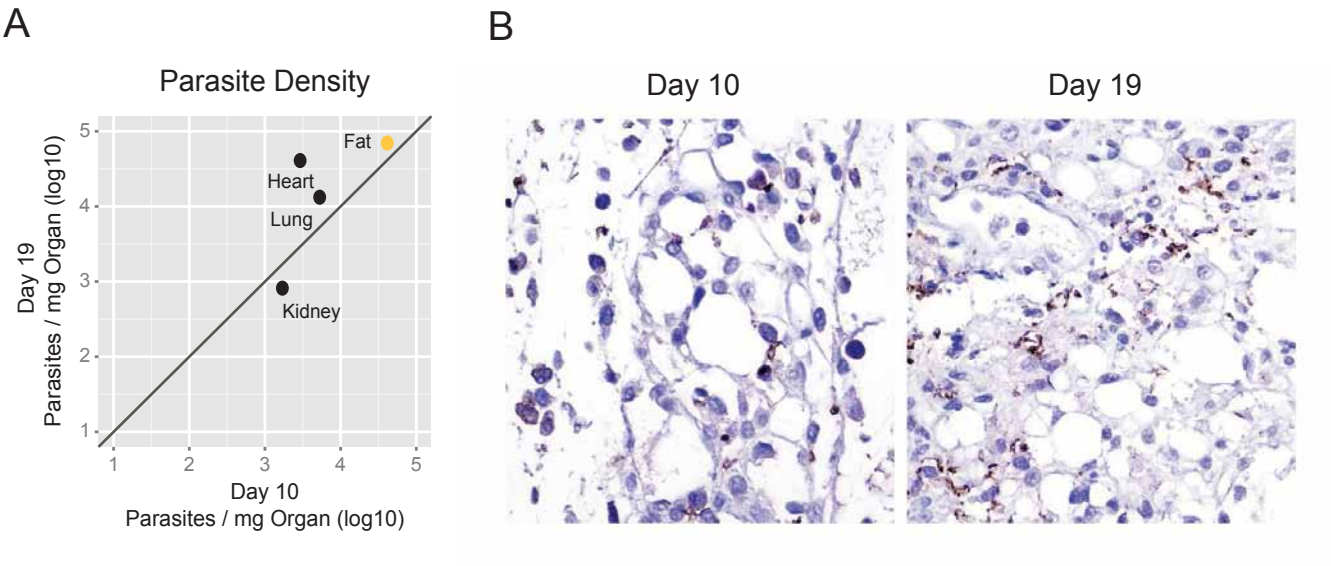


Figure S3. Related to Figure 2. Parasites accumulate in fat when infection is initiated by tsetse bite.

(A) Parasite density in multiple organs in mice naturally infected by the bite of a tsetse fly. Mice were sacrificed at the first and second peaks of parasitemia (10 and 19 days post-infection respectively) and parasite density determined by amplification of gDNA, as described in Figure 2. Each point represents the geometric mean of the parasite density at day 10 (n = 4) and at day 19 (n = 8).
(B) Representative brightfield micrographs of gonadal fat depot immunostained with anti-VSG serum, 10 and 19 days post-infection by tsetse bite. Original magnification, 40x.

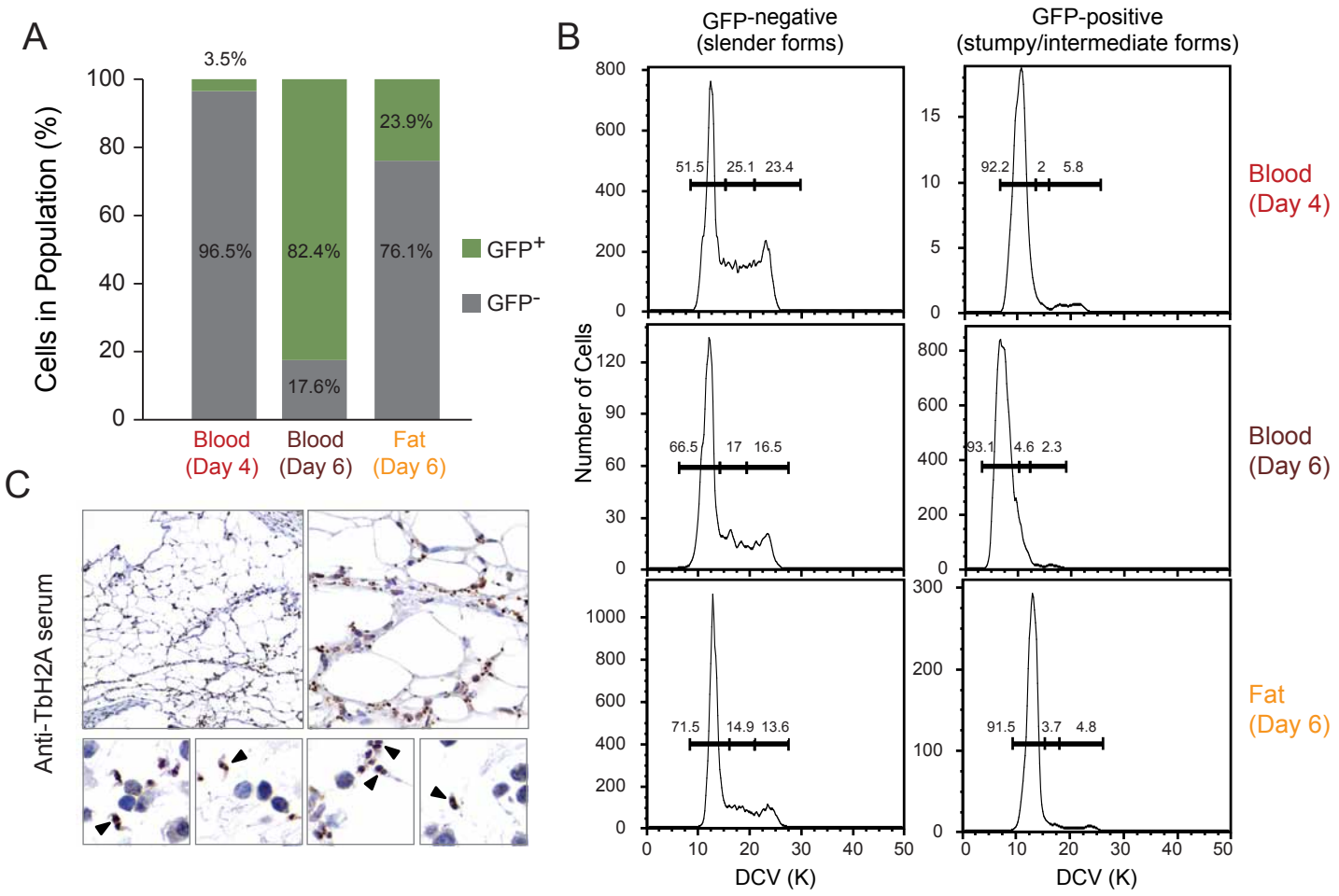


Figure S4. Related to Figure 3. Complementary methods to confirm presence of both replicative and cell-cycle arrested parasites in fat on day 6 post-infection.

(A) GFP Expression and (B) Cell cycle analysis in parasites isolated from blood and adipose tissue, 4 or 6 days post-infection, assayed by FACS. A *GFP::PAD1*$_{3UTR}$ *T. brucei* reporter cell-line was stained with DyeCycle Violet. The histograms represent the distribution and the percentage of parasites in each cell cycle stage. DyeCycle Violet staining validates the conclusions taken from propidium iodide staining analysis.

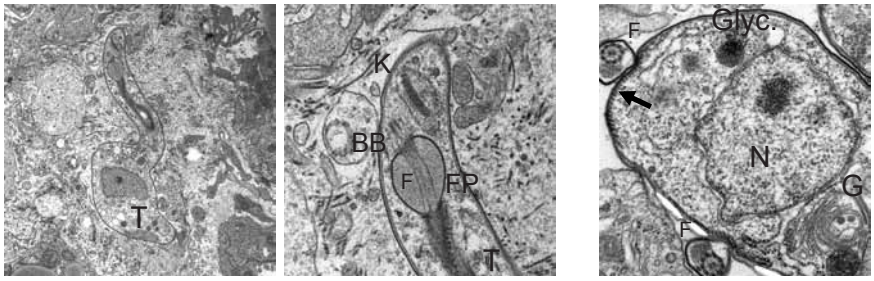
(C) Representative brightfield micrographs of gonadal fat depots on day 6 post-infection, assessed by immunohistochemistry with anti-*T. brucei* H2A rabbit serum (parasites appear in brown). Original magnification, 20x, 40x and 100x. Arrowheads indicate parasites undergoing nuclear division.

A

GFP expression	Tissue	Length (µm)	Width (µm)
GFP-negative (slenders)	Fat	24.57 ± 2.99	2.12± 0.26
	Blood	24.39 ± 2.50	2.15 ± 0.26
GFP-positive (stumpy/interm.)	Fat	21.32 ± 2.73	2.29 ± 0.31
	Blood	18.43 ± 1.81	3.11 ± 0.38

n= 100, Numbers are mean values with ± standard deviations .

B



C

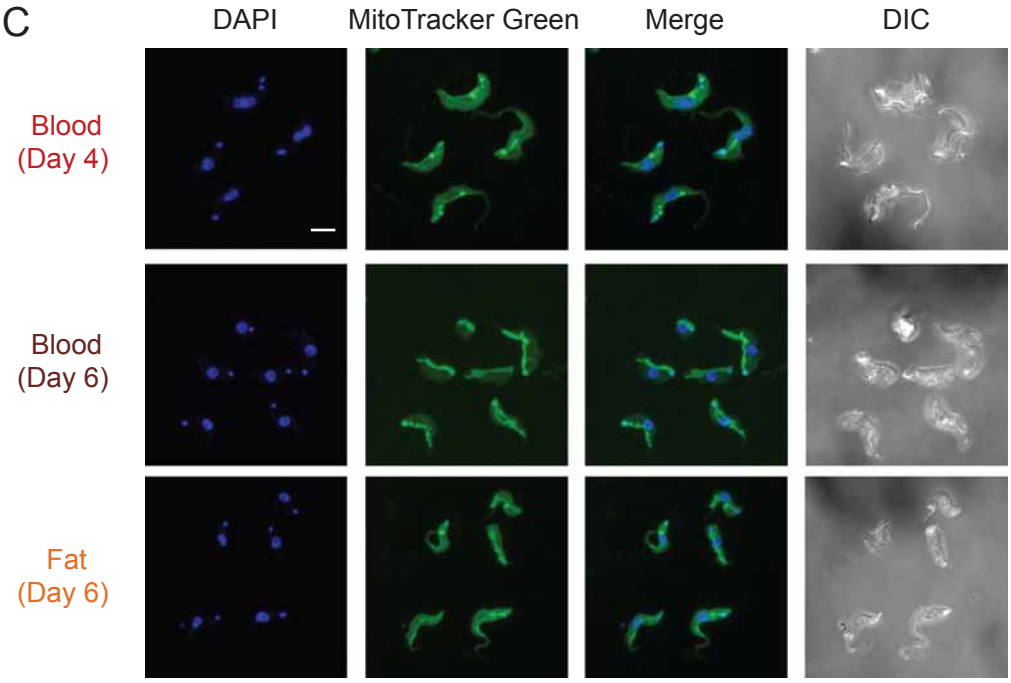


Figure S5. Related to Figure 4. Subcellular organization of parasites from adipose tissue.

(A) Length and width mean values of GFP-negative (slender) and GFP-positive (stumpy/intermediate) parasites isolated from blood or fat.
(B) Transmission electron micrographs of parasites in gonadal adipose tissue (day 28 post-infection).
T, trypanosome; K, kinetoplast; BB, basal body; F, flagellum; FP, flagellar pocket; N, nucleus; Glyc, Glycosomes; Arrow: subpellicular microtubules. Scale bars represent 2 and 0,5 µm in the left and right panels, respectively.
(C) MitoTracker Green, which stains in live cells the mitochondrion membrane, regardless of its membrane potential, was used to assess mitochondrion morphology. DNA was stained by DAPI and images were captured under a confocal microscope. Bloodstream form parasites from day 4 of infection showed a punctate pattern typical of mitochondrion in slender parasites, while on day 6, mitochondrion displayed a tubular structure with a few branches in what appeared to be stumpy forms. The mitochondrion of adipose tissue forms present a tubular structure, but with fewer branches and thinner. All panels are shown with the same magnification. Scale bar represents 5 µm.

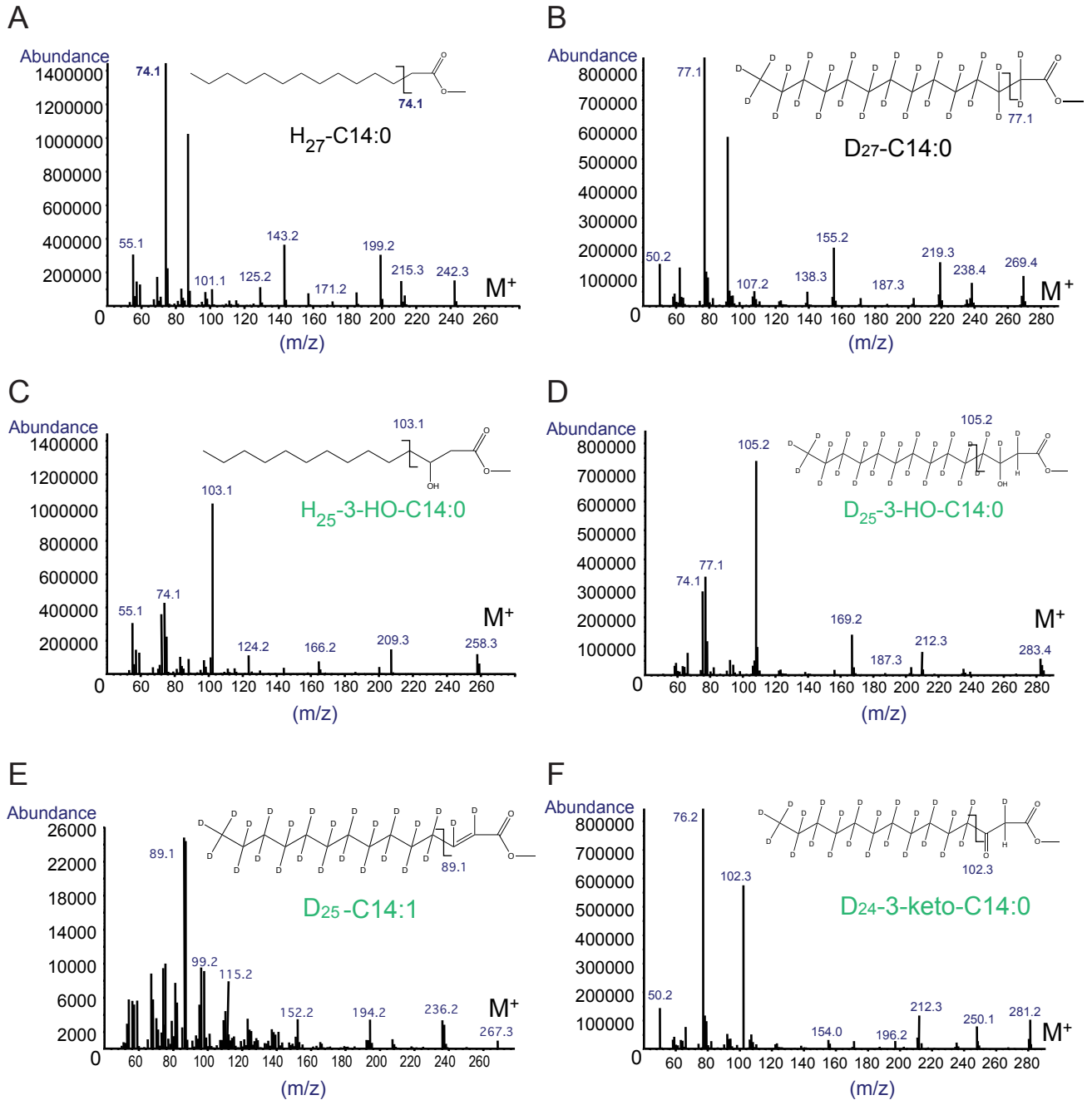


Figure S6. Related to Figure 6. Lipid metabolites identified by GC-MS.

All panels show the structure and fragmentation pattern of methyl ester derivatives of:

(A) myristic acid (C14:0)

(B) fully deuterated-myristic acid (D_{27}-C14:0)

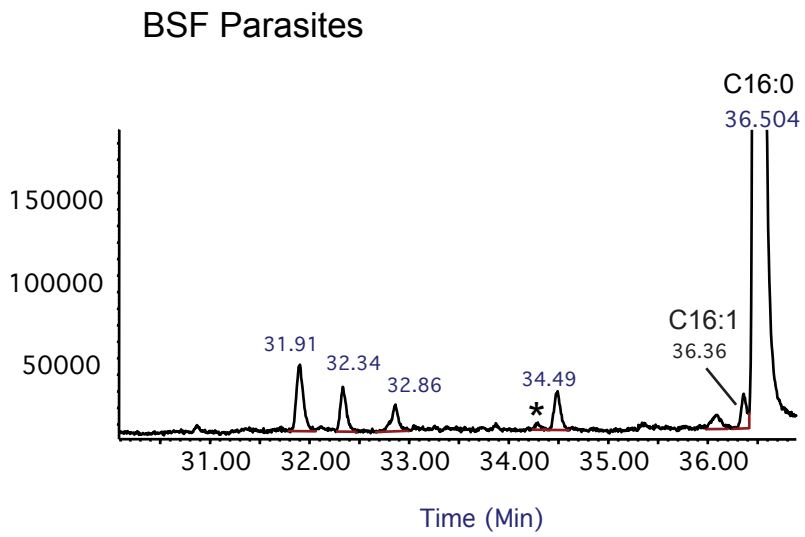
(C) 3-hydroxy- myristic acid (3-HO-C14:0)

(D) 3-hydroxy-deuterated- myristic acid (D_{24}-3-HO-C14:0)

(E) deuterated-myristoleic acid (D_{25}-C14:1)

(F) 3-keto-deuterated- myristic acid (D_{24}-3-keto-C14:0)

A



B

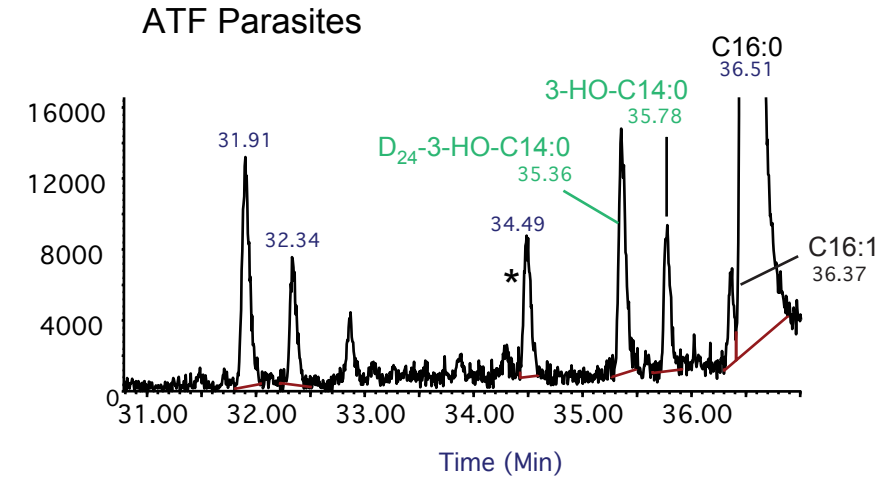


Figure S7. Related to Figure 6. Only ATF parasites produce hydroxyl-fatty acids as part of β-oxidation catabolism of fatty acids.
FAME analysis by GC-MS of D_{27}-Mys labeled and subsequently chased bloodstream (A) and adipose tissue (B) forms. Trace 31-37 minutes showing positions of (3-HO-C14:0) and (D_{24}-3-HO-C14:0) in adipose tissue forms only.

Supplemental Tables and Movies

Table S1, related to Figure 5. Mapping information of RNA-Seq reads in samples from blood and adipose tissue.

Sample	Total # reads	# of reads mapped to mouse genome	% of mouse reads in dataset	# of reads mapped to *T. brucei* genome	% of *T. brucei* reads in dataset
Blood-A	20 491 498	282 806	1,4%	16 004 171	78,1%
Blood-B	22 485 284	255 967	1,1%	18 418 295	81,9%
Fat-A	36 739 561	32 564 158	88,6%	3 396 819	9,2%
Fat-B	37 741 627	36 837 101	97,6%	415 936	1,1%
Fat-C	45 755 070	42 030 335	91,9%	3 311 276	7,2%

Table S2, related to Figure 5. Genes differentially expressed between parasites from adipose tissue and blood. (EXCEL file)

Table S3, related to Figure 5. Classes of transcripts differentially expressed between parasites in adipose tissue and blood. (EXCEL file)

Movie S1, related to Figure 4. Model constructed from hand-drawn contours marking the boundaries of cellular components in a tomogram of a trypanosome isolated from mouse gonadal adipose. The video is projected using Image J and various sub-cellular organelles are highlighted through subjectively attributed colors (plasma membrane in yellow; glycosomes in pink; nucleus in white; mitochondrion in green).

Supplemental Experimental Procedures

Ethical Statements

Animal experiments were performed according to EU regulations and approved by the Animal Ethics Committee of Instituto de Medicina Molecular (iMM) (AEC_2011_006_LF_TBrucei_IMM). The animal facility of iMM complies with the Portuguese law for the use of laboratory animals (Decreto-Lei 113/2013); and follows the European Directive 2010/63/EU and the FELASA (Federation of European Laboratory Animal Science Associations) guidelines and recommendations concerning laboratory animal welfare. The tsetse fly mediated *T. brucei* infection work was performed in compliance with the regulations for biosafety and animal ethics (VPU2014_1) and under approval from the Environmental administration of the Flemish government.

Cell Lines

The majority of the infections described in this manuscript were performed using *T. brucei* AnTat 1.1E, a pleomorphic clone derived from an EATRO1125 clone. AnTat 1.1E 90-13 is a transgenic cell-line encoding the tetracyclin repressor and T7 RNA polymerase (Engstler and Boshart, 2004). Tsetse infections were performed with *T. brucei* AnTAR1 strain. We also used *T. brucei* Lister 427, a monomorphic strain derived from antigenic type MiTat 1.2, clone 221a (Johnson and Cross, 1979). The stumpy reporter cell line *GFP::PAD1*$_{utr}$ derives from AnTat 1.1E 90-13 in which the green fluorescent protein, GFP, is coupled to PAD1 3'UTR. A nuclear localization signal (NLS) targets GFP protein into the nucleus. Prior to infection, *T. brucei* cryostabilates were thawed and parasite mobility was checked under an optic microscope.

Mice Infections

Mice were group-housed in filter-top cages and maintained in a Specific-Pathogen-Free barrier facility. The facility has standard laboratory conditions: 21 to 22°C ambient temperature and a 12 h light/12 h dark cycle. Chow and water were available *ad libitum*. Unless otherwise stated, all infections were performed in wild-type male C57BL/6J mice, 6-10 week old (Charles River, France), by intraperitoneally (i.p.) infection of 2000 *T. brucei* AnTat 1.1E 90-13 parasites.

To test if the presence of parasites in fat was dependent on the model, we performed the following variations in the infection protocol: a) male C57BL/6J mice were infected with a more virulent strain, the *T. b. brucei* strain Lister 427; b) male C57BL/6J mice were infected intravenously (i.v.) in the tail vein; c) a different strain of mice, BALB/c, and 7 week old Wistar rats (Charles River, France) were also infected (the latter with 4000 parasites); d) female C57BL/6J mice and finally e) male C57BL/6J mice naturally infected by tsetse bite. For this latter protocol, freshly emerged *Glossina morsitans morsitans* flies were fed their first blood meal on a *T. brucei* AnTAR1 infected mice at the peak of parasitemia. Subsequently, flies were maintained on commercially available defibrinated horse blood through *in vitro* membrane feeding. Thirty days after the infective blood meal, individual flies were evaluated for the presence of metacyclic trypanosomes in their salivary glands by salivation on pre-warmed (37°C) glass slides. To initiate a natural infection, one individual tsetse fly with a mature salivary gland infection was allowed to probe and feed per mouse.

Clinical Parameters and Organ Collection

All measurements in mice were made between 17:00 and 18:00. For parasite counts, blood samples were taken daily from the tail vein, and parasitemia was determined by manual counting using a Neubauer chamber. Organs/tissues of infected C57BL/6J mice (male and female) were collected at days 6, 13, 21 and 28 post-infection; for the infection with *T. b. brucei* strain Lister 427, organs were collected once parasitemia reached $1x10^8$ parasites/mL; for BALB/c mice, at day 6 post-infection; and for Wistar rats and C57BL/6J mice i.v. infections, at days 6, 13 and 20 post-infection. Animals were sacrificed by CO_2 narcosis, blood was collected by cardiac puncture and perfusion was performed to eliminate blood and parasites from circulation. Briefly, mice were perfused transcardially with pre-warmed heparinized saline (50 mL 1X phosphate buffered saline (1X PBS) with 250 μL of 5000 I.U./mL heparine per animal) using a peristaltic pump, ranging its speed from 2 mL/min to 8 mL/min. Organs were then collected and either used immediately for parasites isolation, snap frozen in liquid nitrogen for molecular analysis, or fixed in 10% neutral-buffered formalin for histopathology.

Transplantations

Donor mice were sacrificed with CO_2 narcosis (at days 21 and 28 post-infection) and the gonadal fat depot, brain, heart, and 600 µL of blood were harvested. Organs were manually homogenized through a 70 µm mesh into 1X PBS. Cell suspension was centrifuged at 1000 g for 10 minutes and resuspended in 800 µL of HMI11. Tissue lysates were injected intraperitoneally in naïve mice. Blood was diluted in 1X PBS and centrifuged for 5 minutes at 2800 g. Cell pellet was diluted in 800 µL of 1X PBS and injected intraperitoneally in naïve mice.

Histology

Formalin-fixed organs were embedded in paraffin and 3µm sections were stained with hematoxylin and eosin (H&E). For immunohistochemistry, 3µm sections were stained for VSG using non-purified rabbit sera anti-*T. brucei* VSG13 antigen (cross-reactive with most VSGs via the C-terminal domain) or anti-*T. brucei* H2A (generated against a recombinant protein) (kind gift of Christian Janzen), diluted 1:5000 and 1:3000, respectively. Briefly, antigen heat-retrieval was performed in a microwave oven (800 w) for 15 minutes with pH 9 Sodium Citrate buffer (Leica Biosystems, MO, USA). Incubation with ENVISION kit (Peroxidase/DAB detection system, Dako Corp, Santa Barbara, CA) was followed by Mayer's hemalumen counterstaining. No staining was observed in the negative control (without primary antibody). Tissue sections were examined by a pathologist (TC), blinded to experimental groups, in a Leica DM2500 microscope coupled to a Leica MC170 HD microscope camera.

For transmission electron microscopy, gonadal fat depot from infected mice (days 6 and 28 post-infection) was collected and fixed for three hours at 4ºC in 0.1 M cacodylate buffer, pH 7.4, containing 2.5% (v/v) glutaraldehyde. After staining for 1 hour with 1% (w/v) osmium tetroxide and 30 minutes with 1% (w/v) uranyl acetate, samples were dehydrated in an ethanol gradient (70-95-100%), transferred to propylene oxide and embedded in EPON™ resin. Semi-thin sections (300-400 nm) were stained with toluidine blue for light microscopy evaluation. Ultra-thin sections (70 nm) were collected in copper slot grids and stained with 2% uranyl acetate and lead citrate (Reynolds recipe). Grids were screened in a Hitachi H-7650 transmission electron microscope at 100 kV acceleration.

For 3D reconstruction of isolated trypanosomes, parasites isolated from gonadal fat depot were centrifuged at 5000 rpm and processed as described above for whole tissue. After embedding, approximately 26 serial ultra-thin sections (70 nm) were collected for each individual parasite. Grids of seven parasites were screened in a Hitachi H-7650 transmission electron microscope at 100 kV acceleration, serial section alignment was achieved using the IMOD software package version 4.7.3 for alignment and modeling (Kremer et al., 1996). ,Videos were projected using ImageJ 4.47v.

Microscopy analysis

Parasites from blood were isolated using a DEAE column (Taylor et al., 1974) and parasites from fat were isolated from gonadal fat depot by incubating the depot in HMI11 for up to 40 minutes and then purified from tissue debris over a DEAE column. For morphometric analysis, $1x10^6$ parasites were settled for 15 minutes to a pre-coated dish with 10% poly-Lysine, fixed with 2% paraformaldehyde for 10 minutes at room temperature, washed with 1X PBS and stained with diamidino-2-phenylindole (DAPI) (5 µg/mL in 1X PBS). Vectashield solution was used to mount the dish. Fluorescence and Phase Contrast images were acquired using Zeiss Cell Observer wide-field microscope. Parasite length was taken essentially as described in (Wheeler et al., 2012). Parasite width was manually scored in the nuclear region of the cell body.

For mitochondrial staining, isolated parasites were incubated in HMI-11 with 700 µM MitoTracker Green for 30 minutes at 37˚C. Excess of MitoTracker stain was removed and cells were chased in HMI-11 medium for 40 minutes at 37˚C. After washing with 1X PBS, cells were fixed and handled as above. Fluorescence and DIC images were acquired using a confocal Laser Point-Scanning Microscope (Zeiss LSM 710).

For tissue fluorescence analysis, the gonadal depot was stained with LipidTox (1:200 v/v in 1X PBS) for 30 minutes at 4ºc with agitation and then fixed in 10% neutral-buffered formalin, Sigma, and washed twice in 1 mL 1X PBS for 30 minutes at 37ºC with 150 rpm horizontal agitation. After, fat samples were embedded and mounted in Fluoromount-G. Fluorescence images were taken using a 40X objective in a Zeiss Cell Observer WF Microscope.

Parasite quantification in organs and tissues

Collected organs/tissues were snap frozen in liquid nitrogen. Genomic DNA (gDNA) was extracted from tissues using NZY tissue gDNA isolation kit (NZYTech, Portugal). The primers used for amplification of 18S rDNA gene of *T. brucei* were 5'–ACGGAATGGCACCACAAGAC–3' and 5'–GTCCGTTGACGGAATCAACC–3' and for the mouse 18S rDNA gene were 5'–TCGAGGCCCTGTAATTGGAA–3' and 5'–CTTTAATATACGCTATTGGAGCTGGAA–3'. By qPCR, the amount of *T. brucei* 18S rDNA present per milligram of organ/tissue were measured and converted into number of parasites using a calibration curve. The curve used to calculate parasite density in most solid organs/tissues was obtained by quantifying *T. brucei* 18S rDNA from serial dilutions of four independent cultures of AnTat1.1E 90-13 of known cell densities. Because PCR amplification from blood genomic DNA is sensitive to hemoglobin, we made a calibration curve specifically for blood samples, which was obtained by quantifying *T. brucei* 18S rDNA from blood from three infected mice of known parasitemia and from a known number of culture parasites diluted in blood. Parasite density in organs/tissues was calculated by dividing the number of parasites by the weight of organ/tissue used for qPCR, while total parasite load was calculated by multiplying the number of parasites/mg by the weight of organ/tissue at the corresponding day of infection. Blood density is about 1.05 kg/L, which means that 1 mL of blood is roughly equivalent 1050 mg. This conversion was taken into account to convert the number of parasites per mL of blood into number of parasites per mg of blood.

Parasite quantification from gDNA was also undertaken using ΔΔCt method. Here *T. brucei* and Mouse 18S rDNA genes were quantified from total genomic DNA. The ratio of the two genes provided a relative parasite density.

Given that trypanosomal gDNA can remain in circulation for up to 14 days, RNA qPCR was also used to quantify parasite load in tissues by calculating the ratio of *T. brucei TbZFP3* transcript to mouse *Gapdh* transcript. RNA was extracted using TRIzol or TRIzol LS reagents (Life Technologies) and cDNA prepared with TaqMan® Reverse Transcription Reagents (Invitrogen). The primers used for amplification of *TbZFP3* gene of *T. brucei* were 5'–CAGGGGAAACGCAAAACTAA–3' and 5'–TGTCACCCCAACTGCATTCT–3' and for mouse *Gapdh* gene were 5'–CAAGGAGTAAGAAACCCTGGACC–3' and 5'– CGAGTTGGGATAGGGCCTCT–3'. Quantitative PCR (qPCR) was performed on an ABI StepOnePlus real-time PCR machine and data was analyzed with the ABI StepOne software.

GFP Expression and Cell Cycle FACS Analysis

To quantify and characterize stumpy population in blood and fat, blood was collected by heart puncture from mice infected for 4 or 6 days, while gonadal fat depot was collected only on day 6 post infection (day 4 yield is too low to obtain enough isolated parasites). Parasites from blood were isolated using a DEAE column (Taylor et al., 1974) and parasites from fat were isolated from gonadal depot by incubating the depot in HMI11 for up to 40 minutes (hemocytometer was used to score the number of released parasites with time). GFP-expression in isolated parasites was analyzed on a BD LSRFortessa™ cell analyzer and data processed using FlowJo.

Cell-cycle analysis was performed using propidium iodide or Vybrant DyeCycle violet. For propidium-iodide, cells were first washed in ice-cold 1X PBS and then fixed with ice-cold 100% ethanol. After a washing step in 1X PBS, 0.5 μL of propidium iodide and 0.5 μL of RNaseA were added per each million of isolated parasites and incubated for 30 minutes at 37ºC. Vybrant DyeCycle violet staining was used in live cells. Essentially, cell suspensions from both blood and fat were washed once in 1X Trypanosome Dilution Buffer (1X TDB). 0.5 μL of DyeCycle violet was added per each million of isolated parasites and incubated for 10 minutes at 37ºC. Intensities were measured with BD LSRFortessa™ cell analyzer.

RNA-Sequencing

Blood from mice infected for 4 days was collected by heart puncture, parasites were purified over a DEAE column (Taylor et al., 1974) and total RNA extracted using TRIzol LS reagent. Mice infected for 6 days were sacrificed and perfused as explained above. Gonadal fat depot was collected and immediately lysed using TRIzol reagent. RNA and cDNA library preparation essentially followed the protocol in (Pena et al., 2014). Samples were sequenced in an Illumina HiSeq2000 platform (EMBL and BGI). Sequenced reads were 100bp paired-end for day 6 fat samples A and B, and 52bp single-end for the remaining samples. Paired-end reads were preprocessed by discarding the second read of each mate-pair, and trimming the first read to 52bp. Sequenced read quality was assessed using the FASTQC quality

control tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The SolexaQA suite of programs (http://www.biomedcentral.com/1471-2105/11/485) (Cox et al., 2010) was used to trim raw reads to their longest contiguous segment above a PHRED quality threshold of 28, and reads smaller than 25 nucleotides long were discarded. Reads were mapped to the *T. brucei* TREU927 (TriTrypDB 8.0) and *Mus musculus* (GRCm38) genomes using Tophat 2 (REF: http://www.genomebiology.com /2013/14/4/R36/abstract) (Kim et al., 2013) with library type unstranded and the total number of reads mapped to each organism was assessed (Table S1). The number of reads mapped to each CDS was quantified in the R software environment using the function summarizeOverlaps from the GenomicAlignments (1.2.2). We found that approximately 0.1% of the reads could not be unambiguously assigned to the mouse or Trypanosome genome. To avoid a possible contamination of the Trypanosome differential expression results with reads originating from mouse RNA, all reads that mapped, uniquely or not, to the mouse genome were discarded for further analysis. Differential gene expression was analysed using DESeq2 (v1.6.3) (Anders and Huber, 2010), edgeR (v3.8.6) (Robinson et al., 2010) and baySeq (2.0.50) (Hardcastle and Kelly, 2010) from Bioconductor (v3.0). Genes were considered differentially expressed if they were detected by at least two of these three algorithms (P adjusted < 0.01). The ArrayExpress accession number for Lister427 culture parasites is E-MTAB-1715. Sequence data generated as part of this study have been submitted to the ArrayExpress database (EMBL-EBI) under accession number E-MTAB-4061.

Myristate Metabolic Labeling

To evaluate myristate incorporation and degradation by *T. brucei*, blood parasites were isolated using a DEAE column (Taylor et al., 1974) and fat parasites were isolated from gonadal depot by incubating it in lipid free 1X Minimum Essential Medium (MEM) for up to 40 minutes (hemocytometer was used to score the number of released parasites with time). MEM was chosen instead of HMI11 to ensure no external stimulus would be given to the parasites. After isolation, parasites were washed in MEM, resuspended in 1 mL of the same medium, placed in an open/vented tube and starved of fatty acids for 30 minutes while in a water bath at 37ºC with 100 rpm horizontal agitation. To the starved parasites, 200 µL of labeled myristate (D_{27}-myristic acid, CDN) pre-coupled with fatty acids free BSA were added. This solution was obtained by adding $2x10^{-3}$g of labeled myristate dissolved in 40 µL 100% ethanol (that was left at room temperature for 20 minutes) to $2x10^{-3}$g of fatty acids free BSA dissolved in 960 µL of MEM (that was left at 60ºC for 20 minutes), after a short spin the supernatant was ready for use. Parasites were once again incubated in a water bath at 37ºC with 100 rpm horizontal agitation for a one hour period (pulse). After pulse, 450 µL of the cell suspension were washed in 1 mL 1X TDB, snap frozen in liquid nitrogen and lyophilized in glass vials. The remaining parasites were centrifuged at 800 g for 5 minutes and re-suspended in 500 µL of MEM and 100 µL of HMI11. This cell suspension was incubated in a water bath at 37ºC with 100 rpm horizontal agitation for one hour (chase). Just as after pulse, cells were washed in 1X TDB and snap frozen in liquid nitrogen and lyophilized in glass vials. After, samples were spiked with an internal standard fatty acid C17:0, Sigma (20 µL of 1 mM) and dried under nitrogen. Acid hydrolysis was conducted using constant boiling HCl 6 M (200 µL) vortexing/sonication followed by incubation for 16 hours at 110 ºC. After cooling, the samples were evaporated to dryness in a speedvac concentrator and dried twice more from 200 µL of methanol:water (1:1). The protonated fatty acids were extracted by partitioning between 500 µL of HCl 20 mM and 500 µL of ether, the aqueous phase was re-extracted with 500 µL of fresh ether and the combined ether phases were dried under nitrogen in a glass vial. These fatty acids were then converted to methyl esters (FAME), by adding diazomethane (3 x 20 µL aliquots) to the dried residue, while on ice. After 30 minutes samples were allowed to warm to room temperature and left to evaporate to dryness in a fume hood. The FAME products were dissolved in 10-20 µL of dichloromethane and 1-2 µL analyzed by GC-MS on a Agilent Technologies (GC-6890N, MS detector-5973) with a ZB-5 column (30 M x 25 mm x 25 mm, Phenomenex), with a temperature program of at 70 ºC for 10 minutes followed by a gradient to 220 ºC at 5 ºC /minute and held at 220ºC for a further 15 minutes. Mass spectra were acquired from 50-550 amu. The FAME and the β-oxidation metabolites were identified based upon retention time and their fragmentation spectra.

Statistical Analysis

Statistical analyses were performed in the free software R: http://www.r-project.org. Statistically significant variation of body weight, food intake between infected and control mice and parasite density in multiple fat depots were determined by using Student's unpaired t test. Statistically significant variation of parasite morphology was determined by Wilcoxon rank sum test. To define the uniformity of the outcome of the two different parasites quantification methods we fitted a Linear Mixed Effects Model

(LME), using the nlme package in R, considering mice as random factors. The bias in parasite density and parasite load for fat relative to non-fat organs was also determined by fitting a linear mixed effect model, using the nlme package in R, considering mice as random factors. For the statistical analysis of RNA-Seq data, see RNA-Sequencing Section above.

Supplemental References

Cox, M. P., Peterson, D. A., and Biggs, P. J. (2010). SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics *11*, 485.

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

Hardcastle, T. J., and Kelly, K. A. (2010). baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. BMC Bioinformatics *11*, 422.

Johnson, J. G., and Cross, G. A. M. (1979). Selective cleavage of variant surface glycoproteins from *Trypanosoma brucei*. Biochem J *178*, 689-697.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Taylor, A. E., Lanham, S. M., and Williams, J. E. (1974). Influence of methods of preparation on the infectivity, agglutination, activity, and ultrastructure of bloodstream trypanosomes. Exp Parasitol *35*, 196-208.